Network Structural Balance Based on Evolutionary Multiobjective Optimization: A Two-Step Approach

Qing Cai, Maoguo Gong, *Senior Member, IEEE*, Shasha Ruan, Qiguang Miao, and Haifeng Du

***Abstract*—Research on network structural balance has been of great concern to scholars from diverse fields. In this paper, a two-step approach is proposed for the first time to address the network structural balance problem. The proposed approach involves evolutionary multiobjective optimization, followed by model selection. In the first step, an improved version of the multiobjective discrete particle swarm optimization framework developed in our previous work is suggested. The suggested framework is then employed to implement network multiresolution clustering. In the second step, a problem-specific model selection strategy is devised to select the best Pareto solution (PS) from the Pareto front produced by the first step. The best PS is then decoded into the corresponding network community structure. Based on the discovered community structure, imbalanced edges are determined. Afterward, imbalanced edges are flipped so as to make the network structurally balanced. Extensive experiments on synthetic and real-world signed networks demonstrate the effectiveness of the proposed approach.**

***Index Terms*—Community structure, evolutionary algorithm (EA), multiobjective particle swarm optimization, signed network, structural balance.**

I. Introduction

NETWORKS are on our minds at the present time [1]. The current interest in networks is part of a broader movement toward research on complex systems [2], [3]. With the development of science and technology, portable electronic devices are ubiquitous. People can build their social relations with the convenience provided by social platforms such as Facebook, Twitter, and WeChat. Social networks are good examples for modeling large-scale social systems. The links between connected users on the social networks give rise to a complex, multidimensional Web of aggregated social behavior [4], [5]. The research on networks may provide new insight into the study of complex systems.

Because structure always affects function, consequently, a substantial volume of work has been done to analyze the structural properties of networks [6]–[9]. Networks have many notable properties, such as the small-world property [10], the scale-free property [11], the community structure property [12], and the structural (or social) balance property [13], because in many interesting real-world social systems, the links between two objects usually display diversity [14]. For example, the relations between two persons may be friendly or antagonistic. Signed social networks, or signed networks for short, whose links are usually labeled with different signs, are an effective paradigm for depicting those social systems. We can get to know how a social system evolves and how it can reach a stable or balanced status by analyzing the properties of the corresponding signed network. Because local balancing influences global structure in social networks [15], the structural balance theory [16] provides us a macroscopic as well as a microscopic view to discover and analyze the features of signed networks.

Structural balance is regarded as a fundamental social process. It illustrates local effects impacting global properties. The structural balance theory [13] reflects the dynamic balance property of signed social networks. This theory can assist the decision makers to observe the evolving process of social systems and can provide valuable suggestions to help to make social systems reach a relatively stable yet harmonious status. The structural balance theory was originated from [17]. Later, it was generalized by Cartwright and Harary [16] and Harary [18], [19] from the perspective of graph theory. Though, there have been more than 50 years since Heider's seminal studies, the research of structural balance now has attracted scholars from diverse fields, such as computer science, sociology, economics, psychology, etc. Substantial amount of creative work has been done to perfect the balance theory and to compute the balance of signed networks. A recent survey on structural balance in signed networks can be found in [20].

Heider's balance theory suggests that a balanced complete graph can be divided into two subsets. However, many real-social networks cannot be represented by complete graphs. Moreover, studies in [21] and [22] suggest that many signed social networks can be divided into more than two clusters. Following these studies, many methods [23]–[27] based on clustering techniques have been devised for realizing

Manuscript received August 25, 2014; revised December 14, 2014 and February 11, 2015; accepted April 12, 2015. Date of publication April 17, 2015; date of current version November 25, 2015. This work was supported in part by the National Natural Science Foundation of China under Grants 61273317, 61422209, and 61473215; in part by the National Program for Support of Top-notch Young Professionals of China; in part by the Specialized Research Fund for the Doctoral Program of Higher Education under Grant 20130203110011; and in part by the Fundamental Research Fund for the Central Universities under Grant K5051202053.

Q. Cai, M. Gong, and S. Ruan are with Key Laboratory of Intelligent Perception and Image Understanding of Ministry of Education, International Research Center for Intelligent Perception and Computation, Xidian University, Xi'an 710071, China (e-mail: gong@ieee.org).

Q. Miao is with the School of Computer Science and Technology, Xidian University, Xi'an 710071, China.

H. Du is with the Center for Administration and Complexity Science, Xi'an Jiaotong University, Xi'an 710049, China (e-mail: haifengdu@mail.xjtu.edu.cn).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TEVC.2015.2424081

structural balance. These methods aim to find a good network partition based on balance evaluation criteria. It should be noted that structural balance can be modeled as optimization problems. Evolutionary algorithms (EAs) [28], [29], inspired by principles from biology, ethology, etc., have been recognized as outstanding techniques for solving diverse kinds of complex optimization problems in many realms. However, to the best of our knowledge, the ideas of EAs have seldom been utilized to address the structural balance problem.

In this paper, the idea of a multiobjective evolutionary algorithm (MOEA) has been adopted to study the structural balance problem for the first time. Consequently, we have suggested a two-step approach for addressing the structural balance problem. The proposed two-step approach involves MOEA [30], [31] and multicriteria decision making [32]. The highlights of the proposed approach are as follows.

1) In the first step of the proposed approach, an improved version of the multiobjective discrete particle swarm optimization (MODPSO) framework suggested in [33] is devised to conduct the multiresolution network clustering. Each single run of the algorithm can yield a set of Pareto solutions (PSs), which correspond to different network community structures. Experiments demonstrate that the improvements made to MODPSO are effective.
2) In the second step, a problem-specific model selection strategy is devised to choose the ultimate solution from the Pareto front (PF) yielded by the first step. A network partition is then obtained. Afterward, we determine and flip all the imbalanced edges so as to make the network structurally balanced. Experiments demonstrate that the proposed approach is promising for addressing the structural balance problem.
3) The proposed approach actually can provide a decision maker with many choices to view and solve network structural balance problem. Moreover, the two-step idea can be easily extended to handle other network issues.

The rest of this paper is organized as follows. Section II gives some related backgrounds including the fundamentals of signed network, the network structural balance theory, and evolutionary multiobjective optimization. Section III states our motivation in detail. The proposed two-step approach for addressing the structural balance problem is presented in Section IV. Section V exhibits the experimental studies. Finally, the conclusion is drawn in Section VI.

II. Preliminary

A. Signed Network Notation

A signed network is commonly modeled by a signed graph that is composed of a set of vertices and edges, as shown in Fig. 1. Usually, we use a full (or dashed) line to represent the friendly (or antagonistic) relation between its two endpoints.

Given a graph denoted by $G = \{V, \text{PE}, \text{NE}\}$, where V is the aggregation of vertices, and PE (or NE) is the aggregation

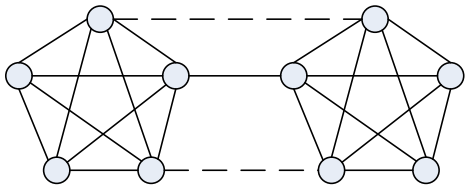

Fig. 1. Graph representation of a signed network. Solid line represents friendship and dashed line denotes antagonism.

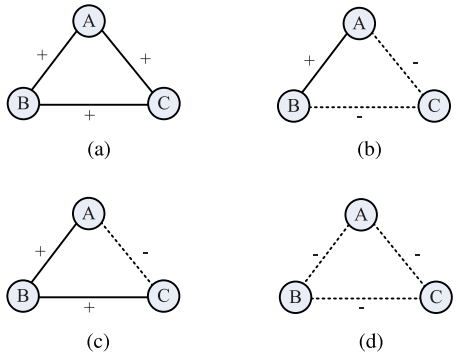

Fig. 2. Four configurations of signed triangles. (a) A, B, and C are mutual friends. (b) A and B are good friends, and they have the mutual enemy C. (c) A and C do not get along with each other, however, they have the same good friend B. (d) A, B, and C are mutual enemies.

of positive (or negative) edges. Thus, the linkage relations of graph G can be reflected by the adjacency matrix A whose element a_{ij} is normally defined as

$$a_{ij} = \begin{cases} \omega_{ij},\ \omega_{ij} > 0 \text{ if } e_{ij} \in \text{PE} \\ \omega_{ij},\ \omega_{ij} < 0 \text{ if } e_{ij} \in \text{NE} \\ \omega_{ij},\ \omega_{ij} = 0 \text{ if } \nexists e_{ij} \end{cases} \tag{1}$$

where ω_{ij} denotes the weight of edge e_{ij}.

Generally, matrix A is symmetric with the diagonal elements 0, but, if the corresponding network is directed, like the e-mail network, A is asymmetric.

B. Structural Balance Theory

Balance theory has a rich and long history. We would like to illustrate the theory with simple words rather than recount the history here. If we label a positive edge with a plus "+" and a negative edge with a minus "−," then for a simple graph with three nodes, there are four basic triadic configurations,[1] as shown in Fig. 2.

According to the structural balance theory first raised by Heider [13], [17], the triads in the first row in Fig. 2 are balanced while those in the second row are unbalanced, because a balanced triad fulfills the adage.

1) A friend of my friend is my friend.
2) An enemy of my enemy is my friend.
3) A friend of my enemy is my enemy.
4) An enemy of my friend is my enemy.

Heider's balance theory provides us with two equivalent ways to view structural balance of complete graphs, i.e., the

[1]Here, we only consider undirected graphs. With regard to the structural balance under directed graph context, please refer to [34].

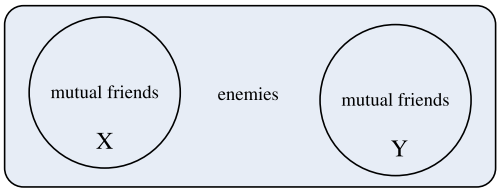

Fig. 3. Structural balance condition of a complete graph.

local view and the global view. From the local view, a complete network is claimed to be balanced if each constituent triangle is balanced as shown in Fig. 2(a) and (b). From the global view, a complete network is balanced iff the nodes of the network can be divided into two subsets [16], [35], as shown in Fig. 3, and each pair of nodes in the same subset has a positive link, whereas, between the two subsets a negative link.

Heider's balance theory does not provide a model for reasoning about the situations in which a network is noncomplete and can be divided into more than two subsets. Inspired by Heider's theory, researchers have extended Heider's theory to arbitrary networks. From a local view, a (noncomplete) graph is balanced if it can be completed by filling edges to form a complete graph that is balanced. From a global view, a graph is claimed to be balanced if it is possible to divide the nodes into two sets X and Y as shown in Fig. 3. Moreover, a signed network is said to be weakly balanced if the nodes can be divided into more than two sets such that every two nodes belonging to the same set are friends, and to different sets are enemies [36].

