Optimal Disintegration Strategy With Heterogeneous Costs in Complex Networks

Ye Deng, Jun Wu, Yu Xiao, Mengxiang Zhang, *Student Member, IEEE*, Yang Yu, and Yang Zhang

***Abstract*—Recently, research in the field of network disintegration, which includes controlling the spread of disease and collapsing terrorist organizations, has found broad applications and attracted increased attention. In this paper, we focus on the network disintegration with heterogeneous cost, in which there may be unequal disintegration costs associated with deleting different nodes. First, we present a cost model for a disintegration strategy with both cost-sensitive and cost-constraint parameters in complex networks. Then, we propose an optimization model for the disintegration strategy with heterogeneous cost and introduce the genetic algorithm to identify the optimal disintegration strategy. Extensive experiments in synthetic and real-world networks suggest that the heterogeneity of the disintegration cost and the tightness of the cost constraint significantly affect the optimal disintegration strategy. We demonstrate that, in contrast to the classical hub node strategy, low-cost nodes play a key role in the optimal disintegration strategies if the cost constraint is tight and the disintegration cost is strongly heterogeneous.**

***Index Terms*—Complex networks, heterogeneous cost, network disintegration, optimal disintegration strategy.**

I. Introduction

COMPLEX networks describe a wide range of systems in nature and society, such as the Internet, power grids, citation networks, epidemic spreading networks, terrorist networks, and rumor-spreading networks [1]–[3]. Most networks are beneficial, and the goal is to preserve their function. Many researchers have focused on designing ways to increase the survivability of such networks [4]–[7]. The original motivation for this paper is to determine how to collapse a network that may be detrimental, such as immunizing residents or communication network to prevent the spread of disease or distribution of a computer virus. Other examples include destabilizing terrorist networks [8], avoiding financial crises [9], controlling the spread of rumors [10], and interrupting cancer networks [11].

Although current research on network disintegration has received less attention than research on network protection, many related studies have contributed to the problem of network disintegration strategies. Early disintegration strategies were related to specific structural metrics of nodes or edges. The most classical strategy is the degree-based strategy [4]. Previous studies have shown that scale-free networks are extremely fragile when they suffer deliberate attacks, in which nodes are deleted by the descending sequence of degree. Considering that the degree is a local property of a node, the betweenness centrality was introduced as a new criterion [12], [13]. Other global metrics were also used to identify nodes that are vital for network disintegration [14], such as the coreness [15] and subgraph centrality [16]. Other methods have been proposed to achieve a better disintegration effect. For example, Chen *et al.* [17] introduced the "equal graph partitioning" immune method. The basic purpose of such an approach is to collapse the network into several components of identical size. Schneider *et al.* [18] introduced an immune method to obtain an optimal susceptible size, whose immune effect is much better than the highest-betweenness strategy. Deng *et al.* [19] developed an optimal disintegration strategy algorithm which is based on the intelligent optimization algorithm to seek a near-global optimal solution. Also, the imperfect information strategy has received increased attention. For instance, Dezső and Barabási [20] introduced a therapeutic method to prevent the spread of a virus with indefinite information, where the possibility of starting treatment to the infested person depends on the corresponding value of the node attribute. Jun *et al.* [21] focused on the disintegration optimization method with imperfect information, which indicates that only a subset of the information of the network can be obtained. Tan *et al.* [22] introduced link prediction into the network disintegration problem using the information of partial nodes.

However, concerning current research progress, most studies about the network disintegration problem assumed that the disintegration cost of deleting each node or edge is identical. Therefore, the total disintegration cost is only limited by the number of nodes that are deleted. In fact, the disintegration cost corresponding to deleted nodes may be heterogeneous for many reasons. In general, key nodes imply a greater disintegration cost. For instance, the expenditure to assassinate

Manuscript received January 15, 2018; accepted April 18, 2018. Date of publication May 11, 2018; date of current version July 17, 2020. This work was supported in part by the National Natural Science Foundation of China under Grant 71371185 and Grant 71690233, and in part by the Program for New Century Excellent Talents in University under Grant NCET-12-0141. This paper was recommended by Associate Editor W. Shen. *(Corresponding author: Jun Wu.)*

Y. Deng, J. Wu, Y. Xiao, Y. Yu, and Y. Zhang are with the College of Systems Engineering, National University of Defense Technology, Changsha 410073, China (e-mail: junwu@nudt.edu.cn).

M. Zhang is with the Business School, Sichuan University, Chengdu 610064, China.

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/TSMC.2018.2832238

Osama Bin Laden should be significantly higher than the case for someone with a lower profile. In this paper, we focus on the optimal disintegration problem based on heterogeneous costs. In the previous works, nodes with the highest degree or betweenness would be preferentially deleted to achieve a more destructive effect. Nonetheless, based on the assumption that the disintegration cost is heterogeneous, the hub node strategy may not be the optimal disintegration strategy among other strategies with cost constraints. The aim is, therefore, to determine how to balance the disintegration effect against the disintegration cost, and to identify the optimal disintegration strategy.

The rest of this paper is structured as follows. In the next section, a cost model is presented for the disintegration strategy. In Section III, an optimization model is proposed for the disintegration strategy. Then, solutions that are based on genetic algorithms (GA) are discussed in Section IV. Experiments to determine the optimal disintegration strategy in synthetic and real-world networks are presented in Sections V and VI, respectively. Finally, discussion and the conclusions are presented in Section VII.

II. Cost Model for Disintegration Strategy in Complex Networks

Complex networks can be described as an undirected graph $G(V, E)$ with node set V and edge set $E \subseteq V \times V$. The number of nodes and edges are denoted by $N = |V|$ and $W = |E|$, respectively. The adjacency matrix of G is denoted by $A(G) = (a_{ij})_{N\times N}$, where $a_{ij} = a_{ji} = 1$ once v_i and v_j are connected. Then, we denote the degree of node v_i by k_i, whose value is equivalent to the number of edges that are connected to node v_i.

In this paper, disintegration methods only consider the node removal strategy, and it is assumed that the connected edges will be deleted once a node is deleted. The disintegration cost of corresponding node v_i is denoted by c_i. We assume that the disintegration cost c_i is a function of node attribute r_i corresponding to node v_i as follows:

$$c_i = r_i^p \tag{1}$$

where $p \geq 0$ is called the cost-sensitive parameter. The node attribute r_i could be regarded as the degree, the betweenness, etc. In more detail, the disintegration cost c_i is homogeneous when $p = 0$, which means that the disintegration costs of each node are identical. A larger value of the cost-sensitive parameter implies that such a disintegration cost is more sensitive.

