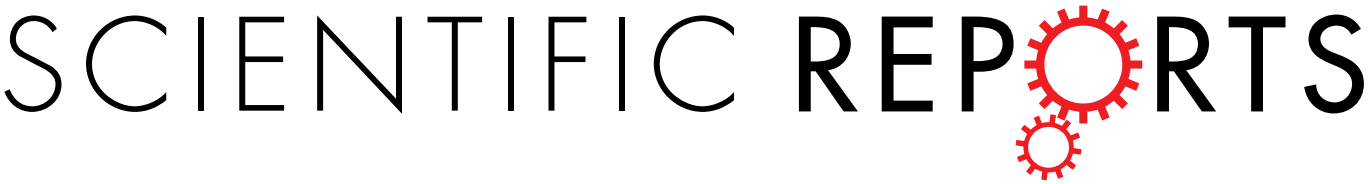

OPEN

Mitigate Cascading Failures on Networks using a Memetic Algorithm

Received: 18 March 2016
Accepted: 15 November 2016
Published: 09 December 2016

Xianglong Tang, Jing Liu & Xingxing Hao

Research concerning cascading failures in complex networks has become a hot topic. However, most of the existing studies have focused on modelling the cascading phenomenon on networks and analysing network robustness from a theoretical point of view, which considers only the damage incurred by the failure of one or several nodes. However, such a theoretical approach may not be useful in practical situation. Thus, we first design a much more practical measure to evaluate the robustness of networks against cascading failures, termed R_{cf}. Then, adopting R_{cf} as the objective function, we propose a new memetic algorithm (MA) named MA-R_{cf} to enhance network the robustness against cascading failures. Moreover, we design a new local search operator that considers the characteristics of cascading failures and operates by connecting nodes with a high probability of having similar loads. In experiments, both synthetic scale-free networks and real-world networks are used to test the efficiency and effectiveness of the MA-R_{cf}. We systematically investigate the effects of parameters on the performance of the MA-R_{cf} and validate the performance of the newly designed local search operator. The results show that the local search operator is effective, that MA-R_{cf} can enhance network robustness against cascading failures efficiently, and that it outperforms existing algorithms.

Many man-made infrastructures such as the Internet, transportation networks, and electric power grids can be represented as complex networks[1]. Because these complex networks play an important role in society, their robustness is pivotal[2–4]. However, most of these infrastructures have been found to be heterogeneous and to have a power-law degree distribution[1,5,6]. With their "heavy-tailed" properties, these complex networks have been found to be robust against random attacks; however, they are rather fragile under malicious attacks, especially cascade-based attacks[7,8].

Cascading failures are common in modern social networks. For example, in electrical power grids, when a power transmission station or a power line goes down, its power will be shifted to the nearby stations (lines). In most cases, neighbouring stations can manage the extra load. However, in some extreme circumstances, these neighbouring stations may become overloaded and fail, resulting in a redistribution of their loads to their neighbours. Ultimately, the redistribution effort may lead to a cascading failure in which a large number of power transmission stations (lines) are overloaded and, consequently, malfunction[9]. Cascading failures may also take place on the Internet. The load on an Internet router represents data packets that must be transmitted per unit of time, and overloading corresponds to congestion[10]. Rerouting data packets from a congested router to another might spread the congestion to a large fraction of subnetworks. Some Internet collapses have been caused by congestion[11]. Another example is a power grid, in which each component is designed to deal with a specific load of power. On August 14, 2003 in Canada and the northeastern United States, a massive power blackout occurred that led to a cascading failure[12]. A similar breakdown occurred in southern Oregon on August 10, 1996[13,14].

Cascading failures in complex networks have been widely studied over the past few decades[15–24]. Different cascading failure models have been proposed to reproduce cascading phenomena. Motter *et al.* first proposed the "*C-L*" model in ref. 18, performing experiments on both random and scale-free networks that focused on cascading triggered by the failure of a single node. The "*C-L*" model obtained good results that were consistent with experts' intuition about how cascading failures occur. Crucitti *et al.*[19] introduced a dynamical model that considered the dynamical redistribution of flow in networks, in which overloaded nodes obstruct network traffic rather than removed. Zhao *et al.*[20] provided a mathematical proof of the "*C-L*" model in scale-free networks

Key Laboratory of Intelligent Perception and Image Understanding of Ministry of Education, Xidian University, Xi'an 710071, China. Correspondence and requests for materials should be addressed to J.L. (email: neouma@mail.xidian.edu.cn)

that analysed the cascading breakdown in scale-free networks in terms of phase transitions. Feng *et al.*[23] proposed an approach of simple, self-consistent probability equations to study cascading behaviours in interdependent networks and showed that this approach can greatly simplify the mathematical analysis of systems ranging from single-layer networks to various types of interdependent networks. Hu *et al.*[24] used a percolation approach to study more realistic coupled networks system in which both interdependent and interconnected links exist and found rich and unusual phase-transition phenomena—including mixed first- and second-order hybrid transitions.