C. Evolutionary Multiobjective Optimization

Optimization has long been an active research area. Many real-world optimization problems involve multiple objectives. A multiobjective optimization problem can be mathematically formulated as[2]

$$\min\ F(x) = (f_1(x), f_2(x), \ldots, f_k(x))^T \tag{2}$$

which subjects to $x = (x_1, x_2, \ldots, x_n) \in \Phi$, where x is called the decision vector, and Φ is the feasible region in decision space.

The objectives in (2) often conflict with one another. Improvement of one objective may lead to deterioration of another. Thus, a single solution, which can optimize all objectives simultaneously, does not exist. For multiobjective optimization problems, the aim is to find good compromises (tradeoffs) which are also called Pareto optimal solutions. To understand the concept, here are some related definitions.

Definition 1 (Pareto Optimality): A point $x^* \in \Phi$ is Pareto optimal if for every $x \in \Phi$ and $I = \{1, 2, \ldots, k\}$ either $\forall i \in I$, $f_i(x) = f_i(x^*)$ or, there is at least one $i \in I$ such that $f_i(x) > f_i(x^*)$.

Definition 2 (Pareto Dominance): Given two points $x, y \in \Phi$, we say that x dominates y (denoted as $x \prec y$), iff

[2]Without loss of generality, we will assume only minimization problems.

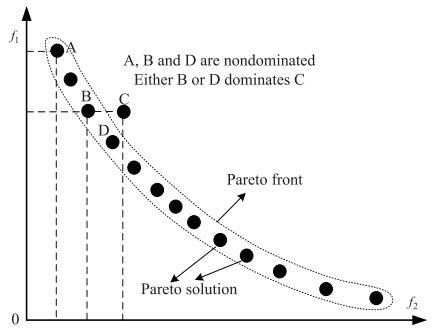

Fig. 4. Graphical illustration of Pareto optimal solution and PF.

$f_i(x) \leq f_i(y)$ for all $i \in I$, and $f_j(x) < f_j(y)$ for at least one $j \in I$. x is nondominated with respect to Φ, if there does not exist another $x' \in \Phi$ such that $x' \prec x$.

Definition 3 (Pareto Optimal Set): The set of all Pareto optimal solutions is called Pareto optimal set which is defined as

$$\text{PS} = \{x \in \Phi | \neg \exists x^* \in \Phi,\ x^* \prec x\}. \tag{3}$$

Definition 4 (Pareto Front): The image of the PS in the objective space is called the PF which is defined as

$$\text{PF} = \{F(x) | x \in \text{PS}\}. \tag{4}$$

Fig. 4 gives an example of the above mentioned definitions. Each dot except that labeled by C in the figure represents a nondominated solution to the optimization problem. The aim of a multiobjective optimization algorithm is to find the set of those nondominated solutions approximating the true PF.

In the last few years, much effort has been devoted to the application of EAs to the development of multiobjective optimization algorithms. Many outstanding MOEAs have been proposed [37]–[46].

III. Motivation

A. Importance of Structural Balance

Structural balance has its important place in network data analysis [47], [48]. Antal *et al.* [35] studied structural balance in international relations. Fig. 5 shows the network data that represent the evolution of alliances in Europe during World War I.

Fig. 5 clearly shows how structural balance penetrates international relations. Structural balance can assist intelligent decision making. The idea of structural balance is relevant to a large area of study.

Because many real-world social networks cannot be represented as complete graphs, and besides, sociometric findings suggest that a social network often can be divided into more than two cliques. The global balance theory, i.e., Heider's theory, provides scholars with an outlet to analyze the balance property of arbitrary signed networks.

B. Limitations of Traditional Methods

In the extant literature, the idea of traditional ways, including graph partitioning, hierarchical clustering, and spectral

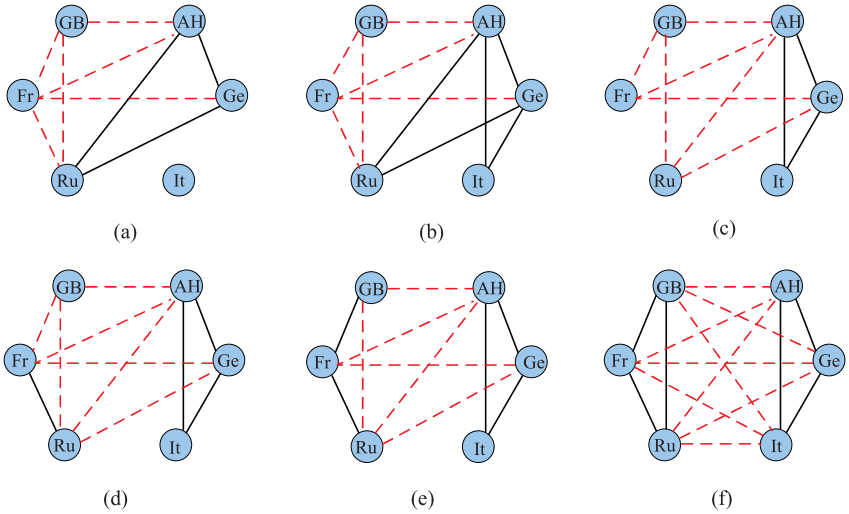

Fig. 5. Evolution of alliances in Europe, 1872–1907. The nations GB, Fr, Ru, It, Ge, and AH are Great Britain, France, Russia, Italy, Germany, and Austria-Hungary, respectively. (a) Three Emperors' league 1872–1881. (b) Triple alliance 1882. (c) German-Russian lapse 1890. (d) French-Russian alliance 1891–1894. (e) Entente Cordiale 1904. (f) British-Russian alliance 1907.

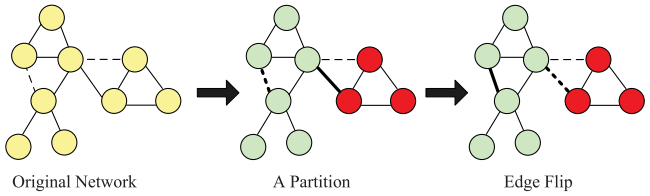

Fig. 6. Traditional way to realize the structural balance for an arbitrary network.

clustering [23]–[27], to realize the structural balance for an arbitrary network is shown in Fig. 6.

In Fig. 6, a technique is utilized to divide the network into small clusters, afterward, all the imbalanced edges[3] have been flipped so as to make the network balanced.

However, there are two drawbacks for the traditional methods. For one thing, it may need to set the clusters in advance. For another thing, each single run of the method can only output one solution. It should be pointed out that many social networks have hierarchical structures, i.e., a network can be divided into diverse kind of partitions based on different standards. Thus, one should have plenty of choices to realize the balanced structure of a network. However, each run of the traditional methods can only provide a single choice.

C. Our Two-Step Idea

To avoid the drawbacks of traditional methods, in this paper we suggest a two-step approach. The basic idea of the proposed approach is graphically illustrated in Fig. 7.

It can be seen from Fig. 7 that in the first step we adopt MOEA to implement multiresolution network community detection. Because each PS represents a certain network

[3]Here, we use the term imbalanced edges to denote two types of edges, i.e., the negative edges within a mutual friends group, the positive edges between different groups.

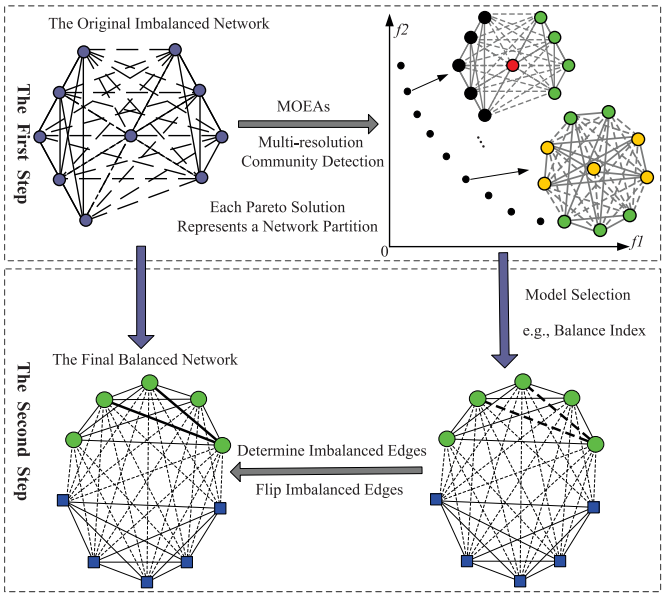

Fig. 7. Graphical illustration of our idea for solving the network structural balance problem.

partition, thus, each single run of the first step will produce a set of different network partitions, which will benefit a decision maker with great convenience. A model selection strategy is developed in the second step so as to recommend to the decision maker a best network partition with respect to the balance property. Afterward, imbalanced edges are determined and changed such that the original network becomes balanced.

We may notice from Fig. 7 that, compared with traditional methods, the proposed two-step approach actually can give many choices to realize the balance structure of a network. Based on different purposes, one can redesign the second step so as to obtain a preferable network balance structure with desired community topology.

Algorithm 1 General Framework of the Proposed Two-Step Approach for Addressing Network Structural Balance Problem

Parameters: maximum algorithm iterations: *gmax,* particle swarm size: *pop*, turbulence probability *pm*, inertia weight: ω, the learning factors: c_1, c_2.
Input: The adjacency matrix A of an imbalanced network G.
Output: Pareto solutions $\boldsymbol{X}^*$, each solution corresponds to a partition of a signed network; the balanced network G^*.
The First Step: multi-resolution network clustering.