In most instances, the total disintegration cost has an upper and lower bound. We define the cost constraint as follows:

$$\hat{C} = \alpha \sum_{i=1}^{N} c_i = \alpha \sum_{i=1}^{N} r_i^p \tag{2}$$

where $\alpha \in [0, 1]$ is called the cost-constraint parameter. A larger value of the cost-constraint parameter α means that the constraint of the disintegration cost is more flexible. In extreme cases, none of the nodes could be deleted when $\alpha = 0$, and all of the nodes belonging to the network could be deleted when $\alpha = 1$.

III. Optimization Model for Disintegration Strategy With Heterogeneous Cost

We denote the set of removed nodes by $\hat{V} \subseteq V$, and the network after disintegration by $\hat{G} = (V - \hat{V}, \hat{E})$. Let $n = |\hat{V}|$ be the disintegration strength. The disintegration strategy is denoted by $X = [x_1, x_2, \ldots, x_N]$, where $x_i = 1$ if $v_i \in \hat{V}$; otherwise, $x_i = 0$. Then, we obtain

$$n = \sum_{i=1}^{N} x_i. \tag{3}$$

The cost of disintegration strategy X is denoted by $C_X = \sum_{v_i \in \hat{V}} c_i$. It can be easily seen that

$$C_X = \sum_{v_i \in \hat{V}} c_i = \sum_{i=1}^{N} x_i c_i = \sum_{i=1}^{N} x_i r_i^p. \tag{4}$$

The measure function of the network performance is denoted by Γ. We assume that if $G_1 = (V_1, E_1)$ is a subgraph of $G_2 = (V_2, E_2)$, i.e., $V_1 \subseteq V_2$ and $E_1 \subseteq E_2$, then $\Gamma(G_1) \leq \Gamma(G_2)$. This monotonicity assumption ensures that the network performance reduces with the process of network disintegration. We define the effect of the disintegration strategy as the degradation of the network performance after node removal $\Phi(X) = \Gamma(G) - \Gamma(\hat{G}) \geq 0$. The optimization target focuses on seeking the optimal disintegration strategy, X^*, that can achieve the optimal disintegration effect. Thus, the optimization model for the disintegration strategy is derived in the following form:

$$\max \ \Phi(X = [x_1, x_2, \ldots, x_N]) \tag{5}$$

$$\text{s.t.} \ \begin{cases} C_X \leq \hat{C} \\ x_i = 0 \text{ or } 1, i = 1, 2, \ldots, N. \end{cases} \tag{6}$$

IV. Solutions Based on Genetic Algorithm

The optimization model for the disintegration strategy belongs to a zero-one integer programming problem. There are some precise mathematical programming techniques, such as the branch and bound algorithm [23], which are used to solve the zero-one integer programming problem. However, for these precise mathematical programming techniques, an objective function with an explicit form is generally required. In this paper, we use $\Phi(X) = \Gamma(G) - \Gamma(\hat{G})$ as the objective function, where Γ represents the relative size of the largest connected component of a network. Although we can calculate Γ given the disintegration strategy X, there is no explicit form for Γ as a function of X. Thus, the exact mathematical programming techniques are not feasible for the problem of network disintegration.

We consider a simple example, where G is an undirected graph with 100 nodes. Although the network contains few nodes, we need to compare 2^{100} solutions to obtain the optimal solution using the method of exhaustion. The number of feasible solutions is less than 2^{100} because of the cost-constraint model; however, the solution space is still extremely large.

This would be wasteful and impossible to implement for large-scale problems. Therefore, we plan to apply GA [24] to solve this zero-one integer programming problem to achieve an approximate optimum solution. In the next section, we propose a complete description of the optimal disintegration-strategy algorithm based on GA. The algorithm procedures are given below. The four primary parts of the procedure are described as coding and decoding, crossover and mutation, fitness function, and selection.

A. Coding and Decoding

We choose the symbol of disintegration strategy $X = [x_1, x_2, \ldots, x_N]$ (1-D binary code) as the coding scheme of chromosome. Denote by $1 \times 2 \times N$ the chromosome

$$x_i = \begin{cases} 1 & \text{if the node is removed} \\ 0 & \text{otherwise} \end{cases} \tag{7}$$

where $i = 1, 2, \ldots, n$. Based on the coding scheme and the fitness function, we design the decoding of chromosome to translate $X = [x_1, x_2, \ldots, x_N]$ 1-D code into the adjacent matrix G to calculate the fitness function.

B. Crossover and Mutation

A crossover operation between parent chromosomes is a natural procedure, and it is customarily set from 0.6 to 1.0. In a crossover, an offspring is obtained by the interchange of parents' information. On the contrary, a mutation is a scarce procedure which represents an abrupt change within an offspring. A mutation could be accomplished through choosing a chromosome among the population and altering a part of its information randomly. The largest advantage of such operation is that it stochastically brings new information into the algorithm procedure, possibly preventing from being trapped in the local optimum. The value of mutation is often less than 0.1. Here, we discuss the crossover and mutation strategy in a simple GA: one-point crossover and bit mutation.

C. Fitness Function

The fitness function should be designed to reflect the goal of the optimal disintegration strategy with the heterogeneous cost problem and the basis for the selection problem. Therefore, we establish the fitness function as follows:

$$F = \Gamma(G) - \Gamma(\hat{G}) \geq 0. \tag{8}$$

As the effect of the disintegration strategy, the higher value of fitness function, the more destructive the disintegration strategy will be.

D. Selection

In this GA, the roulette-wheel selection is chosen as the selection operator [25]. To generate the next generation, the selection of individuals from the current population is based on a probability that is directly proportional to their fitness values. The probabilities of selecting a parent can be seen as spinning a roulette wheel with the size of the segment for each parent being proportional to its fitness.