Based on different cascading failure models, various strategies have been proposed to enhance network robustness against cascading failures. Koç *et al.*[25] proposed a robust metric for cascading failures on power grid networks; an entropy-based metric was introduced in ref. 26. Wang *et al.*[27] studied cascading failures on the Internet. Based on a new cascading edge model, they proposed some methods to protect the Internet from cascading failures. However, all the above methods focused only on cascades triggered by removing one or two nodes, and such methods cannot evaluate the overall robustness of networks against cascading failures and may not be useful in many practical applications. Additionally, these methods rarely take the cost involved in updating the real-world systems into account.

Considering only the cascading failure resulting from removing individual nodes in networks is insufficient because many of the remaining nodes are still connected; therefore, the network still maintains its integrity to a certain extent. In contrast, in this paper, we first design a new robustness measure to evaluate the overall robustness of networks against cascading failures. In this robustness measuring scheme, the network is attacked through cascading failures repeatedly until the entire network collapses. During this process, after each cascade attack, the remaining large network components are calculated.

Based on this measure, we propose a memetic algorithm (MA) called MA-R_{cf} that enhances network robustness against cascading failures. Memetic algorithms form a popular branch of evolutionary algorithms (EAs) that successfully combine global and local searches and have been shown to be more efficient and more effective than traditional EAs for many problems[28–30]. In a previous study, we designed a new memetic algorithm named MA-RSF_{MA}, which improves the robustness of scale-free networks against malicious attacks that achieved a good performance[31]. Thus, the algorithm proposed here, MA-R_{cf}, is based on the framework developed for MA-RSF_{MA} and makes use of the properties of cascading failures to design new operators; that is, a new local search operator that considers the characteristic of cascading failures is designed for MA-R_{cf}. In MA-R_{cf}, the degree distribution of networks is also kept unchanged to minimize the costs of updating real-world systems. Both synthetic and real-world networks are used to validate the performance of the MA-R_{cf}. The experimental results show that the MA-R_{cf} can enhance the network robustness against cascading failures efficiently. Moreover, some properties of robust networks are also analysed.

Methods

Robustness Measure for Cascading Failures. A network can be modelled as a graph, $G=(V, E)$, where $V=\{1, 2, \ldots, N\}$ is a set of N nodes and $E=\{e_{jk}|\ j, k\in V$ and $j\neq k\}$ is a set of M links. In ref. 18, Motter *et al.* proposed the "*C-L*" model for cascading failures in which, for a given network, at each time step, one unit of the relevant quantity (such as energy or goods) is exchanged between each pair of nodes and transmitted along the shortest connecting path. The "load" at a node consists of the total number of shortest paths passing through it[32,33]. Each node carries the maximum load that it can handle, and in man-made networks, node capacity is limited by economic costs. The capacity C_i of node i and its initial load L_i have the following proportional relation:

$$C_i = (1 + \alpha)L_i^0,\ i = 1, 2, \ \ldots, N, \tag{1}$$

where the constant α is a tolerance parameter, and L_i^0 is the initial load of the ith node. Initially, the network operates in a free-flow state insofar as $\alpha \geq 0$. However, the failure of a node for any reason triggers the dynamics of the redistribution of loads. When the load at a node becomes larger than the node's capacity, the node fails. This forces the load previously carried by that node to shift to its neighboring nodes, which in turn, can cause them to fail. Consequently, subsequent failures can occur, and this step-by-step process is a cascading failure[18].

In ref. 34, Schneider *et al.* proposed an effective robustness measure, R, to evaluate networks' ability to resist targeted attacks on individual nodes. The R measure is based on the "giant component," namely, the largest connected component left in the network after each node removal. To calculate R, the network must be attacked until only separated nodes are left. Thus, R can evaluate the robustness of entire networks. Therefore, we combine the "*C-L*" model with R to design a new measure, R_{cf}, which can evaluate the overall robustness of networks against cascading failures. With the original property of "*C-L*" model in mind, the process for calculating R_{cf} is described below.

Step 1. $\boldsymbol{S_{sum}} \leftarrow 0$ and $\boldsymbol{t} \leftarrow 1$, where $\boldsymbol{S_{sum}}$ is the accumulated size of the giant components and $\boldsymbol{t}$ is the index of cascaded attack rounds;
Step 2. Calculate the initial load L_i^0 and $\boldsymbol{C_i}$ of each node;
Step 3. Remove the node with the maximum load and all the edges connected to it;
Step 4. If the number of remaining nodes is equal to 1, go to Step 6;
Step 5. Recalculate the L_i^t of each remaining node:
If a remaining node $\boldsymbol{i}$ is overloaded, namely, $\boldsymbol{L_i > C_i}$, then remove node $\boldsymbol{i}$ and the edges connected to it, then, go to Step 4;
If no remaining node is overloaded, calculate the relative size of the giant component $\boldsymbol{S^t}$, $\boldsymbol{S_{sum} \leftarrow S_{sum} + S^t}$, $\boldsymbol{t \leftarrow t + 1}$, then, go to Step 3;
Step 6. Calculate the robustness measure $\boldsymbol{R_{cf}}$ against cascading failures as follows,