1.1 Population Initialization:
  a) Position initialization: $\boldsymbol{P} = \{\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_{pop}\}^T$, where $\boldsymbol{x}_i$ is the position vector of the ith particle.
  b) Velocity initialization: $\boldsymbol{V} = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_{pop}\}^T$, where $\boldsymbol{v}_i$ is the velocity vector of the ith particle.
  c) Generate a uniformly distributed weighted vectors: $\boldsymbol{W} = \{\boldsymbol{w}_1, \boldsymbol{w}_2, \ldots, \boldsymbol{w}_{pop}\}^T$.
  d) Personal best solution initialization: $\boldsymbol{Pbest} = \{\boldsymbol{pb}_1, \boldsymbol{pb}_2, \ldots, \boldsymbol{pb}_{pop}\}^T$, where $\boldsymbol{pb}_i = \boldsymbol{x}_i$.
1.2 Initialize reference point $\boldsymbol{z}^*$.
1.3 Initialize neighborhood $B(i)$ based on Euclidean distance.
1.4 Population Update:
  a) Set $iter = 0$.
  b) for $i = 1, 2, \ldots, popsize$, do
    i) Randomly select one particle from the T neighbors as the leader.
    ii) Calculate the new velocity $\boldsymbol{v}_i^{t+1}$ and the new position $\boldsymbol{x}_i^{t+1}$ for the ith particle. See our previous work in [33].
    iii) If $iter < gmax \cdot pm$, implement turbulence operation on $\boldsymbol{x}_i^{t+1}$.
    iv) Evaluate $\boldsymbol{x}_i^{t+1}$.
    v) Update neighborhood solutions with $\boldsymbol{x}_i^{t+1}$. See Algorithm 2 for more information.
    vi) Update reference point $\boldsymbol{z}^*$.
    vii) Update personal best solution $\boldsymbol{pb}_i$.
  c) If $iter < gmax$, then $iter++$ and go to 1.4b.

The Second Step: structural balance transformation.

2.1 Select the ultimate Pareto solution BS from the PF yielded by the first step, using the proposed model selection strategy depicted in subsection IV-B.
2.2 Decode BS to a network partition $\Omega = (c_1, c_2, \ldots, c_k)$.
2.3 Determine the negative edges in c_i, $i = 1, \ldots, k$.
2.4 Determine the positive edges between c_i and c_j, $i \neq j$.
2.5 Flip all the imbalanced edges.

IV. Methodology

A. General Framework

In this paper, we have improved the MODPSO algorithm previously proposed in [33]. In the first step of the proposed approach, we employ the improved MODPSO to conduct multiresolution network clustering. In the second step of the proposed approach, a problem-specific PS selection strategy is devised. After obtaining the final network partition, we determine and flip the imbalanced edges. The framework of the proposed two-step approach for addressing network structural balance problem is presented in Algorithm 1.

Algorithm 2 Pseudocodes of the Improved Subproblems Update Strategy Suggested in This Paper

1) Set $P = \{1, \ldots, pop\}$. Set $c = 0$.
2) For each index $j \in P$, if $g^{te}(\boldsymbol{x}_i^{t+1}|\boldsymbol{w}_j, \boldsymbol{z}^*) \leq g^{te}(\boldsymbol{x}_j^{t+1}|\boldsymbol{w}_j, \boldsymbol{z}^*)$, then $c = c + 1$. //$g^{te}(\cdot)$ is the Tchebycheff decomposition approach used in [41].
3) **if** $c \leq n_r$ **then**
  a) For each index $j \in P$, if $g^{te}(\boldsymbol{x}_i^{t+1}|\boldsymbol{w}_j, \boldsymbol{z}^*) \leq g^{te}(\boldsymbol{x}_j^{t+1}|\boldsymbol{w}_j, \boldsymbol{z}^*)$, then set $\boldsymbol{x}_j^{t+1} = \boldsymbol{x}_i^{t+1}$ and $F(\boldsymbol{x}_j^{t+1}) = F(\boldsymbol{x}_i^{t+1})$.
4) **else**
  - **for** indexes $j \in P$ with $g^{te}(\boldsymbol{x}_i^{t+1}|\boldsymbol{w}_j, \boldsymbol{z}^*) \leq g^{te}(\boldsymbol{x}_j^{t+1}|\boldsymbol{w}_j, \boldsymbol{z}^*)$ **do**
    - Rank the Euclidean distances between $F(\boldsymbol{x}_i^{t+1})$ and $F(\boldsymbol{x}_j^{t+1})$ with an ascend order.
    - Choose the former n_r solutions, then set $\boldsymbol{x}_j^{t+1} = \boldsymbol{x}_i^{t+1}$ and $F(\boldsymbol{x}_j^{t+1}) = F(\boldsymbol{x}_i^{t+1})$.
  - **end for**
5) **end if**

It should be pointed out that, in Algorithm 1, the initialization step, the two objective functions, i.e., signed ratio association (SRA) and signed ratio cut (SRC), and the particle status update principles are the same as those in [33]. Small change has been made to the turbulence operation used in step 1.4. In our previous work, we assign the label of a vertex to all its neighboring vertices. In this paper, we randomly choose a label from the neighborhood and assign it to a vertex. In our experiments, we find the revised turbulence operation works better.

To better preserve population diversity, we have modified the subproblems update strategy. After a new offspring solution $\boldsymbol{x}_i^{t+1}$ is generated, we use this solution to update only n_r subproblems from the whole population. The pseudocode of the modified strategy is shown in Algorithm 2.

B. Model Selection

The first step of our proposed approach can return a set of equally good solutions. Each solution represents a certain network partition. In order to recommend to the decision maker a best solution (BS), in the second step we have developed a problem-specific model selection strategy. The flow chart of the devised model selection strategy is shown in Fig. 8.

In Fig. 8, the energy function $H(s)$, proposed by Facchetti *et al.* [14], [49], is an index to evaluate the degree of balance. The function $H(s)$ reads

$$H(s) = \sum_{i,j} \big(1 - \omega_{ij}(s_i \cdot s_j)\big)/2 \tag{5}$$

where $s_i \in \{\pm 1\} = \mathbb{B}_2$ is the sign[4] of node i. If a network is balanced, its energy function value is zero.

[4]According to Heider's balance theory, a balanced network can be divided into two groups. Then we label the nodes in one group with $+1$ and the nodes in the other group with -1.

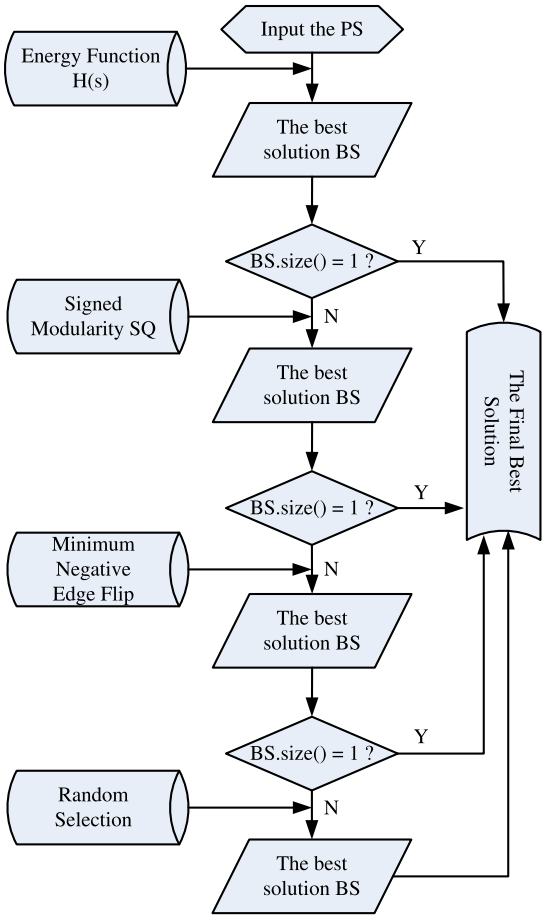

Fig. 8. Flow chart of the suggested model selection strategy.

Given that a network G has a partition $\Omega = (c_1, c_2, \ldots, c_k)$. We further define $\Gamma(\cdot) = s_i \cdot s_j$. Because, if nodes i and j are in the same group, then $\Gamma(\cdot) = 1$, otherwise, $\Gamma(\cdot) = -1$. Then $H(s)$ can be rewritten as

$$\begin{aligned} H(s) &= \sum_{m=1}^{k} \sum_{i,j \in c_m} \big(1 - \omega_{ij}\Gamma(\cdot)\big)/2 \\ &\quad + \sum_{m=1}^{k} \sum_{n=1, n\neq m}^{k} \sum_{i \in c_m, j \in c_n} \big(1 - \omega_{ij}\Gamma(\cdot)\big)/2 \\ &= \sum_{m=1}^{k} \sum_{i,j \in c_m} \left|\omega_{ij}^{-}\right| + \sum_{m=1}^{k} \sum_{n=1, n\neq m}^{k} \sum_{i \in c_m, j \in c_n} \left|\omega_{ij}^{+}\right| \\ &= \sum_{m=1}^{k} \left|e_m^{-}\right| + \sum_{m=1}^{k} \sum_{n=1, n\neq m}^{k} \left|e_{mn}^{+}\right| \end{aligned} \tag{6}$$

where $|e_m^-|$ denotes the number of negative edges within community c_m and $|e_{mn}^+|$ represents the number of positive edges between community c_m and c_n.

From the above equation, we can notice that, the value of $H(s)$ equals the number of imbalanced edges of a network.[5] Consequently, if a signed network has smaller energy value, it will need to flip much less edges to reach balance. Therefore, after obtaining the PSs, we first utilize the $H(s)$ index to choose the solutions that have minimum $H(s)$ value. Because different PSs may have the same $H(s)$ value, if there are more than one BS, then we use the signed modularity index (SQ) to choose the BSs that have the maximum SQ value. The SQ index is proposed by Gómez *et al*. [50]. It is a metric to evaluate the goodness of a network community structure. The SQ index reads

$$\mathrm{SQ} = \frac{1}{2\left(w^+ + w^-\right)} \sum_{i,j} \left[w_{ij} - \left(\frac{w_i^+ w_j^+}{2w^+} - \frac{w_i^- w_j^-}{2w^-} \right) \right] \delta(i,j) \tag{7}$$

where $w_i^+ (w_i^-)$ denotes the sum of all positive (negative) weights of node i. If nodes i and j are in the same community, $\delta(i,j) = 1$, otherwise, $\delta(i,j) = 0$.

A larger value of SQ indicates the network has a strong community structure. However, different community structures may correspond to the same SQ value. In other words, the number of BS may exceed one. From the perspective of sociology, two friends may easily turn into enemies, whereas two enemies may hardly become friends. Putting it another way, it is more difficult to flip a negative edge into positive edge than the reverse way. With respect to this, when the number of BS is larger than one, we use the minimum negative edge flip criterion to choose the BSs from BS, i.e., choose the PS with the corresponding imbalanced network structure that needs much less negative edge flips. After all these, if the number of BS is still larger than one, then we randomly choose one solution as the output.