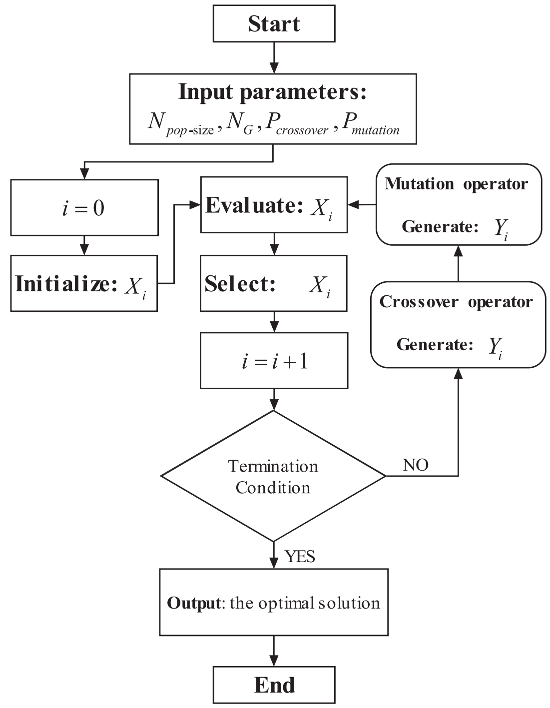


Fig. 1. Algorithm diagram of the optimal disintegration strategy.

Algorithm 1 Optimal Disintegration Strategy With Heterogeneous Cost

Input ODS problem data, GA parameter: $N_{pop-size}$, N_G, $P_{crossover}$ and $P_{mutation}$;
Output The optimal solution;
$i = 0$; //i:generation number
initialize X_i; //X_i:population
evaluate X_i using the fitness function F and keep the best solution;
while (Not terminating condition) **do**
  generate Y_i from X_i using the crossover operator; //Y_i:offspring
  generate Y_i from X_i using the mutation operator;
  evaluate Y_i using the fitness function F and keep the best solution;
  select X_{i+1} from X_i and Y_i using the roulette-wheel selection;
  $i = i + 1$
end while
Output The optimal solution;

Each step of the algorithm is described below, and the algorithm diagram of the optimal disintegration strategy is also shown in Fig. 1.

Step 1: Input ODS problem data and GA parameters: population size $N_{\text{pop-size}}$, generation N_G, crossover probability $P_{\text{crossover}}$, and mutation probability P_{mutation}.
Step 2: Initialize $N_{\text{pop-size}}$ chromosomes.
Step 3: Compute the fitness of the chromosome based on the fitness function F.
Step 4: Select the chromosomes by running the roulette-wheel selection.

TABLE I
PERCENTAGE OF REMOVED NODES θ WITH $d_i \leq \bar{d}$ AND $d_i > \bar{d}$ IN THE OPTIMAL DISINTEGRATION STRATEGIES IN A RANDOM SCALE-FREE NETWORK WITH DEGREE DISTRIBUTION $p(k) = (\lambda - 1)m^{\lambda-1}k^{-\lambda}$, WHERE $N = 100$, $\lambda = 2.5$, AND $m = 2$

p	α=0.2		α=0.4		α=0.6		α=0.8	
	$\theta_{d_i \leq \bar{d}}$	$\theta_{d_i > \bar{d}}$	$\theta_{d_i \leq \bar{d}}$	$\theta_{d_i > \bar{d}}$	$\theta_{d_i \leq \bar{d}}$	$\theta_{d_i > \bar{d}}$	$\theta_{d_i \leq \bar{d}}$	$\theta_{d_i > \bar{d}}$
0	63.2%	36.8%	68.3%	31.7%	80.0%	20.0%	82.0%	18.0%
0.5	92.0%	8.0%	85.0%	15.0%	82.5%	17.5%	82.6%	17.4%
1.0	97.1%	2.9%	87.8%	12.2%	83.32%	16.7%	85.9%	14.1%
1.5	100.0%	0.0	91.4%	8.6%	84.1%	15.9%	86.9%	13.1%
2.0	100.0%	0.0	92.3%	7.7%	85.9%	14.1%	91.0%	9.0%

Step 5: Update the chromosomes using crossover and mutation operations based on the crossover and mutation probability $P_{\text{crossover}}$ and P_{mutation}, respectively.

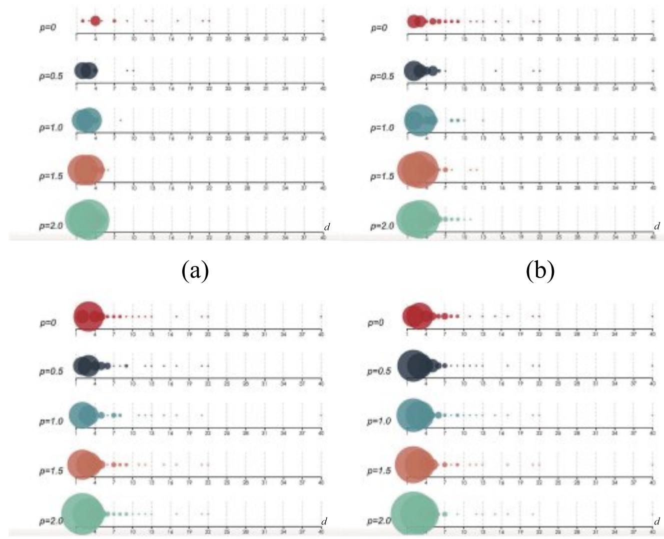
Step 6: Check the termination condition; if false, go to step 3; otherwise, go to step 7.

Step 7: Return the best chromosome as the optimal solution, let this chromosome be the solution of the optimal disintegration strategy because some applications need the decision to be made in network disintegration, which is the trade-off between the cost and disintegration effect. The pseudocode of the algorithm is given as Algorithm 1.

V. EXPERIMENTS IN SCALE-FREE NETWORKS

In this section, we analyze the optimal disintegration strategies that are based on the optimization model and algorithm introduced above. We also generate random scale-free networks with degree distributions using the configuration model [2], and we use the node degree k_i as the referential attribute r_i and the relative size of the largest connected component as the measure function of the network performance Γ. Then, the calibration of GA is in accordance with the results of preliminary experiments that were conducted to observe and compare the behavior of the algorithm for different parameter settings [26], [27]. In this case, the GA parameters are as follows: population size: $N_{\text{pop-size}} = 50$, generation: $N_G = 1000$, crossover probability: $P_{\text{crossover}} = 0.4$, and mutation probability: $P_{\text{mutation}} = 0.1$. We used MATLAB 2017a to run the simulations of scale-free networks on a PC with an Intel Core i5-3470 CPU having 3.2 GHz and 4.0 GB of RAM.