$$R_{cf} = \frac{S_{sum}}{N} = \frac{1}{N}\sum_{t=1}^{T} S^t \quad (2)$$

where T is the total number of rounds that a cascading failure-based attack needs to destroy the entire network, reducing it to only one node. Obviously, T may vary even for networks of the same size; therefore, the normalization factor $1/N$ ensures comparability for the robustness of networks with different sizes.

Memetic Algorithm to Enhance R_{cf}. Memetic algorithms have been shown to be highly capable of searching for the optimal solution in optimization problems[28–30]. In ref. 31, we designed a new memetic algorithm named MA-RSF$_{MA}$ to improve the robustness of scale-free networks against malicious attacks, and it achieved a good performance[31]. Thus, based on the framework of MA-RSF$_{MA}$, in this paper, we propose a new memetic algorithm, MA-R$_{cf}$, to enhance the overall robustness of networks against cascading failures. By considering the intrinsic property of cascading failures, we design a new local search operator for MA-R$_{cf}$ that takes R_{cf} as its objective function while keeping the degree of each node unchanged. Next, we introduce the representation of chromosomes and the initialization process. Then, we describe the evolutionary operators, including the newly designed local search operator. Finally, we summarize the entire framework of the MA-R$_{cf}$ algorithm.

Representation and initialization. In the MA-R$_{cf}$ algorithm, each chromosome represents a network. Initially, MA-R$_{cf}$ has a population with W chromosomes. The initialization process for MA-R$_{cf}$ is the same as that used for MA-RSF$_{MA}$[31]. During the initialization, because we need to preserve the number of links and the degree of each node, each chromosome is generated by randomly adjusting a fraction of the edges in the initial network, G_0—that is, the connections of two randomly chosen edges that have no common nodes are swapped in the network. During the initialization, the goal is to generate different networks with the same degree distribution; therefore, any edge-swapping operations that can keep the network connected are accepted without checking whether the swap improves the robustness of the network. The details of this initialization process are summarized in Algorithm 1 (also refer to ref. 31 for more information).

Algorithm 1 Population Initialization

Input:

W: Population size;

G_0: Initial network;

Output:

$\boldsymbol{P}^1 = \{G_1^1, G_2^1, \ldots, G_W^1\}$: Population for the 1st generation;

for $i = 1$ ***to*** W ***do***

  $G_i^1 \leftarrow G_0$;

  for $j = 1$ ***to*** M ***do*** /*M is the number of nodes in the network; */

    Randomly select two edges e_{kl} and e_{mn} from G_i^1, where m, n are different than k, l and e_{km} and e_{ln} do not exist in G_i^1;

    Remove e_{kl} and e_{mn} from G_i^1, and add e_{km} and e_{ln} to G_i^1;

  if (G_i^1 is not connected)

    Remove e_{km} and e_{ln} from G_i^1, and add e_{kl} and e_{mn} back to G_i^1;

  end if;

end for;

end for;

Evolutionary operators. In evolutionary algorithms, crossover operators are often performed to exchange genetic information among the individuals in the population. In ref. 31, a new crossover operator that operates on complex networks is designed that exchanges the structures of two networks effectively without changing their degree distributions. We also employ this crossover operator in this paper. This crossover operator (whose details are summarized in Algorithm 2) acts on two randomly selected parent chromosomes and generates a pair of child chromosomes. Please refer to ref. 31 for more information about this crossover operator.

Algorithm 2 Crossover Operator

Input:

G_{p1} and G_{p2}: Two parent chromosomes;

p_c: Crossover rate;

Output:

G_{c1} and G_{c2}: Two child chromosomes;

$G_{c1} \leftarrow G_{p1}$, $G_{c1} \leftarrow G_{p1}$;