C. Complexity Analysis

As shown in Algorithm 1, the time complexity of the proposed approach lies in two parts, i.e., the first and the second steps. As discussed in [33], the time complexity for the first step would be $O(\,\mathrm{pop} \cdot \mathrm{gmax} \cdot (m+n))$, where n and m are the node and edge numbers of a network, respectively.

In the second step, the worst case of the model selection step needs $O(\,\mathrm{pop} \cdot (m+n))$ basic operations. Decoding an individual needs $O(n \log n)$ basic operations. The determination and changes of imbalanced edges can be computed in $O(m)$ time at most.

According to the operational rules of the symbol O, the overall time complexity of the proposed two-step approach is $O(\,\mathrm{pop} \cdot \mathrm{gmax} \cdot (m + n \log n))$ in the worst case.

V. Experimental Study

In this section, we apply our proposed approach to synthetic and real-world signed networks. And we also compare our method with several other MOEAs. The C++ code of our algorithm can be downloaded from http://see.xidian.edu.cn/faculty/mggong/index.htm. The experiments were performed on a 3.2 GHz Intel Core i3 CPU 550 machine with 4 GB memory. The operating system is Windows 7, and the compiler is VC++ 6.0.

[5] In (5) and (6), we only compute the upper triangular matrix, because the adjacency matrix is symmetrical.

TABLE I
STATISTICS OF THE SIGNED NETWORKS

Network	n	m	m^+	m^-	$\overline{k}$	D
SPP	10	45	18	27	9.000	1.0000
GGS	16	58	29	29	7.250	0.4833
EGFR	329	779	515	264	4.736	0.0144
Macrophage	678	1,425	947	478	4.204	0.0062
Yeast	690	1,080	860	220	3.130	0.0045
Ecoli	1,461	3,215	1,879	1,336	4.401	0.0030
WikiElec	7,114	100,321	78,792	21,529	28.204	0.0040
Slashdot	77,357	466,666	352,890	113,776	12.065	1.5597

m^+ and m^- denote the numbers of positive and negative edges, respectively. $\overline{k}$ is the averaged node degree. D represents the density of a network.

A. Signed Network Data Sets

1) Benchmark Networks: The network generator proposed by Lancichinetti, Fortunato and Radicchi (LFR) [51] is a reliable model for benchmarking. However, this model is originally designed for unsigned networks.

In this paper, we have extended the LFR model to signed context. A signed LFR model is depicted by SLFR (n, k_{avg}, $k_{\max}$, γ, β, $s_{\min}$, $s_{\max}$, μ, $p-$, $p+$), where n is the number of nodes; k_{avg} and $k_{\max}$ are the averaged and maximum degree of a node, respectively; γ and β are the exponents for the power law distribution of node degree and community size, respectively; $s_{\min}$ and $s_{\max}$ are the minimum and maximum community size, respectively. μ is a mixing parameter. Each node shares a fraction $1-\mu$ of its links with the other nodes of its community and a fraction μ with the other nodes of the network. $p-$ is the fraction of negative edges within communities and $p+$ is the fraction of positive edges between different communities.

In our experiments, the used signed LFR model is SLFR (500, 10, 20, 2, 1, 20, 40, μ, $p-$, $p+$), with μ ranging from 0.1 to 0.6 at intervals of 0.1, and $p-$ and $p+$, respectively, ranging from 0 to 1 at intervals of 0.2. Consequently, we have generated 216 benchmark networks whose ground truths are known.

2) Real-World Networks: Eight real-world signed networks are employed to test the performance of the proposed algorithm. The networks are the Slovene parliamentary party (SPP) network [52], the Gahuku-Gama subtribes (GGS) network [53], the epidermal growth factor receptor pathway (EGFR) network [54], the Macrophage network (Macrophage) [55], the Yeast network (Yeast) [56], the gene regulatory network of the Escherichia coli (Ecoli, http://regulondb.ccg.unam.mx, version 6.3) [57], the election network of Wikipedia (WikiElec) [58], and the social network of the blog Slashdot (Slashdot) [58]. The statistics of each network are given in Table I.

B. Validation Experiments

Because, in our proposed approach, the output of the first step is the input of the second step, in this section, we first will check the abilities of the first step for community detection, then we will check whether the second step can solve the structural balance problem effectively.

1) Community Detection Validation: We run the first step of our proposed approach for 30 independent trials for each of the 216 generated benchmark networks. Because the ground truth of each benchmark network is known, we employ the normalized mutual information (NMI) [59], index to evaluate the goodness of a network partition.

Given that A and B are two partitions of a network, respectively, C is a confusion matrix, C_{ij} equals to the number of nodes shared in common by community i in partition A and by community j in partition B. Then NMI (A, B) is written as

$$\text{NMI} = \frac{-2\sum_{i=1}^{C_A}\sum_{j=1}^{C_B} C_{ij}\log\left(C_{ij}\cdot n/C_{i.}C_{.j}\right)}{\sum_{i=1}^{C_A} C_{i.}\log(C_{i.}/n)+\sum_{j=1}^{C_B} C_{.j}\log\left(C_{.j}/n\right)} \tag{8}$$

where C_A (or C_B) is the number of clusters in partition A (or B), $C_{i.}$ (or $C_{.j}$) is the sum of elements of C in row i (or column j). NMI $(A, B) = 1$ means that A and B are identical and NMI $(A, B) = 0$ indicates that A and B are completely different.

For each single run we record the maximum NMI value from the PF. The statistical results are presented in Fig. 9.

As can be seen from Fig. 9, when μ is not bigger than 0.3 and $p-$ is not bigger than 0.2, we can successfully detect the ground truths of the networks. As μ increases, the community structures become vague and we fail to obtain the right partitions.

We further can notice from the figure that when $p-$ increases, the decline of the NMI values is obvious. However, from the viewpoint of $p+$, the decline of the NMI values is not obvious. This phenomenon indicates that from the perspective of community detection, our method is robust to noises of positive edges between communities.

2) Structural Balance Validation: The above validation experiments indicate that the first step is promising for community detection. In this section, we will validate the performance of the second step.

In order to better visualize network structures, we test our method on the two small networks as shown in Fig. 5(d) and (e). The PFs obtained by the first step with one run are shown in Fig. 10.

We can see from the PFs that for the first network, the first step obtains two nondominated solutions. We get the optimal solution with the $H(s)$ value of 0 which indicates the proposed approach has discovered the network partition that is structurally balanced. For the second network, we obtain only one optimal solution with the $H(s)$ value of 1 which means one edge flip is needed to make the network balanced. The balanced structures of the two networks are presented in Fig. 11.

The balance result shown in Fig. 11(b) is interesting. The result suggests that, in order to realize structural balance, the negative edge between nodes "Ru" and "GB" should be changed into a positive one. In other words, the result suggests that Great Britain should cease fire with Russia.

According to the history, in 1907, a bipartite agreement between Russia and Great Britain was established, that then bound Russia, Great Britain, and France into the

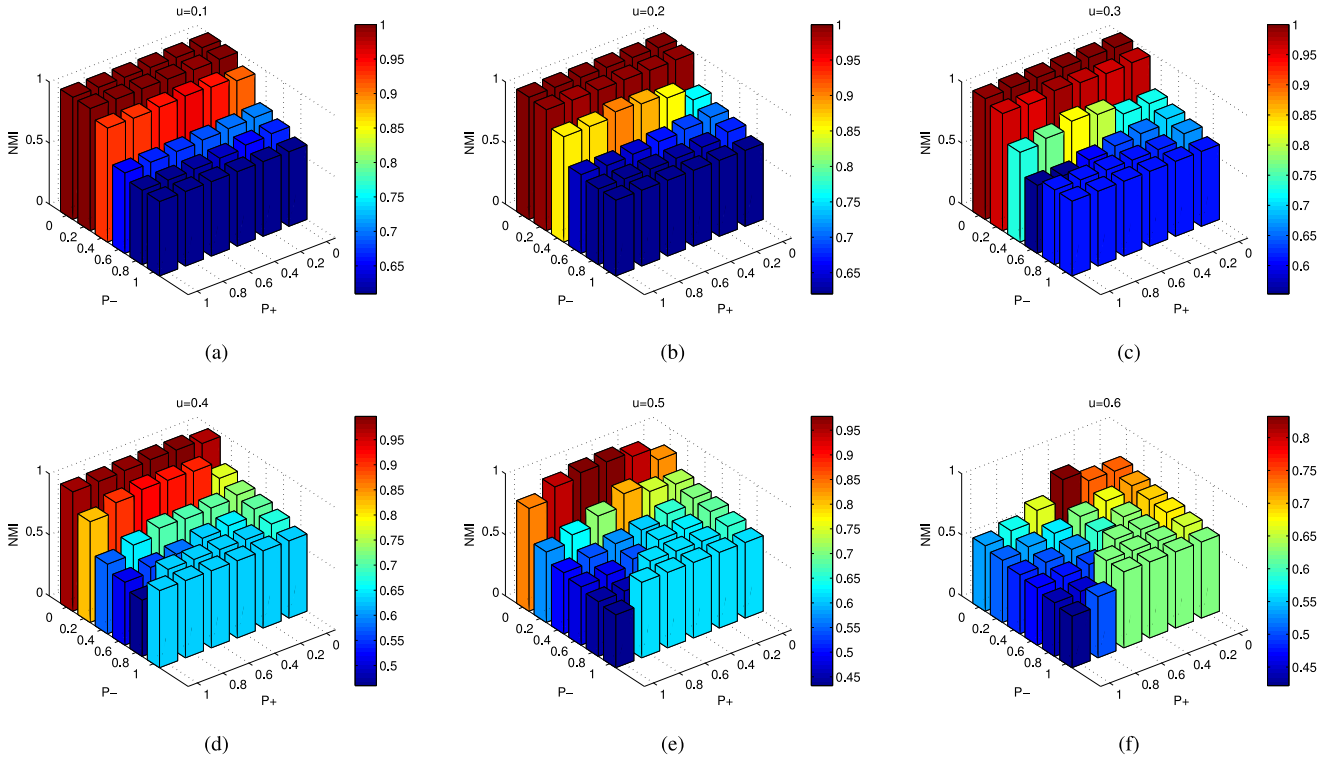

Fig. 9. Averaged NMI values for the combinations of the three parameters μ, $p-$, and $p+$ when experimenting on the 216 signed benchmark networks. (a) for $\mu = 0.1$, (b) for $\mu = 0.2$, (c) for $\mu = 0.3$, (d) for $\mu = 0.4$, (e) for $\mu = 0.5$, (f) for $\mu = 0.6$.