To display the optimal disintegration strategies, we first show the proportion of removed nodes with different degrees in the optimal disintegration strategies in Fig. 2. The size of each circle is proportional to the number of removed nodes whose degree belongs to the corresponding domain. We find that for the homogeneous cost ($p = 0$), there are many hub nodes with high degree in the optimal disintegration strategies, which corresponds to the classical hub-node strategies. However, in the case of heterogeneous cost with $p > 0$, low-cost nodes play a key role in the optimal disintegration strategies. In particular, considering the tight constraint of the disintegration cost, for example, $\alpha = 0.2$, most removed nodes are low-degree nodes. Furthermore, we calculated the percentage of removed nodes θ whose degree are below the average degree $d_i \leq \bar{d}$ and the percentage of removed nodes whose degree are above the average degree $d_i > \bar{d}$. In Table I, the results indicate that $\theta_{d_i \leq \hat{d}}$ increases gradually with increasing value of p, so the low-degree nodes become an increasingly important part of the optimal disintegration strategy. For example, the values of $\theta_{d_i \leq \hat{d}}$ for the parameters $\alpha = 0.4$ in Table I are 68.3%, 85.0%, 87.8%, 91.4%, and 92.3%, which add up to almost 100%. As shown in Table I, we obtained similar results in other parameter settings.



(a) (b) (c) (d)

Fig. 2. Proportion of removed nodes with different degrees in the optimal disintegration strategies in a random scale-free network with degree distribution $p(k) = (\lambda - 1)m^{\lambda-1}k^{-\lambda}$, where $N = 100$, $\lambda = 2.5$, and $m = 2$. The size of each circle is proportional to the number of removed nodes whose degree belongs to the corresponding domain.

To investigate the impact of the heterogeneity of the disintegration cost and the tightness of the cost constraint on the tendency of optimal disintegration strategies, we show the average degree of removed nodes $\hat{d}$ in the optimal disintegration strategies as functions of the cost-sensitive parameter p

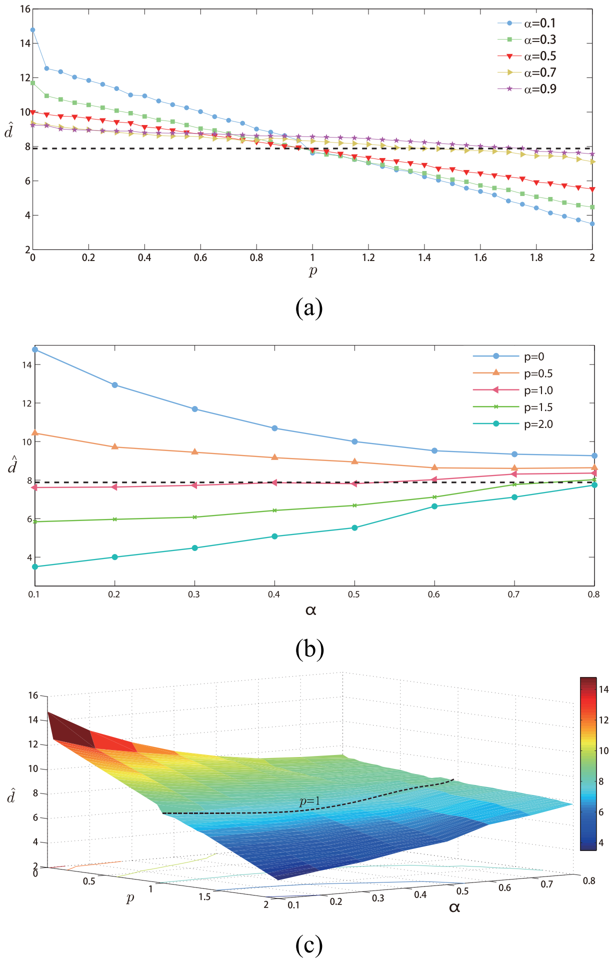


Fig. 3. Average degree of removed nodes $\hat{d}$ in the optimal disintegration strategies as a function of the (a) cost-sensitive parameter, the (b) cost-constraint parameter, and both, in the (c) 3-D surface plot. The original network is the same as that used in Fig. 2. The dotted lines in (a) and (b) are the reference lines, which correspond to the average degree of the original network. The dotted lines in (c) represent the critical line with $p = 1$.

and the cost-constraint parameter α in Fig. 3. From Fig. 3(a), we see that, whether the cost constraint is loose or tight, $\hat{d}$ decreases monotonically with p. This indicates that the optimal disintegration strategies consist more low-degree nodes as the heterogeneity of disintegration cost increases. From Fig. 3(b), we see that, whether the disintegration cost is homogeneous or heterogeneous, $\hat{d}$ converges to the average degree $\bar{d}$ of the original network. This indicates that when the cost constraint is sufficiently loose, the difference between various cost-sensitive parameters will gradually disappear. Moreover, from Fig. 3(a) and (b), we find that, if the disintegration cost c_i is linearly related to the degree d_i ($p = 1$), $\hat{d}$ remains almost constant with α. If the disintegration cost c_i is sublinearly related to the degree d_i ($p < 1$), $\hat{d}$ is greater than the average degree of the original network, and decreases with α. If the disintegration cost c_i is superlinearly related to the degree d_i ($p > 1$), $\hat{d}$ is less than the average degree of the original network and increases with α. From Fig. 3(c), we observe that there is a watershed at $p \approx 1$. For the case of weak heterogeneity of disintegration cost ($p < 1$), the surface drops gradually as α increases. For the case of strong heterogeneity

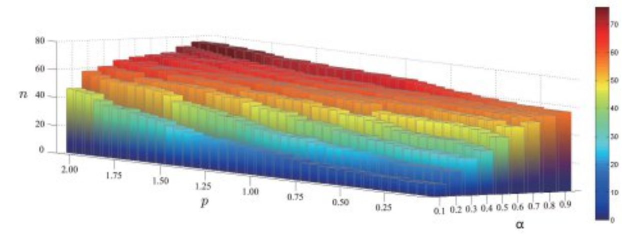


Fig. 4. Number of removed nodes in the optimal disintegration strategies with various cost-sensitive parameters and cost-constraint parameters. The deeper color indicates a higher value of n, and each concrete value n is shown at the lump. The original network is the same as that used in Fig. 2.