Continued

for $i = 1$ ***to*** N ***do***/*N is the number of edges in the network; */
  if $(U(0, 1) < p_c)$/*$U(0, 1)$ is a uniformly distributed random real number in [0, 1];*/
    Determine $V_i^{G_{c1}}$ and $V_i^{G_{c2}}$, which are the sets of neighbours of node i in G_{c1} and G_{c2}, ;respectively,
    $\overline{V}_i^{G_{c1}} \leftarrow V_i^{G_{c1}} - (V_i^{G_{c1}} \mathrm{I}\, V_i^{G_{c2}})$ and $\overline{V}_i^{G_{c2}} \leftarrow V_i^{G_{c2}} - (V_i^{G_{c1}} \mathrm{I}\, V_i^{G_{c2}})$;
    for each node $j \in \overline{V}_i^{G_{c1}}$ ***do***
      Randomly select a node $k \in \overline{V}_i^{G_{c2}}$;
      Remove e_{ij} from G_{c1} and e_{ik} from G_{c2};
      Add e_{ik} to G_{c1} and e_{ij} to G_{c2};
      Randomly select another edge e_{kl} that connects to node k in G_{c1} but where e_{jl} does not exist in G_{c1};
      Remove e_{kl} and add e_{jl} in G_{c1};
      Randomly select another edge e_{jm} that connects to node j in G_{c2} but where e_{km} does not exist in G_{c2};
      Remove e_{jm} and add e_{km} in G_{c2};
      $\overline{V}_i^{G_{c2}} = \overline{V}_i^{G_{c2}} - \{k\}$;
    end for;
  end if;
end for;

The local search operator is another important operator in MAs. In ref. 35, Tanizawa *et al.* found that networks with an "onion-like" structure, where nodes with almost the same degree are connected to each other, are more robust under targeted attacks than those without such onion-like structures. Inspired by this, to search for networks that are the most robust against cascading failures, we design a new local search operator that lets nodes with similar loads connect to each other. The principle of this operator is simple: if a node with small load connects to a node with a very large load, the small-load node will crash immediately if the large-load node fails, because the small-load node has insufficient capacity to handle the extra load. Therefore, connecting nodes with similar loads to each other have a high probability of avoiding such situations. Suppose edges e_{ij} and e_{pq} are selected and are swapped to e_{ip} and e_{jq}. This swap is accepted only if

$$\left|L_i - L_p\right| + \left|L_j - L_q\right| < \beta \times \left(\left|L_i - L_j\right| + \left|L_p - L_q\right|\right), \tag{3}$$

where L_i, L_j, L_p and L_q are the loads of the corresponding nodes, and β is a parameter in the range of [0, 1] that controls the acceptance level for the difference in loads between nodes. The smaller the value of β is, the stronger the constraint is and, thus, the larger the reduction in load differences is. Consequently, this operator can effectively guarantee that nodes with similar loads will be connected to each other, which enhances the search for load-balanced networks. The details of this local search operator are given in Algorithm 3.

Algorithm 3 Local Search Operator

Input:
  G: One chromosome;
  p_l: Local search probability;
  β: Predefined parameter;
Output:
  G: Chromosome after performing the local search operator;

for (each edge e_{ij} in G) ***do***
  if $(U(0, 1) < p_l)$/*$U(0, 1)$ is a uniformly distributed random real number in [0, 1];*/
    Randomly select another existing edge e_{pq} in G;
    if (equation (3) is satisfied)
      $G^* \leftarrow G$;
      Remove e_{ij} and e_{pq} from G^*;
      Add e_{ip} and e_{jq} to G^*;
      if $(R_{cf}(G^*) > R_{cf}(G))$
        $G \leftarrow G^*$;
      end if;
    end if;
  end if;
end for;

MA-R_{cf} uses binary tournament selection in each generation to select the chromosomes for the next population. Binary tournament selection involves a "tournament" between two chromosomes chosen randomly from the population in which the winner is the chromosome whose fitness is better. Binary tournament selection is a popular method for selecting better chromosomes from a population in an evolutionary algorithm.

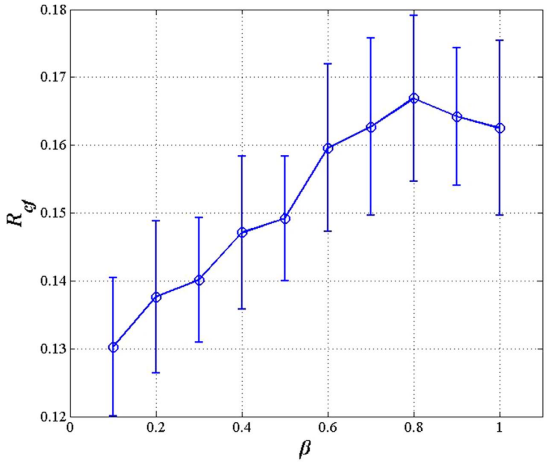

Figure 1. The effect of parameter β on the performance of MA-R_{cf}.

N	Initial	Without (Average ± Standard Deviation)	With (Average ± Standard Deviation)
100	0.0751	0.1617 ± 0.0135	0.1669 ± 0.0128
200	0.0521	0.1202 ± 0.0127	0.1309 ± 0.0119
300	0.0395	0.0924 ± 0.0097	0.1072 ± 0.0085
500	0.0263	0.0783 ± 0.0074	0.0874 ± 0.0067

Table 1. The R_{cf} values of BA networks obtained by MA-R_{cf} with and without the local search operator. The results are averaged over 10 independent runs.