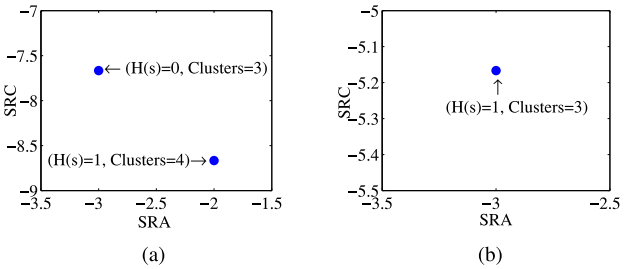

Fig. 10. PFs obtained by the first step with one run on the two small networks. The $H(s)$ values and the clusters are recorded. (a) is the result for the network shown in Fig. 5(d), and (b) is for the network shown in Fig. 5(e).

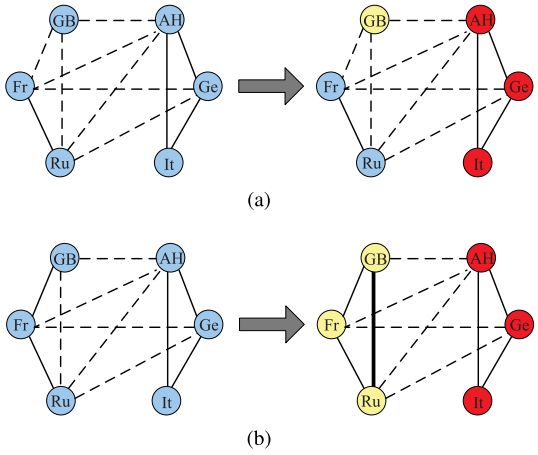

Fig. 11. Structural balance results for the two signed networks. (a) Network 1. (b) Network 2.

Triple Entente. From this viewpoint, we can see that the second step of our proposed approach is effective. The balance results are helpful for decision making.

TABLE II
PARAMETER SETTINGS OF THE ALGORITHMS

Algorithm	pop	$gmax$	pc	pm	T	n_r	reference
two-step	100	100	—	0.9	10	2	[ours]
MODPSO	100	100	—	0.9	40	—	[33]
NSGA-II	100	100	0.9	0.9	—	—	[39]
MOEA/D	100	100	0.9	0.9	40	—	[61]

C. Comparisons With Other MOEAs

In the first step, we have improved the MODPSO algorithm suggested in our previous work. Because the output of the first step impacts the final balance result, it is necessary to test whether the improvements made to MODPSO are effective.

In this section, comparisons between the two-step and several other MOEAs which optimize the same objective functions as two-step does are made. The compared MOEAs are the MODPSO algorithm, the fast and elitist multiobjective genetic algorithm (NSGA-II) proposed in [38], and the multiobjective evolutionary algorithm based on decomposition (MOEA/D) devised in [60]. The parameter settings of the algorithms are listed in Table II.

We test all the algorithms on the signed networks listed in Table I, and we run each algorithm 30 times. We record the PFs for each algorithm and the hypervolume index (IH) is adopted to estimate the performance of the MOEAs.

The IH, also known as the S metric or the Lebesgue measure, was originally proposed by Zitzler and Thiele [61]. The hypervolume of a set of solutions measures the size of the portion of objective space that is dominated by those solutions collectively. Generally, hypervolume is favored because it captures in a single scalar both the closeness of the solutions to the optimal set and, to some extent, the spread of the solutions

TABLE III
HYPERVOLUMES OF THE PFS GENERATED BY THE TWO-STEP, MODPSO, MOEA/D, AND NSGA-II AFTER 30 INDEPENDENT RUNS FOR THE EIGHT SIGNED NETWORKS

network	IH	two-step	MODPSO	MOEA/D	NSGA-II
SPP	Max	**1.4212**	1.4212	1.4169	0.9558
	Mean	**1.4210**	1.3923	1.3729	0.6313
	Std	0.0013	0.0551	0.0265	0.1271
	RST	×	≈	+	+
GGS	Max	**1.3497**	1.2126	1.2315	0.8081
	Mean	**1.3146**	1.0019	1.0142	0.6894
	Std	0.0148	0.1701	0.0992	0.1394
	RST	×	+	+	+
EGFR	Max	**1.2039**	0.8695	0.6948	0.4461
	Mean	**1.1330**	0.7957	0.5865	0.4114
	Std	0.0318	0.0347	0.0722	0.0185
	RST	×	+	+	+
Macrophage	Max	**1.2552**	0.9815	0.7750	0.4495
	Mean	**1.2065**	0.8653	0.6314	0.4251
	Std	0.0294	0.0472	0.0758	0.0130
	RST	×	+	+	+
Yeast	Max	**1.2138**	1.056	0.9008	0.5482
	Mean	**1.1695**	0.8766	0.6346	0.4967
	Std	0.0321	0.0758	0.1217	0.0263
	RST	×	+	+	+
Ecoli	Max	**1.1923**	0.9200	0.8929	0.5337
	Mean	**1.1257**	0.8054	0.6781	0.4789
	Std	0.0421	0.0705	0.1203	0.0255
	RST	×	+	+	+
WikiElec	Max	0.1468	0.0475	**0.2456**	0.0831
	Mean	0.0980	0.0462	**0.1464**	0.0811
	Std	0.0325	0.0006	0.0595	0.0009
	RST	×	+	−	+
Slashdot	Max	**0.8428**	0.1989	0.4941	0.3510
	Mean	**0.6121**	0.1172	0.4827	0.2979
	Std	0.1336	0.0537	0.0091	0.0748
	RST	×	+	+	+

The symbols "≈", "+" and "−" denote that the performance of the two-step approach is similar to, significantly better than, and worse than that of the compared method, respectively.

across objective space. If A is a set of nondominated solutions, the hypervolume function of A is defined as

$$\mathrm{IH}(A, \boldsymbol{y}_{\mathrm{ref}}) = \Lambda\left(\bigcup_{\boldsymbol{y}\in A}\left\{\boldsymbol{y}' | \boldsymbol{y} \prec \boldsymbol{y}' \prec \boldsymbol{y}_{\mathrm{ref}}\right\}\right),\ A \subseteq \mathbb{R}^m \quad (9)$$

where Λ denotes the Lebesgue measure; $\boldsymbol{y}_{\mathrm{ref}} \in \mathbb{R}^m$ denotes the reference point that should be dominated by all Pareto-optimal solutions. In our experiments, when calculating the IH index, the hypervolumes have been normalized and the reference point $\boldsymbol{y}_{\mathrm{ref}}$ is set to (1.2, 1.2).

We also perform a statistical analysis using the Wilcoxon's rank sum test (RST). In the two-sided RST experiments, the significance level is set to 0.05 with the null hypothesis that the two independent samples (the two IH sets) come from distributions with equal means. The statistical results are summarized in Table III.

It can be seen from the table that the proposed approach outperforms the compared MOEA-based methods. The two-step approach obtains larger IH values on the eight real-world signed networks. The RSTs also suggest that the two-step approach works better. We can observe from the table that the two-step approach outperforms the MODPSO method which was proposed in [33].

Fig. 12 displays the PFs with the largest IH values obtained by the two-step, MODPSO, MOEA/D, and NSGA-II. It is graphically obvious that our proposed two-step approach has better performance than the compared methods.

An MOEA generally concerns two key components, the reproduction (i.e., generate offspring) and the replacement (i.e., update parents). When integrating PSO into an MOEA, the replacement strategy should be taken into thorough consideration. MODPSO uses the Chebyshev decomposition method. When applying MODPSO to implement the multiobjective network clustering, several outstanding PSs which are better than the majority of the solutions in the population appear easily. In this situation, if we update the neighboring subproblems with these outstanding solutions, population diversity will be lost. However, in the MODPSO, the particle position status update rule is a differential-like operator. Because the leader particle is randomly selected from the neighboring subproblems, even if the newly generated solution is excellent, the population diversity can still be preserved by the differential-like particle position status update rule. This is the reason why the PF obtained by MODPSO is better than that obtained by MOEA/D.

MODPSO simply updates all the neighboring subproblems. When the newly generated solution is better enough, the neighboring subproblems will be replaced by the newly generated solution, and consequently, diversity will be lost. To overcome this drawback, in this paper we have modified the replacement strategy. Instead of updating all the neighboring subproblems, we only update n_r subproblems from the population. In the MODPSO, each particle represents a subproblem, the personal BS update is actually an update of the subproblem of itself. If we update subproblems from the population, we may avoid the situation that all the neighboring solutions are replaced by one offspring solution, and diversity will be enhanced. It can be seen from both Fig. 12 and Table III that the improved replacement strategy in this paper works better.

We can draw the conclusion from the above experiments that the improved MODPSO framework is efficient. The effectiveness of the first step ensures the ultimate performance. In the next section, we will test the performance of the two-step approach on addressing the structural balance problem.

D. Structural Balance Experiments

For each signed network, we only analyze four PSs on the PF. These solutions have the smallest $H(s)$ values. The statistical results are summarized in Table IV.

When experimenting on the SPP network, the first step of our proposed approach yields only three PSs, as shown in Fig. 12(a). After the second step, We choose the solution with the smallest $H(s)$ value as the final result. Fig. 13 displays the structural balance result of the SPP network.

For the SPP network, we obtain a partition with three clusters, as shown in Fig. 13(b). The $H(s)$ value is 2, which means two edges need to be flipped so as to make the network balanced. After flipping the two positive edges into negative ones, we obtain a balanced network, as shown in Fig. 13(c).

It should be pointed out that, the ground truth of the SPP network is that it has been divided into two parts as shown in Fig. 14(a).