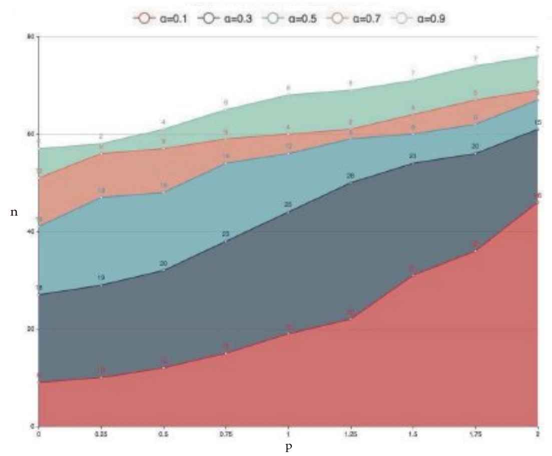


Fig. 5. Specific data of the number of removed nodes in the optimal disintegration strategies with cost-constraint parameter $\alpha = 0.1, 0.3, 0.5, 0.7$, and 0.9. We also emphasize the increase in the number of removed nodes, and the original network is the same as that used in Fig. 2.

of disintegration cost ($p > 1$), the surface rises gradually as α increases.

To analyze the strategy tendency in detail, we calculate the number of removed nodes n with various cost-sensitive parameters and cost-constraint parameters and show the 3-D histogram in Fig. 4. We observe a clear upward trend of n for large values of p and α. These results indicate that the more heterogeneous the disintegration cost and the looser the cost constraint, more nodes will be removed in the optimal disintegration strategy. The specific data are also presented in Fig. 5, where we emphasize the increase in the number of removed nodes with various values of α and p. We find that when α is set as 0.1, 0.3, and 0.5, the values of the disintegration strength n increases rapidly with increasing values of p. Combined with the preceding analysis that the average degree $\hat{d}$ decreases monotonically with p, it is in good agreement with the previous results that the optimal disintegration strategies consist of more low-degree nodes as the heterogeneity of disintegration cost increases.

VI. Experiments in Real Network

A. Data Description

The study of network disintegration is important for many real-world systems such as rumor spreading in online social networks and terrorist network disintegration. To verify the feasibility of our method, we implemented

TABLE II
BASIC STATISTICS OF REAL NETWORKS

Node ID	Degree	Closeness Centrality	Betweenness Centrality	Pagerank	Clustering Coefficient	Number of Triangles	Eigenvector Centrality
1	29	0.5833	392.3521	0.0486	0.3325	135	1
2	2	0.3727	0	0.0051	1	1	0.1448
3	27	0.5727	320.8381	0.0450	0.3675	129	0.9705
4	10	0.45	0	0.0176	1	45	0.4524
5	10	0.45	0	0.0176	1	45	0.4524
6	7	0.4344	61.2699	0.0154	0.1904	4	0.1981
7	22	0.5	267.6658	0.0397	0.3679	85	0.7410
8	6	0.3987	12.6291	0.0130	0.6	9	0.2351
9	4	0.3888	0	0.0089	1	6	0.1786
10	2	0.3405	0	0.0057	1	1	0.0726
11	18	0.4736	91.6632	0.0309	0.5294	81	0.7047
12	10	0.4090	0	0.0168	1	45	0.4713
13	10	0.4090	0	0.0168	1	45	0.4713
14	10	0.4090	0	0.0168	1	45	0.4713
15	11	0.4117	13.1624	0.0191	0.8363	46	0.4779
16	11	0.4117	51.5214	0.0194	0.8181	45	0.4749
17	16	0.4666	58.6457	0.0269	0.6166	74	0.6529
18	15	0.4736	175.8186	0.0302	0.4666	49	0.4888
19	6	0.3684	17.2088	0.0144	0.4666	7	0.0824
20	8	0.3230	91.5437	0.0224	0.5	14	0.0439
21	5	0.375	27.8841	0.0122	0.6	6	0.0741
22	3	0.3662	39.4486	0.0084	0.3333	1	0.0511
23	16	0.4809	232.2502	0.0315	0.45	54	0.4894
24	8	0.4012	147.2830	0.0184	0.4285	12	0.1963
25	4	0.2971	0	0.0113	1	6	0.0318
26	11	0.4632	14.1952	0.0193	0.8727	48	0.4859
27	12	0.4736	25.6789	0.0207	0.7878	52	0.5513
28	14	0.4960	246.3920	0.0274	0.5384	49	0.4798
29	3	0.3387	0	0.0077	1	3	0.0791
30	12	0.4598	52.7410	0.0220	0.7272	48	0.4660
31	5	0.3620	25.2415	0.0127	0.5	5	0.0921
32	5	0.2490	0	0.01390	1	10	0.0274
33	5	0.3559	1.75	0.0118	0.8	8	0.0977
34	5	0.3559	1.75	0.0118	0.8	8	0.0977
35	4	0.2971	0	0.0113	1	6	0.0318
36	2	0.3333	1.125	0.0059	0	0	0.0515
37	11	0.4883	268.9847	0.0197	0.6727	37	0.4798
38	17	0.5080	63.4107	0.0283	0.6029	82	0.7231
39	1	0.3264	0	0.0040	0	0	0.0372
40	1	0.3230	0	0.0040	0	0	0.0368
41	10	0.4468	6.0294	0.0176	0.8222	37	0.4691
42	1	0.3028	0	0.0045	0	0	0.0117
43	1	0.3230	0	0.0038	0	0	0.0514
44	4	0.3962	0	0.0084	1	6	0.2156
45	1	0.3230	0	0.0040	0	0	0.0368
46	1	0.2451	0	0.0047	0	0	0.0061
47	1	0.3351	0	0.0038	0	0	0.0546
48	2	0.3405	0.3928	0.0056	0	0	0.0632
49	4	0.2971	0	0.0113	1	6	0.0318
50	6	0.3559	46.4912	0.0156	0.4	6	0.0864
51	8	0.4285	0	0.0140	1	28	0.4212
52	10	0.4090	0	0.0168	1	45	0.4713
53	2	0.3368	0	0.0053	1	1	0.0896
54	2	0.3058	1.25	0.0060	0	0	0.0428
55	13	0.4532	164.2050	0.0258	0.4487	35	0.4855
56	1	0.3333	0	0.0040	0	0	0.0361
57	11	0.4315	448.5200	0.0280	0.2181	12	0.1191
58	6	0.315	9.9140	0.0161	0.9333	14	0.0381
59	6	0.315	9.9140	0.0161	0.9333	14	0.0381
60	6	0.315	9.9140	0.0161	0.9333	14	0.0381
61	6	0.315	9.9140	0.0161	0.9333	14	0.0381
62	2	0.315	0	0.0070	1	1	0.0398
63	2	0.315	0	0.0070	1	1	0.0398
64	2	0.3073	0	0.0066	1	1	0.0200

experiments in the terrorist network of the Madrid train bombing (MTB) with $N = 64$ nodes and $W = 243$ edges. The data can be downloaded from http://konect.uni-koblenz.de/. Basic statistics of the network are shown in Table II.