Algorithm 4 MA-R_{cf}

Input:

W: Population size;

G_0: Initial network;

p_c: Crossover probability;

p_l: Local search probability;

Output:

G^*: Chromosome with the highest robustness (R_{cf});

$\mathbf{P}^1 \leftarrow$ **Population_Initialization** (W, G_0) and $t \leftarrow 1$;

while (terminal criteria are not satisfied) *do*

$\mathbf{P}_c^t \leftarrow \varnothing$ //$\mathbf{P}_c^t$ is the child population of $\mathbf{P}^t$

Repeat

Randomly choose two chromosomes G_i^t and G_j^t from $\mathbf{P}^t$ that have not been selected;

$\left(G_{ci}^t, G_{cj}^t\right) \leftarrow$ **Crossover_Operator** $\left(G_i^t, G_j^t, p_c\right)$;

$\mathbf{P}_c^t \leftarrow \mathbf{P}_c^t \cup \left\{G_{ci}^t, G_{cj}^t\right\}$;

Until (all chromosomes in $\mathbf{P}^t$ have been selected);

Calculate the robustness of each chromosome in $\mathbf{P}_c^t$;

for $i = 1$ *to* W ***do***

Select a chromosome G^t from $\mathbf{P}^t$ and $\mathbf{P}_c^t$ using roulette wheel selection based on the robustness of all chromosomes;

Conduct the local search operator on G^t with probability p_l;

end for;

$\mathbf{P}^{t+1} \leftarrow$ **2-Tournament_Selection**($\mathbf{P}^t$, $\mathbf{P}_c^t$);

$t \leftarrow t+1$;

end while.

Implementation of MA-R_{cf}. In MA-R_{cf}, the initialization operator is first used to generate an initial population with W chromosomes. In each generation, the crossover operator is applied to the population first, and then, the local search operation is conducted. After performing the crossover operator, a new child population is obtained.

N	Algorithms	Best	Worst	Average ± Standard deviation
100	Hill Climbing	0.1231	0.0912	0.1035 ± 0.0113
	Simulated Annealing	0.1356	0.1037	0.1187 ± 0.0109
	MA-R_{cf}	0.1797	0.1476	**0.1669 ± 0.0128**
200	Hill Climbing	0.0926	0.0747	0.0813 ± 0.0066
	Simulated Annealing	0.1021	0.0893	0.0922 ± 0.0053
	MA-R_{cf}	0.1427	0.0891	**0.1309 ± 0.0119**
300	Hill Climbing	0.0821	0.0668	0.0729 ± 0.0059
	Simulated Annealing	0.0932	0.0720	0.0864 ± 0.0066
	MA-R_{cf}	0.1157	0.0892	**0.1072 ± 0.0085**
500	Hill Climbing	0.0782	0.0595	0.0651 ± 0.0054
	Simulated Annealing	0.0883	0.0692	0.0793 ± 0.0059
	MA-R_{cf}	0.1021	0.0722	**0.0874 ± 0.0067**

Table 2. The R_{cf} of BA networks of different sizes obtained by the three tested algorithms. The results are averaged over 10 independent runs.

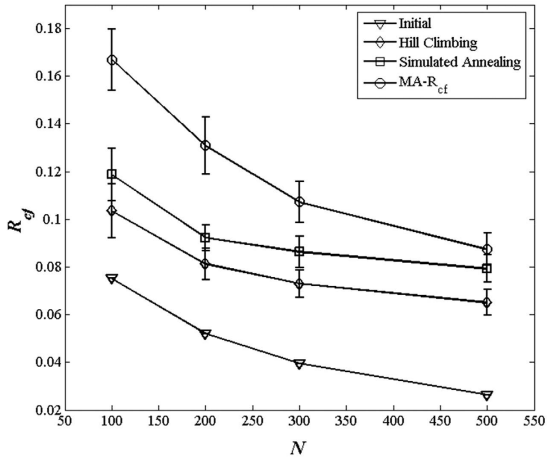

Figure 2. A comparison between MA-R_{cf} and existing algorithms on BA networks of different sizes.

Then, the local search operator and the binary tournament selection operator are applied to both the parent and child populations to generate the child population for the next generation. Finally, the best chromosome in the last population is the most robust network found. The framework of MA-R_{cf} is summarized in Algorithm 4.

Results

In this section, because scale-free networks have become an important type of network, experiments are conducted on both synthetic scale-free networks and real-world networks to validate the performance of MA-R_{cf}. We also study some of the network properties of the robust networks obtained by MA-R_{cf}. The synthetic scale-free networks were generated using the BA model[5], and their average degree was set to 4. In ref. 31, Zhou *et al.* showed that MA-RSF_{MA} can improve the robustness of scale-free networks against malicious attacks effectively; consequently, we also compare networks optimized by MA-R_{cf} with those optimized by MA-RSF_{MA} to investigate the different properties of both optimized networks.