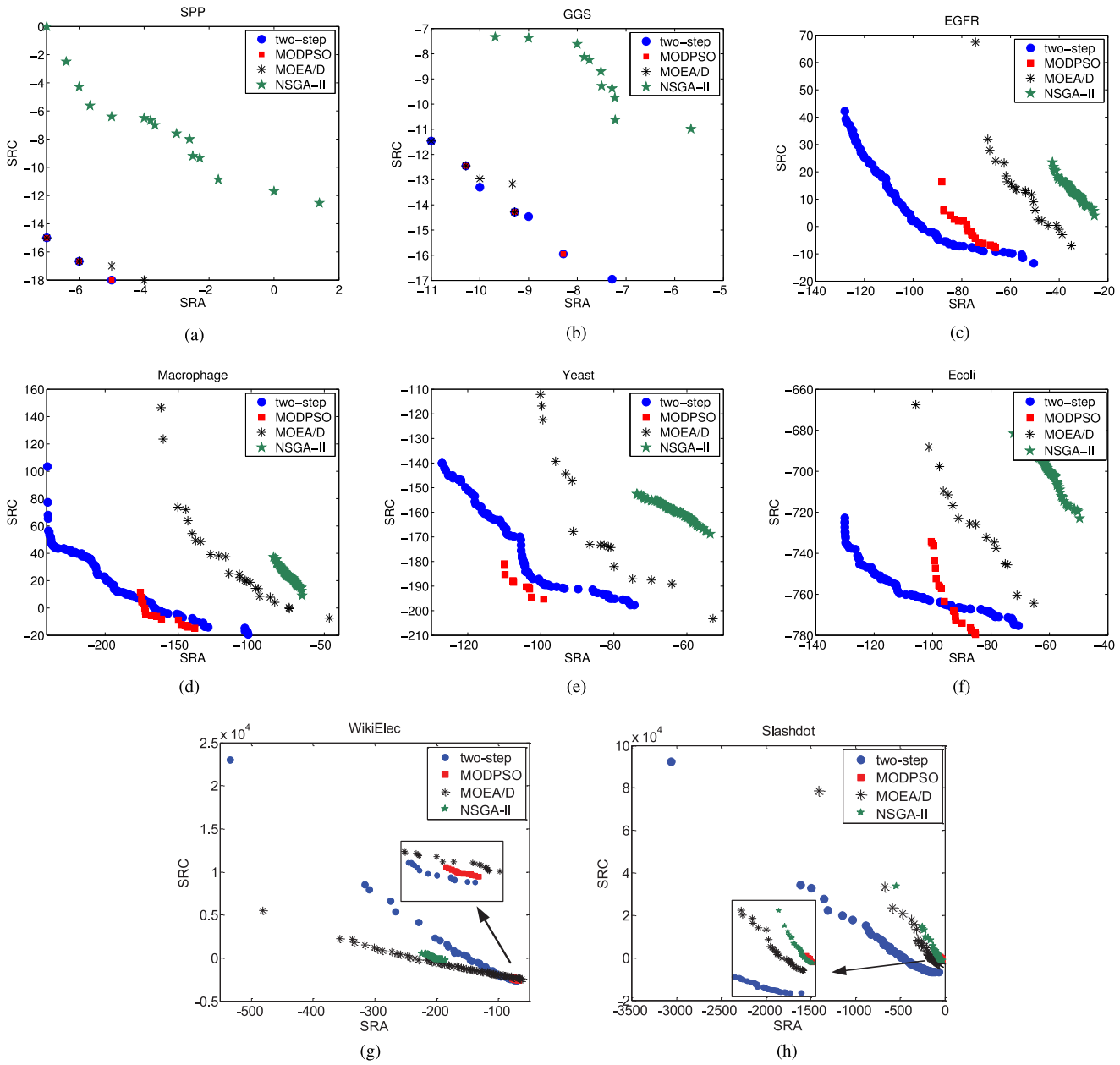

Fig. 12. Comparisons between the PFs with largest IH values obtained by the four MOEAs. (a) SPP. (b) GGS. (c) EGFR. (d) Macrophage. (e) Yeast. (f) Ecoli. (g) WikiElec. (h) Slashdot.

It can been seen from Fig. 14(a) that nodes "SNS," "LDS," "ZS-ESS," "DS," and "ZLSD" belong to the same community. In the final population of our proposed approach, we have found the solution that corresponds to the ground truth partition of the SPP network, however, this solution is dominated by other solutions. We have reasons to believe this domination is reasonable. From the perspective of MOEA, the solution is indeed dominated by other solutions. From the perspective of structural balance theory, although this solution corresponds to the same $H(s)$ value of 2 which means that we can flip two edges to make the network balanced and the balanced structure is shown in Fig. 14(b), it will need more effort to do so. It can be seen from Fig. 14 that we need to change the negative edge between nodes SNS and DS and the edge between nodes SNS and ZS-ESS into a positive edge so as to make the network balanced. However, from the view point of sociology, it is too hard for two enemies to become friends, but it is rather easy for two friends to become enemies. In other words, it is more difficult to flip a negative edge into a positive edge than the reverse way. Moreover, as for the topology of the SPP network, the node SNS is special. It has negative links with nodes DS and ZS-ESS. Our proposed two-step approach simply detects this node as a separate cluster.

For the GGS network, our proposed approach discloses four communities. The corresponding energy function value is 4, which indicates four edges need to be changed. The detected community structure and the balanced network topology are exhibited in Fig. 15.

TABLE IV
STRUCTURAL BALANCE EXPERIMENTAL RESULTS OBTAINED BY OUR PROPOSED APPROACH

network	$H(s)$	SQ	cluster	E^{+-}	E^{-+}	$H_b(s)$
SPP	2	0.4086	3	2	0	0
	5	0.3228	4	5	0	0
	7	0.2685	5	7	0	0
	×	×	×	×	×	×
GGS	4	0.3870	4	4	0	0
	7	0.3148	5	7	0	0
	8	0.3511	5	8	0	0
	9	0.2678	6	9	0	0
EGFR	283	0.2848	73	212	71	0
	289	0.2789	76	219	70	0
	289	0.2779	73	217	72	0
	291	0.2728	73	220	71	0
Macrophage	461	0.3284	74	168	293	0
	462	0.3280	77	173	289	0
	464	0.3269	76	179	285	0
	465	0.3266	81	179	286	0
Yeast	146	0.6065	109	121	25	0
	149	0.6013	108	124	25	0
	154	0.5987	113	130	24	0
	155	0.5934	113	130	25	0
Ecoli	443	0.3989	401	387	56	0
	444	0.3967	401	384	60	0
	474	0.3907	408	415	59	0
	476	0.3873	408	416	60	0
WikiElec	18,753	0.0016	915	18,715	38	0
	18,754	0.0017	914	18,660	94	0
	18,759	0.0017	936	18,636	123	0
	18,870	0.0018	927	18,604	266	0
Slashdot	86,197	0.0035	2,810	98	86,099	0
	86,216	0.0029	2,906	487	85,729	0
	86,309	0.0057	2,887	539	85,770	0
	86,474	0.0073	2,927	997	85,475	0

E^{+-} (or E^{-+}) denotes the number of positive (or negative) edges that need to be flipped into negative (or positive) edges. $H_b(s)$ is the energy function value of the network to which edge flips have been made.

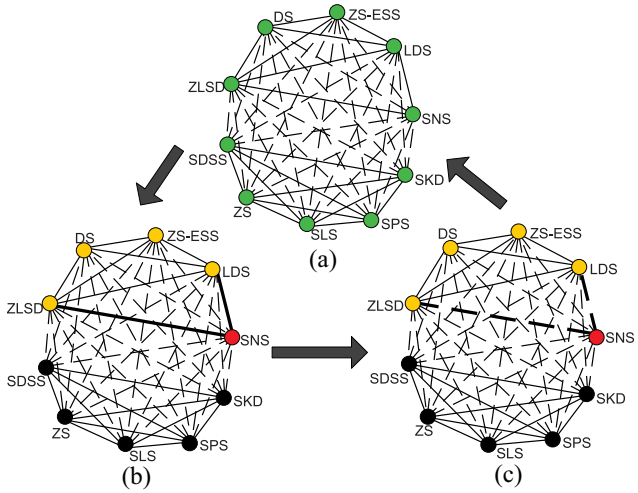

Fig. 13. Imbalanced and balanced topology structures of the SPP network. (a) Original network topology. (b) Network partition. (c) Balanced network topology.

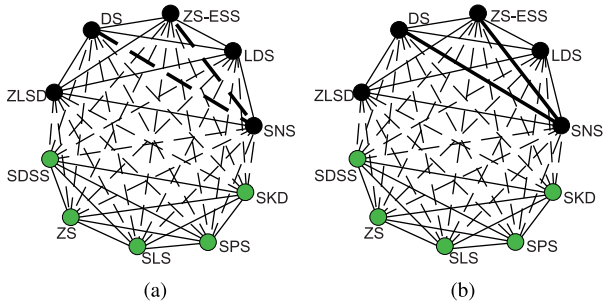

Fig. 14. Structures of the SPP network. (a) Ground truth partition and (b) corresponding balanced structure.

From Fig. 15, we can clearly see that the proposed approach has found four imbalanced edges, exhibited in Fig. 15(b).

Fig. 16 displays the ground truth of the GGS network and the corresponding balanced network structure. The GGS network is originally divided into three communities. However, the solution denoting the true partition is dominated by the other solutions. Although, by flipping the two edges connecting node "MASIL" with nodes "NAGAM" and "UHETO" we can also obtain a balanced network structure as shown in Fig. 16(b), we think that this kind of structure is not stable in real life. In Fig. 16, the four nodes NAGAM, "SEUVE," UHETO, and "NOTOH" form a quadrangle which is unstable in real life since a triangular relation is commonly recognized as the most stable status. The proposed approach separates the two members NAGAM and SEUVE from the original tribe to form a new subtribe.

Because the scales of the remaining six signed networks are large, it is hard to display their balanced network structures. However, from what is recorded in Table IV we can note that, after flipping the imbalanced edges discovered by our proposed approach, the $H_b(s)$ values became 0, which indicates that the changed networks have gotten structurally balanced. The WikiElec and Slashdot networks are large in size, however, the obtained $H(s)$ values for the two networks are comparable to those reported in [14].

The above experiments demonstrate that the proposed two-step approach for addressing the network structural balance problem is effective. Our proposed method can simultaneously give a macroscopic (community structure) and a microscopic view (imbalanced edges) to analyze a signed social network.

E. Discussion on Parameters

Because the output of the first step of our proposed approach affects the final decision, thus, it is necessary to investigate the effects of the parameters on the performance of the approach.

In our proposed approach, in order to maintain the diversity of the population, we have improved the replacement operation, in which n_r is a key parameter. The value of n_r is significant in balancing diversity and convergence of the population. The n_r parameter was first proposed in [62], in which n_r is set to $n_r = 0.01 \cdot \text{popsize}$. To check its impact, we calculate the hypervolumes of the PFs generated by the two-step with n_r ranging from 1 to 5 at intervals of 1. The results are recorded in Table V.

The parameter n_r cannot be too large. On one hand, a large n_r value will lead to the loss of diversity, on the other hand, a large n_r value will consume a lot of computational time. The results in Table V suggest that the influence of the n_r parameter is slight. Generally, $n_r = 2$ seems to be the best.