International terrorism remains a grave threat to politics, economics, culture, religion, and society. While all forms of terrorism result in tragedies, the MTB was driven by extremists and resulted in many casualties and significant property loss. The MTB was organized by Al-Qaeda, and it involved many terrorist organizations to constitute a global terrorist network. In a purely topological model, a terrorist network is considered to be a network which is composed of nodes (individual terrorists) and connected by edges (possible connections between

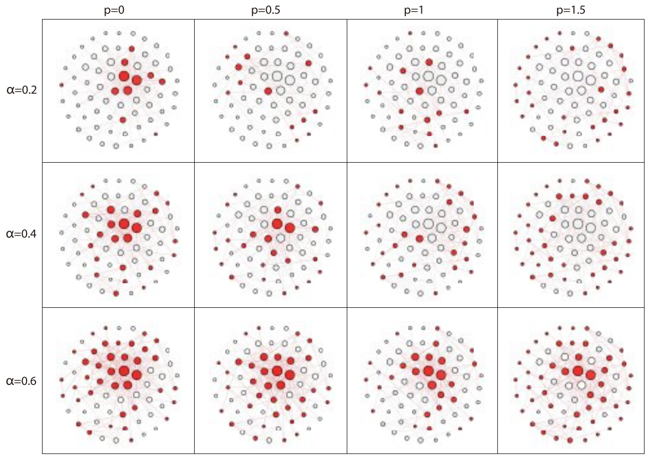


Fig. 6. Visualization of optimal disintegration strategies in MTB terrorist network with various cost-sensitive parameters and cost-constraint parameters. The size of each node is proportional to its degree. The red solid circles indicate the removed nodes.

these individuals). Experiments using terrorist network may help authorities to develop efficient and effective disintegration strategies and measures [28].

B. Result

Here, we used the node degree k_i as the referential attribute r_i and the relative size of the largest connected component as the fitness function of the network performance. The modeling software and the hardware are the same as experiments in the model network. In this part, the number of decision variables is 64, and the number of equality constraint is 1. We also initialize $N_{\text{pop-size}}$ chromosomes randomly, which satisfy the cost constraint.

We visualized the optimal disintegration strategy for the real-world network in Fig. 6. The size of each node is proportional to its degree. The high-degree nodes are put in the core of each layout, and low-degree nodes are put in the periphery of each layout. In Table III, we also show the node ID of removed nodes with various cost-sensitive parameters α and cost-constraint parameters p. Combined with the basic statistics in Table II, we observe some meaningful results.

1) The more heterogeneous the disintegration cost and the looser the cost constraint, the more terrorists will be killed.
2) In the case of the homogeneous cost ($p = 0$), the anti-terrorism operations follow the classical hub-node strategy. It is well known that high-degree terrorists are killing preferentially.
3) If the disintegration cost c_i is sublinearly related to the degree d_i ($p < 1$), as shown in Table II, the removed nodes set $\hat{V}$ always contains some hub nodes, e.g., nodes 1, 3, and 7, who have a high value of degree centrality and betweenness centrality, so that killing the key members should be considered in the action plan on combating terrorism, such as Jamal Zougam (node ID: 1), who was suspected of involvement in over 2000 charges of murder, and has been given a 42 922-year sentence.

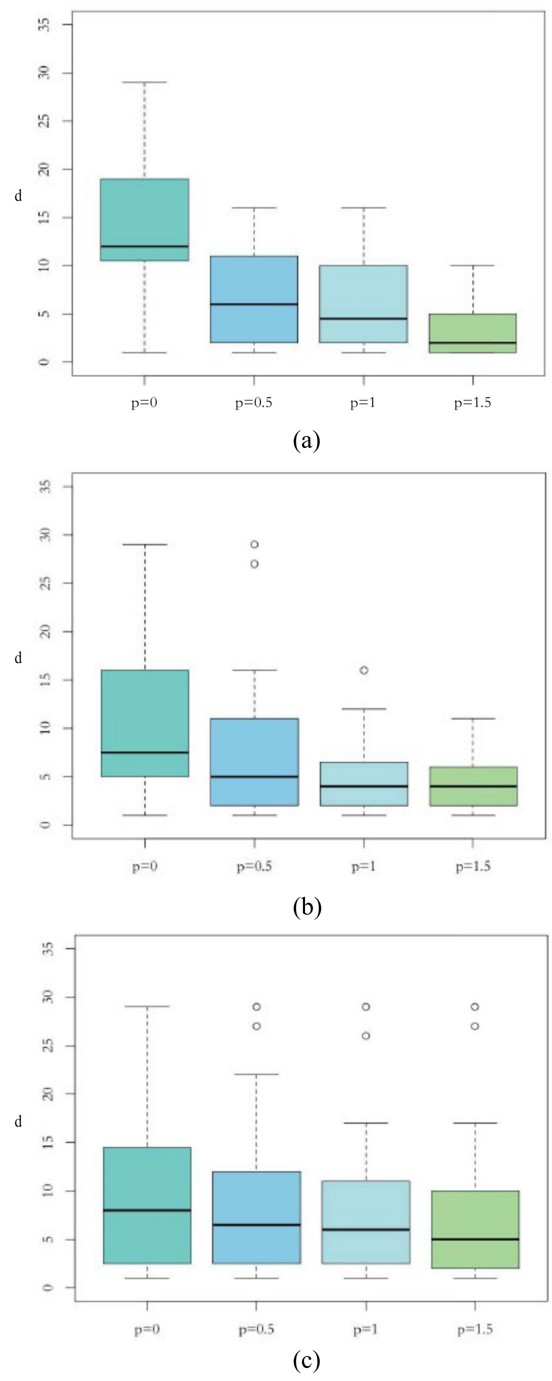


Fig. 7. Box plot of the degree of removed nodes. Cost-constraint parameter (a) $\alpha = 0.2$, (b) $\alpha = 0.4$, and (c) $\alpha = 0.6$. For each box plot, the top bar is the maximum degree between removed nodes, the lower bar is the minimum among removed nodes, and the middle bar is the median value. The original network is the same as that used in Fig. 2.

4) If the disintegration cost c_i is superlinearly related to the degree d_i ($p > 1$), the "small potatoes" would play a key role in the optimal disintegration strategies, e.g., nodes 2 and 8–10, which suggests that the core of anti-terrorism operations comprises marginal individuals in the organization, except for sufficiently loose cost constraints.