The parameter α in (1) reflects the capacity of a node to handle its load. A larger α indicates a stronger node. The value of α is always assumed to be in the range [0, 1]: a $\alpha > 1$ is unrealistically large[18,19]. In this work, we assume that ability of a node to handle its loads is average (neither very strong nor very weak). Thus, in the following experiments, α is set to a median value, 0.5.

In the local search operator, the tolerance parameter β controls the percentage by which loads can differ between connected nodes. Therefore, we first conducted an experiment to find an appropriate value for β. This experiment used BA networks with 100 nodes. The robustness obtained by MA-R_{cf} under different values of β is plotted in Fig. 1. The results are averaged over ten independent runs on each sampled point. As Fig. 1 shows, MA-R_{cf} achieves the highest robustness when β equals 0.8. Thus, β is set to 0.8 in the following experiments.

The other parameters of MA-R_{cf} were set as follows: the population size W was set to 10, the crossover probability p_c and the local search probability p_l were set to 0.8 and 0.5, respectively, and the maximum number of objective function evaluations was set to 5×10^4.

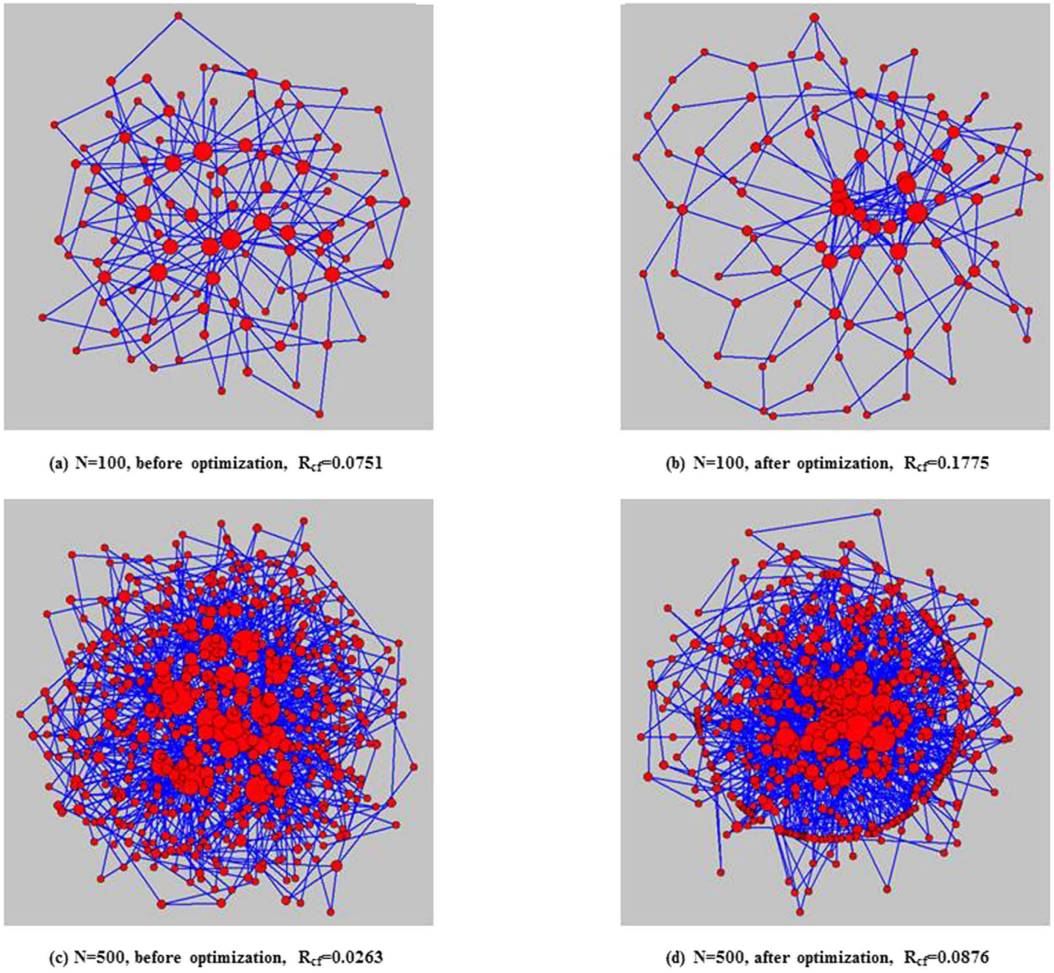

Figure 3. The network topology before and after optimization by MA-R_{cf}. The size of each node is proportional to its degree.