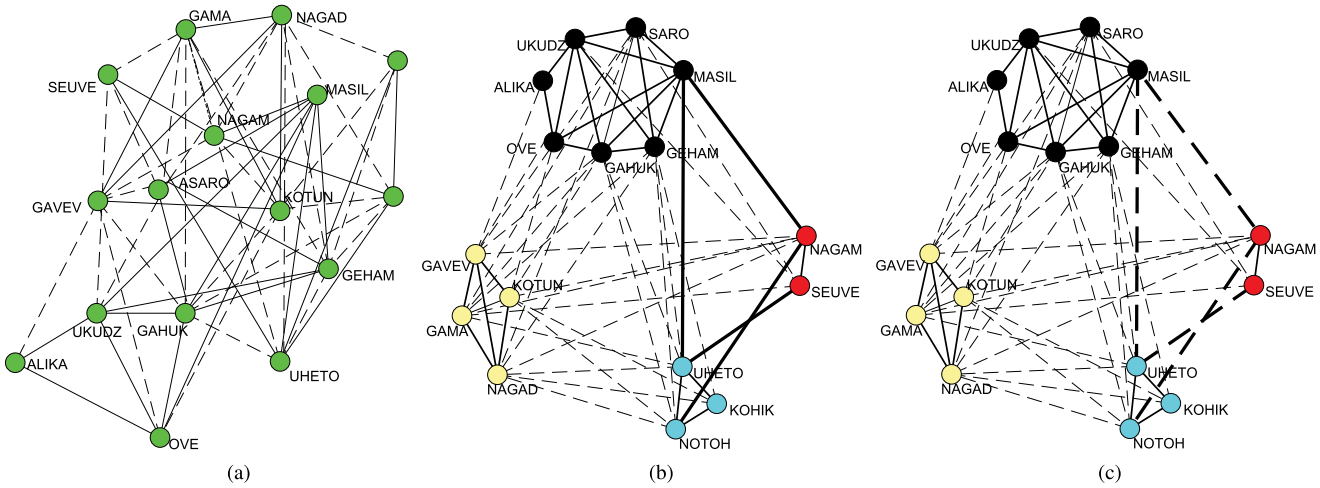

Fig. 15. Topology structures of the GGS network. (a) Original network structure. (b) Network partition. (c) Balanced network topology.

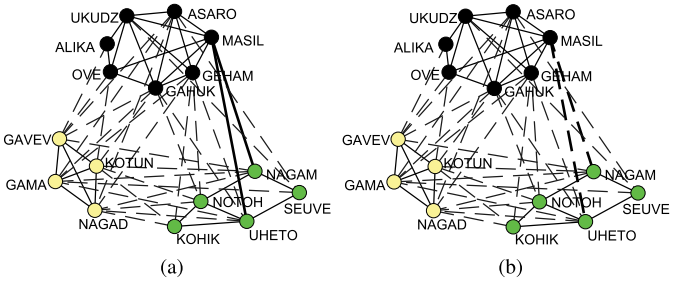

Fig. 16. Structures of the GGS network. (a) Ground truth partition and (b) corresponding balanced structure.

TABLE V
HYPERVOLUMES OF THE PFS GENERATED BY THE TWO-STEP WITH DIFFERENT SETTINGS OF THE PARAMETER n_r. FOR EACH n_r, WE RUN THE TWO-STEP APPROACH FOR 30 INDEPENDENT RUNS FOR THE EIGHT SIGNED NETWORKS

network	IH	n_r=1	n_r=2	n_r=3	n_r=4	n_r=5
SPP	Mean	1.4212	1.4210	1.4210	1.4212	1.4212
	Std	0.0020	0.0013	0.0013	0.0019	0.0021
GGS	Mean	1.3053	1.3146	1.3120	1.3060	1.2934
	Std	0.0147	0.0148	0.0176	0.0282	0.0381
EGFR	Mean	1.1215	1.1330	1.1377	1.1415	1.0604
	Std	0.0376	0.0318	0.0476	0.0430	0.1184
Macrophage	Mean	1.1919	1.2065	1.2111	1.2044	1.1718
	Std	0.0257	0.0294	0.0261	0.0392	0.0743
Yeast	Mean	1.1470	1.1695	1.1303	1.1394	1.1263
	Std	0.0301	0.0321	0.0323	0.0448	0.0391
Ecoli	Mean	1.0894	1.1257	1.1149	1.1141	1.0458
	Std	0.0414	0.0421	0.0358	0.0483	0.1064
WikiElec	Mean	0.0934	0.0980	0.1057	0.1305	0.1276
	Std	0.0251	0.0325	0.0282	0.0196	0.0147
Slashdot	Mean	0.5957	0.6121	0.6096	0.6073	0.6007
	Std	0.1738	0.1336	0.1124	0.1434	0.1401

VI. CONCLUSION

Network analysis is one of the theoretical underpinnings of social computing and big data. The past decades have witnessed the prosperity of the research on network structure analysis, such as the network robustness analysis, the community structure analysis, the structural balance analysis, etc. Network structural balance has grown into a large area of study.

This paper tries to address the network structural balance problem with an optimization view. A two-step approach was proposed for the first time to solve the problem. In the first step of the proposed approach, we employed an improved version of the multiobjective discrete particle swarm optimization framework suggested in our previous work to cluster a signed network. Each single run of the method can yield a Pareto set containing a set of equally good solutions. Each solution represents a kind of network partition. In the second step, we devised a problem-specific model selection strategy to select the BS, then we determined and flipped all the imbalanced edges based on the discovered network community structure. Extensive experiments on eight real-world signed networks have demonstrated the effectiveness of the proposed approach.

The proposed two-step approach actually acts as a paradigm for addressing the network structural balance problem. It is meaningful and constructive for sociologists. However, help from scientists from the field of sociology, computer science, psychology, etc., is still needed to perfect the paradigm. In our experimental studies, because the Slashdot network is large in size, our experiments on that network are conducted on a distributed computing platform consisting of six computers, each of which has four cores and 8 GB memory. In order to handle big networks, for the first step of our proposed paradigm, more researches on scaling up evolutionary computing techniques to solve large scale global optimization problems can be done. For the second step of the paradigm, knowledge from social science is desired to improve the model selection strategy. We expect that complex network analysis's scope will continue to expand and its applications to multiply. Our future work will focus on more in-depth analysis of network issues. Such analysis is expected to shed light on how networks change the real world.

REFERENCES

[1] M. E. J. Newman, *Networks: An Introduction*. Oxford, U.K.: Oxford Univ. Press, 2010.
[2] S. H. Strogatz, "Exploring complex networks," *Nature*, vol. 410, no. 6825, pp. 268–276, 2001.

[3] M. E. J. Newman, "Complex systems: A survey," *Am. J. Phys.*, vol. 79, no. arXiv: 1112.1440, pp. 800–810, 2011.

[4] G. Palla, A.-L. Barabási, and T. Vicsek, "Quantifying social group evolution," *Nature*, vol. 446, no. 7136, pp. 664–667, 2007.

[5] M. Szell, R. Lambiotte, and S. Thurner, "Multirelational organization of large-scale social networks in an online world," *Proc. Nat. Acad. Sci. USA*, vol. 107, no. 31, pp. 13636–13641, 2010.

[6] M. E. J. Newman, "The structure and function of complex networks," *SIAM Rev.*, vol. 45, no. 2, pp. 167–256, 2003.

[7] S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, and D.-U. Hwang, "Complex networks: Structure and dynamics," *Phys. Rep.*, vol. 424, nos. 4–5, pp. 175–308, 2006.

[8] R. Cohen and S. Havlin, *Complex Networks: Structure, Robustness And Function*. Cambridge, MA, USA: Cambridge Univ. Press, 2010.

[9] S. Fortunato, "Community detection in graphs," *Phys. Rep.*, vol. 486, nos. 3–5, pp. 75–174, 2010.

[10] D. J. Watts and S. H. Strogatz, "Collective dynamics of 'small-world' networks," *Nature*, vol. 393, no. 6684, pp. 440–442, 1998.

[11] A.-L. Barabási and R. Albert, "Emergence of scaling in random networks," *Science*, vol. 286, no. 5439, pp. 509–512, 1999.

[12] 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, Jun. 2002.

[13] F. Heider, "Social perception and phenomenal causality," *Psychol. Rev.*, vol. 51, no. 6, pp. 358–374, 1944.

[14] G. Facchetti, G. Iacono, and C. Altafini, "Computing global structural balance in large-scale signed social networks," *Proc. Nat. Acad. Sci. USA*, vol. 108, no. 52, pp. 20953–20958, 2011.

[15] A. Srinivasan, "Local balancing influences global structure in social networks," *Proc. Nat. Acad. Sci. USA*, vol. 108, no. 5, pp. 1751–1752, 2011.

[16] D. Cartwright and F. Harary, "Structural balance: A generalization of Heider's theory," *Psychol. Rev.*, vol. 63, no. 5, pp. 277–293, 1956.

[17] F. Heider, "Attitudes and cognitive organization," *J. Psychol.*, vol. 21, no. 1, pp. 107–112, 1946.

[18] F. Harary, "On local balance and *N*-balance in signed graphs," *Michigan Math. J.*, vol. 3, no. 1, pp. 37–41, 1955.

[19] F. Harary, "A matrix criterion for structural balance," *Naval Res. Logist. Quart.*, vol. 7, no. 2, pp. 195–199, 1960.

[20] X. Zheng, D. Zeng, and F.-Y. Wang, "Social balance in signed networks," *Inf. Syst. Front.*, pp. 1–19, Jan. 2014. [Online]. Available: http://dx.doi.org/10.1007/s10796-014-9483-8

[21] S. Wasserman, *Social Network Analysis: Methods and Applications*, vol. 8. Cambridge, MA, USA: Cambridge Univ. Press, 1994.

[22] D. Easley and J. Kleinberg, *Networks, Crowds, and Markets: Reasoning About a Highly Connected World*. Cambridge, MA, USA: Cambridge Univ. Press, 2010.

[23] P. Doreian and A. Mrvar, "Partitioning signed social networks," *Social Netw.*, vol. 31, no. 1, pp. 1–11, 2009.

[24] J. Kunegis *et al.*, "Spectral analysis of signed graphs for clustering, prediction and visualization," in *Proc. SDM*, vol. 10. Columbus, OH, USA, 2010, pp. 559–570.