TABLE III
NODE ID OF PROPORTION OF REMOVED NODES WITH VARIOUS COST-SENSITIVE PARAMETERS α AND COST-CONSTRAINT PARAMETERS p IN REAL NETWORK EXPERIMENTS

p	ID of Removed Nodes	Computing time(s)
α=0.2, p=0	1,3,7,12,18,23,27,30,46,57,63	1.693
α=0.2, p=0.5	8,9,22,23,26,31,33,35,39,41,46,47,54,55	1.708
α=0.2, p=1.0	16,18,19,22,23,31,42,43,51,53,57,60,61,63	1.828
α=0.2, p=1.5	2,8,9,10,14,21,22,24,25,29,31,32,36,40,42,42,43,44,45,46,47,48,49,50,54,56,58,62,63	1.798
α=0.4, p=0	1,3,6,7,8,11,17,19,20,23,25,26,37,38,39,45,47,48,49,50,55,57,60,61	1.708
α=0.4, p=0.5	1,3,6,15,18,22,23,27,29,31,36,37,44,47,49,51,53,54,59,62,63	1.783
α=0.4, p=1.0	2,4,5,6,8,10,23,24,27,29,31,34,35,36,37,43,44,45,47,49,50,53,54,56,59,60,63	1.750
α=0.4, p=1.5	2,4,5,6,8,9,10,12,13,16,19,21,22,25,29,31,33,34,35,36,39,40,42,43,44,45,46,47,48,49,50,53,54,56,59,60,62,63,64	1.816
α=0.6, p=0	1,2,3,5,6,7,10,11,12,13,16,17,18,21,22,23,23,25,26,33,34,37,38,40,41,43,44,45,48,50,51,53,54,55,57,58,59,61,62,63	1.764
α=0.6, p=0.5	1,3,4,7,8,9,11,15,16,17,18,19,22,23,25,26,32,34,36,37,38,39,40,42,43,44,45,48,50,51,52,53,55,56,57,59	1.827
α=0.6, p=1.0	1,3,4,7,10,11,12,14,16,17,18,22,24,28,32,35,36,38,40,42,45,46,49,51,57,61,63	1.793
α=0.6, p=1.5	1,2,3,4,6,9,10,12,13,14,16,18,19,20,21,22,24,25,26,28,30,31,32,33,35,36,38,39,40,42,43,45,46,47,48,49,51,53,54,56,57,58,59,61,62,63,64	1.779

To visually summarize and compare these groups of data, we show the box plot of the degree of removed nodes in Fig. 7. As shown in Fig. 7(a)–(c), we find that the average degree of removed nodes shows a declining trend. With increasing values of p, the values of the top bar also decrease dramatically, which indicates that the low-cost nodes take over the optimal disintegration strategies and gradually play a key role in the optimal disintegration strategies, especially for the situation in which there is a tight constraint of the disintegration cost.

VII. CONCLUSION

The problem of network disintegration intrinsically involves identifying a set of nodes or edges that should be removed to have a maximal disintegration effect. Seeking an optimal disintegration strategy among large numbers of alternative strategy sets with a limited budget is an important and challenging problem. Most works that are related to disintegration strategies in complex networks assumed that the cost to delete each node (edge) is uniform, while they neglected the crucial fact that the cost to delete each node (edge) may be heterogeneous. In this paper, we have addressed the problem of the optimal disintegration strategy with heterogeneous disintegration cost, where there are different costs to delete each node.

We presented a cost model for disintegration strategy in complex networks. In this model, we assumed that the disintegration cost of a node is a function of a specific referential attribute of the node, which can be controlled by the cost-sensitive parameter p. Then, we proposed an optimization model for the disintegration strategy with a heterogeneous cost and solved it based on GA. Finally, we explored the optimal disintegration strategies in both synthetic and real-world networks.

Results of extensive experiments performed in scale-free networks indicate that the heterogeneity of the disintegration cost and the tightness of the cost constraint significantly affect the optimal disintegration strategy. We observed a critical point of the cost-sensitive parameter, at which the disintegration cost c_i is linearly related to the degree d_i ($p = 1$). We demonstrated that the optimal disintegration strategies are determined by the cost-sensitive parameter p and the cost-constraint parameter α.

1) If the cost constraint is loose, there is no obvious difference between the optimal disintegration strategies with various cost-sensitive parameters.
2) If the cost constraint is tight and the disintegration cost is weakly heterogeneous ($p < 1$), we should remove a minority of hub nodes to achieve the optimal disintegration effect.
3) If the cost constraint is tight and the disintegration cost is strongly heterogeneous ($p > 1$), the optimal disintegration strategies should contain more low-degree nodes.

Experiments using data from a real terrorist network verify the feasibility and validity of our method. We believe that our finding will also help the decision-makers.

REFERENCES

[1] R. Albert and A.-L. Barabási, "Statistical mechanics of complex networks," *Rev. Mod. Phys.*, vol. 74, no. 1, pp. 47–97, 2002.
[2] M. E. J. Newman, "The structure and function of complex networks," *SIAM Rev.*, vol. 45, no. 2, pp. 167–256, 2003.
[3] X. F. Wang, "Complex networks: Topology, dynamics and synchronization," *Int. J. Bifurcat. Chaos.*, vol. 12, no. 5, pp. 885–916, 2002.
[4] R. Albert, H. Jeong, and A.-L. Barabási, "Error and attack tolerance of complex networks," *Nature*, vol. 406, no. 6794, pp. 378–382, 2000.
[5] B. Shargel, H. Sayama, I. R. Epstein, and Y. Bar-Yam, "Optimization of robustness and connectivity in complex networks," *Phys. Rev. Lett.*, vol. 90, no. 6, 2003, Art. no. 068701.
[6] E. Zio and G. Sansavini, "Vulnerability of smart grids with variable generation and consumption: A system of systems perspective," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 43, no. 3, pp. 477–487, May 2013.
[7] J. Wu, M. Barahona, Y.-J. Tan, and H.-Z. Deng, "Spectral measure of structural robustness in complex networks," *IEEE Trans. Syst., Man, Cybern. A, Syst., Humans*, vol. 41, no. 6, pp. 1244–1252, Nov. 2011.
[8] J. Raab and H. B. Milward, "Dark networks as problems," *J. Public Admin. Res. Theory*, vol. 13, no. 4, pp. 413–439, 2003.
[9] T. Kobayashi and K. Hasui, "Efficient immunization strategies to prevent financial contagion," *Sci. Rep.*, vol. 4, p. 3834, Jan. 2014.
[10] R. M. Tripathy, A. Bagchi, and S. Mehta, "A study of rumor control strategies on social networks," in *Proc. Int. Conf. Inf. Knowl. Manag.*, Toronto, ON, Canada, 2010, pp. 1817–1820.
[11] A. P. Quayle, A. S. Siddiqui, and S. J. M. Jones, "Preferential network perturbation," *Physica A Stat. Mech. Appl.*, vol. 371, no. 2, pp. 823–840, 2006.
[12] P. Holme, B. J. Kim, C. N. Yoon, and S. K. Han, "Attack vulnerability of complex networks," *Phys. Rev. E, Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top.*, vol. 65, no. 5, 2002, Art. no. 056109.
[13] M. Bellingeri, D. Cassi, and S. Vincenzi, "Efficiency of attack strategies on complex model and real-world networks," *Physica A Stat. Mech. Appl.*, vol. 414, pp. 174–180, Nov. 2014.
[14] L. Lü *et al.*, "Vital nodes identification in complex networks," *Phys. Rep.*, vol. 650, pp. 1–63, Sep. 2016.
[15] M. Kitsak *et al.*, "Identification of influential spreaders in complex networks," *Nat. Phys.*, vol. 6, no. 11, pp. 888–893, 2010.
[16] E. Estrada and J. A. Rodriguez-Velazquez, "Subgraph centrality in complex networks," *Phys. Rev. E, Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top.*, vol. 71, no. 5, 2005, Art. no. 056103.