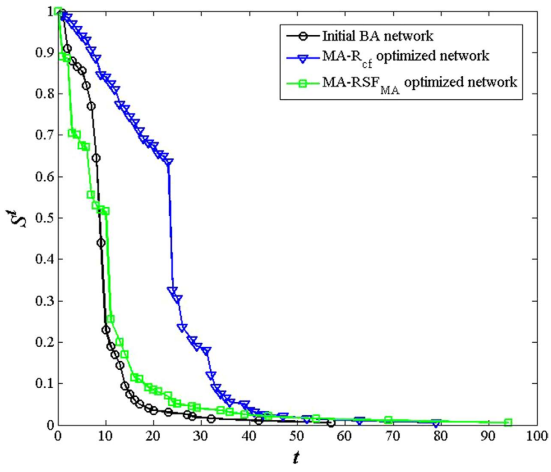

Figure 4. The change in the relative size of the giant component S^t over a series of cascaded attack circles t. The BA network has 200 nodes. The R_{cf} values of the initial BA network, the optimized network obtained by MA-RSF_{MA} and the optimized network obtained by MA-R_{cf} were 0.0521, 0.0558 and 0.1319, respectively.

Experiments on Synthetic Networks. In this experiment, scale-free networks with different scales were used to test the performance of MA-R_{cf}. First, an experiment was carried out to test the effectiveness of the local search operator. In this experiment, versions of MA-R_{cf} both with and without the local search operator were

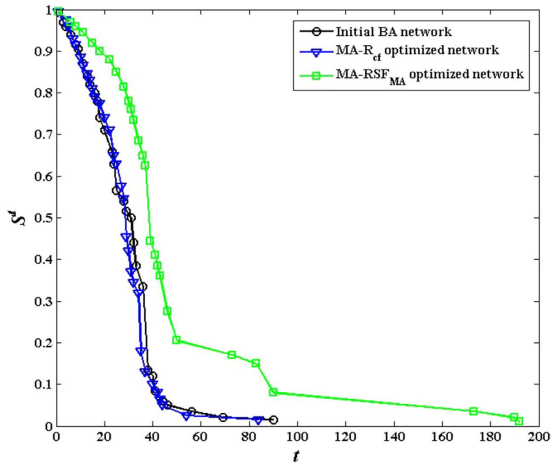

Figure 5. The change in the relative size of the giant component S^t under high node degree attack circles t. The BA network has 200 nodes. The R_{cf} values of the initial BA network, the optimized network obtained by MA-RSF_{MA} and the optimized network obtained by MA-R_{cf} were 0.0521, 0.0558 and 0.1319, respectively.

N		A (Average ± Standard Deviation)	L (Average ± Standard Deviation)	C (Average ± Standard Deviation)
100	Before Optimization	−0.1489 ± 0.0000	3.1681 ± 0.0000	0.3563 ± 0.0000
	MA-RSF_{MA}	0.6206 ± 0.0242	8.1168 ± 0.2203	0.2186 ± 0.0093
	MA-R_{cf}	0.3956 ± 0.0132	3.8307 ± 0.1841	0.3139 ± 0.0125
200	Before Optimization	−0.2194 ± 0.0000	3.5645 ± 0.0000	0.3124 ± 0.0000
	MA-RSF_{MA}	0.3951 ± 0.0176	7.9266 ± 0.2312	0.2057 ± 0.096
	MA-R_{cf}	0.1734 ± 0.0118	3.5726 ± 0.1873	0.3115 ± 0.0103
300	Before Optimization	−0.2344 ± 0.0000	3.6327 ± 0.0000	0.3001 ± 0.0000
	MA-RSF_{MA}	0.3480 ± 0.0198	8.2787 ± 0.2241	0.1801 ± 0.0102
	MA-R_{cf}	0.1470 ± 0.0093	3.7308 ± 0.1254	0.2943 ± 0.0089
500	Before Optimization	−0.2507 ± 0.0000	3.8264 ± 0.0000	0.2836 ± 0.0000
	MA-RSF_{MA}	0.3651 ± 0.0142	8.1968 ± 0.2101	0.1869 ± 0.0114
	MA-R_{cf}	0.1033 ± 0.0076	3.8304 ± 0.1219	0.2814 ± 0.0081

Table 3. The changes in important network properties of BA networks with different sizes before and after optimization by MA-R_{cf} and MA-RSF_{MA}, including the assortative coefficients (A), the average shortest path length (L) and the global communication efficiency (C). The values of the optimized networks were averaged over 10 independent runs.

tested on BA networks with 100, 200, 300, and 500 nodes. The obtained robustness values are listed in Table 1, which shows that the MA-R_{cf} version with the local search operator always performs better than the version without the local search operator. Therefore, the local search operator in MA-R_{cf} is effective.

Next, some experiments were conducted to test the ability of MA-R_{cf} to search for the most robust networks. Network structure optimization is a hard optimization problem. In existing works, the hill climbing algorithm[34] and the simulated annealing algorithm[36,37] are widely used to address this problem. Thus, we compared the performance of MA-R_{cf} with that of the hill climbing algorithm[34] and the simulated annealing algorithm[36,37]. The maximum number of objective function evaluations for these two algorithms was also set to 5×10^4 to obtain the results of these three algorithms at the same computational cost.