[25] E. Terzi and M. Winkler, "A spectral algorithm for computing social balance," in *Algorithms and Models for the Web Graph*. Berlin, Germany: Springer, 2011, pp. 1–13.

[26] L. Wu, X. Ying, X. Wu, A. Lu, and Z.-H. Zhou, "Spectral analysis of *k*-balanced signed graphs," in *Advances in Knowledge Discovery and Data Mining*. Berlin, Germany: Springer, 2011, pp. 1–12.

[27] M. Kim and K. S. Candan, "SBV-Cut: Vertex-cut based graph partitioning using structural balance vertices," *Data Knowl. Eng.*, vol. 72, pp. 285–303, Feb. 2012.

[28] I. Boussaïd, J. Lepagnot, and P. Siarry, "A survey on optimization metaheuristics," *Inf. Sci.*, vol. 237, pp. 82–117, Jul. 2013.

[29] M. Črepinšek, S.-H. Liu, and M. Mernik, "Exploration and exploitation in evolutionary algorithms: A survey," *ACM Comput. Surv.*, vol. 45, no. 3, pp. 1–33, 2013.

[30] V. Guliashki, H. Toshev, and C. Korsemov, "Survey of evolutionary algorithms used in multiobjective optimization," *Prob. Eng. Cybern. Robot.*, vol. 60, pp. 42–54, Jun. 2009.

[31] A. Zhou *et al.*, "Multiobjective evolutionary algorithms: A survey of the state of the art," *Swarm Evol. Comput.*, vol. 1, no. 1, pp. 32–49, 2011.

[32] E. Triantaphyllou, *Multi-Criteria Decision Making Methods: A Comparative Study*. New York, NY, USA: Springer, 2000.

[33] M. Gong, Q. Cai, X. Chen, and L. Ma, "Complex network clustering by multiobjective discrete particle swarm optimization based on decomposition," *IEEE Trans. Evol. Comput.*, vol. 18, no. 1, pp. 82–97, Feb. 2014.

[34] J. Leskovec, D. Huttenlocher, and J. Kleinberg, "Signed networks in social media," in *Proc. SIGCHI Conf. Human Factors Comput. Syst.*, Atlanta, GA, USA, 2010, pp. 1361–1370.

[35] T. Antal, P. L. Krapivsky, and S. Redner, "Social balance on networks: The dynamics of friendship and enmity," *Phys. D*, vol. 224, no. 1, pp. 130–136, 2006.

[36] J. A. Davis, "Clustering and structural balance in graphs," *Human Relat.*, vol. 20, no. 2, pp. 181–187, 1967.

[37] J. D. Knowles and D. W. Corne, "Approximating the nondominated front using the Pareto archived evolution strategy," *Evol. Comput.*, vol. 8, no. 2, pp. 149–172, 2000.

[38] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan, "A fast and elitist multiobjective genetic algorithm: NSGA-II," *IEEE Trans. Evol. Comput.*, vol. 6, no. 2, pp. 182–197, Apr. 2002.

[39] E. Zitzler, M. Laumanns, and L. Thiele, "SPEA2: Improving the strength Pareto evolutionary algorithm," in *Proc. Evol. Methods Design Optim. Control Appl. Ind. Prob.*, Barcelona, Spain, 2002, pp. 95–100.

[40] C. Coello, G. Pulido, and M. Lechuga, "Handling multiple objectives with particle swarm optimization," *IEEE Trans. Evol. Comput.*, vol. 8, no. 3, pp. 256–279, Jun. 2004.

[41] Q. Zhang and H. Li, "MOEA/D: A multiobjective evolutionary algorithm based on decomposition," *IEEE Trans. Evol. Comput.*, vol. 11, no. 6, pp. 712–731, Dec. 2007.

[42] M. Gong, L. Jiao, H. Du, and L. Bo, "Multiobjective immune algorithm with nondominated neighbor-based selection," *Evol. Comput.*, vol. 16, no. 2, pp. 225–255, 2008.

[43] S. Bandyopadhyay, S. Saha, U. Maulik, and K. Deb, "A simulated annealing-based multiobjective optimization algorithm: AMOSA," *IEEE Trans. Evol. Comput.*, vol. 12, no. 3, pp. 269–283, Jun. 2008.

[44] J. Bader and E. Zitzler, "HypE: An algorithm for fast hypervolume-based many-objective optimization," *Evol. Comput.*, vol. 19, no. 1, pp. 45–76, 2011.

[45] X.-S. Yang, "Multiobjective firefly algorithm for continuous optimization," *Eng. Comput.*, vol. 29, no. 2, pp. 175–184, 2013.

[46] K. Deb and H. Jain, "An evolutionary many-objective optimization algorithm using reference-point-based nondominated sorting approach, part I: Solving problems with box constraints," *IEEE Trans. Evol. Comput.*, vol. 18, no. 4, pp. 577–601, Aug. 2014.

[47] R. Singh, S. Dasgupta, and S. Sinha, "Extreme variability in convergence to structural balance in frustrated dynamical systems," *Europhys. Lett.*, vol. 105, no. 1, 2014, Art. ID 10003.

[48] E. Tereanu, M. A. Tuladhar, and M. A. Simone, *Structural Balance Targeting and Output Gap Uncertainty*, Int. Monetary Fund, Washington, DC, USA, 2014.

[49] G. Facchetti, G. Iacono, and C. Altafini, "Exploring the low-energy landscape of large-scale signed social networks," *Phys. Rev. E*, vol. 86, no. 3, 2012, Art. ID 036116.

[50] S. Gómez, P. Jensen, and A. Arenas, "Analysis of community structure in networks of correlated data," *Phys. Rev. E*, vol. 80, Jul. 2009, Art. ID 016114.

[51] A. Lancichinetti, S. Fortunato, and F. Radicchi, "Benchmark graphs for testing community detection algorithms," *Phys. Rev. E*, vol. 78, no. 4, 2008, Art. ID 046110.

[52] S. Kropivnik and A. Mrvar, "An analysis of the Slovene parliamentary parties network," in *Developments in Statistics and Methodology*, A. Ferligoj and A. Kramberger, Eds., Metodološki zvezki, Ljubljana: FDV, 1996, pp. 209–216.

[53] K. E. Read, "Cultures of the central highlands, New Guinea," *Southwest. J. Anthropol.*, vol. 10, no. 1, pp. 1–43, 1954.

[54] K. Oda, Y. Matsuoka, A. Funahashi, and H. Kitano, "A comprehensive pathway map of epidermal growth factor receptor signaling," *Mol. Syst. Biol.*, vol. 1, no. 1, pp. 1–17, 2005.

[55] K. Oda *et al.*, "Molecular interaction map of a macrophage," *AfCS Res. Rep.*, vol. 2, no. 14, pp. 1–12, 2004.

[56] R. Milo *et al.*, "Network motifs: Simple building blocks of complex networks," *Science*, vol. 298, no. 5594, pp. 824–827, 2002.

[57] S. S. Shen-Orr, R. Milo, S. Mangan, and U. Alon, "Network motifs in the transcriptional regulation network of Escherichia coli," *Nature Genet.*, vol. 31, no. 1, pp. 64–68, 2002.

[58] J. Leskovec and A. Krevl. (Jun. 2014). *SNAP Datasets: Stanford Large Network Dataset Collection*. [Online]. Available: http://snap.stanford.edu/data

[59] L. Danon, A. Diaz-Guilera, J. Duch, and A. Arenas, "Comparing community structure identification," *J. Stat. Phys.*, vol. 78, no. 9, 2005, Art. ID P09008.

[60] M. Gong, L. Ma, Q. Zhang, and L. Jiao, "Community detection in networks by using multiobjective evolutionary algorithm with decomposition," *Phys. A*, vol. 391, no. 15, pp. 4050–4060, 2012.
[61] E. Zitzler and L. Thiele, "Multiobjective optimization using evolutionary algorithms—A comparative case study," in *Parallel Problem Solving From Nature*. Berlin, Germany: Springer, 1998, pp. 292–301.
[62] Q. Zhang, W. Liu, and H. Li, "The performance of a new version of MOEA/D on CEC09 unconstrained MOP test instances," in *Proc. IEEE Congr. Evol. Comput.*, Trondheim, Norway, 2009, pp. 203–208.

Qing Cai received the B.S. degree in electronic information engineering from Wuhan Textile University, Wuhan, China, in 2010. He is currently working toward the Ph.D. degree in pattern recognition and intelligent systems with the School of Electronic Engineering, Xidian University, Xi'an, China.

His research interests include computational intelligence, complex network analytics, and recommender systems.

Maoguo Gong (M'07–SM'14) received the B.S. degree in electronic engineering and the Ph.D. degree in electronic science and technology from Xidian University, Xi'an, China, in 2003 and 2009, respectively.

Since 2006, he has been a Teacher with Xidian University, where he was promoted to be an Associate Professor and a Full Professor, both with exceptive admission, in 2008 and 2010, respectively. He is leading or has completed ten projects, as the principal investigator, funded by the National Natural Science Foundation of China and the National High Technology Research and Development Program (863 Program) of China. His research interests include computational intelligence with applications to optimization, learning, data mining, and image understanding. He has published over 50 papers in journals and conferences, and holds 12 granted patents as the first inventor.

Dr. Gong received the prestigious National Top Young Talent of China by the Central Organization Department of China, the Excellent Young Scientist Foundation by the National Natural Science Foundation of China, and the New Century Excellent Talent in University by the Ministry of Education of China. He is the Vice Chair of the IEEE Computational Intelligence Society Task Force on Memetic Computing, an Executive Committee Member of the Chinese Association for Artificial Intelligence, and a Senior Member of the Chinese Computer Federation.

Shasha Ruan received the B.S. degree in electronic information science and technology engineering from Henan University of Science and Technology, Henan, China, in 2013. She is currently working toward the M.S. degree in pattern recognition and intelligent systems with the School of Electronic Engineering, Xidian University, Xi'an, China.

Her research interests include multiobjective optimization and complex network analytics.

Qiguang Miao received the M.Eng. and Ph.D. degrees in computer science from Xidian University, Xi'an, China.

He is a Professor with the School of Computer Science and Technology, Xidian University. His research interests include intelligent image processing and multiscale geometric representations for images.

Haifeng Du received the B.S. degree in fluid transmission and control and the Ph.D. degree in mechatronics engineering from Xi'an Jiaotong University, Xi'an, China, in 1996 and 2001, respectively.

He is a Professor with the School of Public Policy and Management, Xi'an Jiaotong University. His research interests include public management and social network analytics.