[17] Y. Chen, G. Paul, S. Havlin, F. Liljeros, and H. E. Stanley, "Finding a better immunization strategy," *Phys. Rev. Lett.*, vol. 101, no. 5, 2008, Art. no. 058701.

[18] C. M. Schneider, T. Mihaljev, S. Havlin, and H. J. Herrmann, "Suppressing epidemics with a limited amount of immunization units," *Phys. Rev. E, Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top.*, vol. 84, no. 6, 2011, Art. no. 061911.

[19] Y. Deng, J. Wu, and Y.-J. Tan, "Optimal attack strategy of complex networks based on tabu search," *Physica A Stat. Mech. Appl.*, vol. 442, pp. 74–81, Jan. 2016.

[20] Z. Dezső and A.-L. Barabási, "Halting viruses in scale-free networks," *Phys. Rev. E, Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top.*, vol. 65, no. 5, 2002, Art. no. 055103.

[21] L. Jun, W. Jun, L. Yong, D. Hong-Zhong, and T. Yue-Jin, "Optimal attack strategy in random scale-free networks based on incomplete information," *Chin. Phys. Lett.*, vol. 28, no. 6, 2011, Art. no. 068902.

[22] S.-Y. Tan, J. Wu, L. Lü, M.-J. Li, and X. Lu, "Efficient network disintegration under incomplete information: The comic effect of link prediction," *Sci. Rep.*, vol. 6, Mar. 2016, Art. no. 22916.

[23] H. Crowder, E. L. Johnson, and M. Padberg, "Solving large-scale zero-one linear programming problems," *Oper. Res.*, vol. 31, no. 5, pp. 803–834, 1983.

[24] D. Whitley, "A genetic algorithm tutorial," *Stat. Comput.*, vol. 4, no. 2, pp. 65–85, 1994.

[25] B. L. Miller and D. E. Goldberg, "Genetic algorithms, selection schemes, and the varying effects of noise," *Evol. Comput.*, vol. 4, no. 2, pp. 113–131, Jun. 1996.

[26] H. Aytug and G. J. Koehler, "Stopping criteria for finite length genetic algorithms," *INFORMS. J. Comput.*, vol. 8, no. 2, pp. 183–191, 1996.

[27] P. C. Pendharkar and G. J. Koehler, "A general steady state distribution based stopping criteria for finite length genetic algorithms," *Eur. J. Oper. Res.*, vol. 176, no. 3, pp. 1436–1451, 2007.

[28] J. Qin, J. J. Xu, D. Hu, M. Sageman, and H. Chen, "Analyzing terrorist networks: A case study of the global salafi jihad network," in *Proc. Int. Conf. Intell. Security Informat.*, Atlanta, GA, USA, 2005, pp. 287–304.

Ye Deng received the B.S. degree in communication engineering from Sichuan University, Chengdu, China, in 2013, and the M.S. degree in management science and engineering from the National University of Defense Technology, Changsha, China, in 2015, where he is currently pursuing the Ph.D. degree.

His current research interests include complex networks, especially the disintegration strategy of complex networks.

Jun Wu received the B.S. degree in management science from Sichuan University, Chengdu, China, in 2002, and the Ph.D. degree in management science from the National University of Defense Technology, Changsha, China, in 2008.

From 2007 to 2008, he was a visiting Ph.D. student with the Institute for Mathematical Sciences, Imperial College London, London, U.K. From 2016 to 2017, he was an Academic Visitor with the Department of Computer Science, University of California at Davis, Davis, CA, USA. He joined the National University of Defense Technology, in 2008, where he is currently a Professor. His current research interests include complex networks, especially the structural robustness of complex networks.

Yu Xiao received the B.S. degree in communication engineering from Sichuan University, Chengdu, China, in 2014, and the M.S. degree in management science and engineering from the National University of Defense Technology, Changsha, China, in 2016, where he is currently pursuing the Ph.D. degree in management science and engineering.

His current research interests include multicriteria decision analysis, rank aggregation, graph theory, and voting theory.

Mengxiang Zhang (S'14) received the B.S. and M.S. degrees in system science from Sichuan University, Chengdu, China, in 2013 and 2016, respectively, where he is currently pursuing the Ph.D. degree in management science and engineering.

He has published several peer-reviewed papers in international journals. His current research interests include algorithms, uncertain decision making, and system engineering.

Mr. Zhang is a Student Member of the American Society of Civil Engineers.

Yang Yu received the B.S. and M.S. degrees in management science and engineering from the National University of Defense Technology, Changsha, China, in 2015 and 2017, respectively.

She is a Team Member of the Complex System Laboratory, College of System Engineering, National University of Defense Technology. Her current research interests include robustness of complex networks, characteristic of directed networks, and optimal disintegration strategies of networks.

Yang Zhang received the B.S. and M.S. degrees in management science and engineering from the National University of Defense Technology, Changsha, China, in 2015 and 2017, respectively.

He has published several peer-reviewed papers in international journals. His current research interests include university ranking, ranking aggregation, and complex network.