We tested BA networks with 100, 200, 300, and 500 nodes. The best, worst and average values of R_{cf} of the three algorithms over 10 independent runs are reported in Table 2. In addition, the corresponding curves of the average robustness obtained by the different algorithms are plotted in Fig. 2. As shown, MA-R_{cf} obtains the highest robustness values among these algorithms on all test networks. That is, MA-R_{cf} always finds more network structures more robust to cascading failures than do the other algorithms.

It is useful to study the robustness of the network structures obtained MA-R_{cf}. Thus, the network topologies of BA networks before and after optimized by MA-R_{cf} are plotted in Fig. 3, where the size of each node is proportional to its degree. As shown, before optimization, low degree nodes are often connected to nodes with high degrees; consequently, the entire network is composed of numerous star networks with hub nodes. However, the optimized networks which have higher R_{cf}, the low degree nodes are more likely to be connected to other low degree nodes and the high degree nodes are more likely to be connected to other high degree nodes. The entire network is a hub-node-connected structure. Considering the property of cascading failures, it is easy to

Network	Algorithms	Best	Worst	Average ± Standard deviation
WE Power	Hill Climbing	0.1221	0.1072	0.1135 ± 0.0063
	Simulated Annealing	0.1256	0.1097	0.1187 ± 0.0059
	MA-R_{cf}	0.1494	0.1126	**0.1330 ± 0.0082**
US Air	Hill Climbing	0.0463	0.0378	0.0415 ± 0.0047
	Simulated Annealing	0.0481	0.0382	0.0422 ± 0.0051
	MA-R_{cf}	0.0537	0.0401	**0.0475 ± 0.0062**

Table 4. The robustness of networks after optimization with different algorithms on two real world networks. The results shown were averaged over 10 independent runs. The initial R_{cf} values of the WE Power network and US air network were 0.1022 and 0.0251, respectively.

understand why hub-node-connected networks have a stronger ability to resist cascade failures: when a hub node fails, the neighbouring hub nodes can withstand the additional loads effectively, preventing the spread of cascading failures.

Next, we carried out an experiment to test how well the networks obtained by MA-R_{cf} resist cascading failures. We simulated the process of cascaded failures on BA networks with 200 nodes until the size of the giant component decreased to 1. The decreasing progress of the size of the giant component S^t is plotted in Fig. 4. As shown, along with the increasing cascade attack circle t, the MA-R_{cf} optimized network, which has a higher R_{cf} value, protects the giant component more effectively. The area between the curve of the "Initial BA network" and the curve of the "MA-R_{cf} optimized network" in Fig. 4 represents the amount of cascade failure mitigation, which corresponds to improving network robustness against cascade failures by 153%. These results show that the networks obtained by MA-R_{cf} can resist cascading failures effectively.

In ref. 31, Zhou *et al.* found that the onion-like network in which nodes with similar degree connect to each other can contribute to resisting high node degree attacks, we plotted the decreasing process of the size of S^t of networks optimized by MA-RSF_{MA} under cascading failures in Fig. 4, which has 200 nodes. In Fig. 5, we separately plotted the decreasing process of S^t of networks optimized by both MA-R_{cf} and MA-RSF_{MA} under high node degree attacks. In each attack circle, the node with largest degree and all the edges connected to it are removed. To perform a fair comparison, the parameters for MA-RSF_{MA} were set to the same as those for MA-R_{cf}, namely, the population size was set to 10, the crossover probability and the local search probability were set to 0.8 and 0.5, respectively, and the maximum number of objective function evaluations was set to 5×10^4. In Fig. 4, under cascaded attack circles, the size of the S^t of the MA-RSF_{MA} optimized networks decreases as fast as that of the initial BA network—even more sharply after the first several attack rounds. Moreover, under high node degree attack circles, the size of the S^t of the MA-R_{cf} optimized network decreases as fast as initial BA networks (see Fig. 5 for more details). In other words, the MA-R_{cf} algorithm cannot contribute to resisting high node degree attacks. These two experiments show that although the network structures optimized by these two algorithms have some similarity, their ability to resist cascading failures is significantly different.

We are also interested in whether other important network properties might have changed because of the optimization. Therefore, in Table 3 shows the results of assessing the assortativity coefficient[38], the average shortest path length and the global communication efficiency[39] of networks both before and after being optimized by MA-R_{cf}. As shown, before the optimization, the BA networks have negative assortativity coefficient values that become positive after the optimization. In other words, the correlation degree of the networks changes from disassortativeness to assortativeness after the optimization. This occurs because the optimization process promotes the connection of high degree nodes with other high degree nodes. After the optimization, the average shortest path length of BA networks is slightly increased while the global communication efficiency has a slight decrease, which means that the optimization process has no significant effect on network communication efficiency.

Because the networks obtained by MA-RSF_{MA} are also assortative, it is interesting to study the difference in terms of the network properties of networks obtained by both MA-R_{cf} and MA-RSF_{MA}; these properties are listed in Table 3. We can see that when both algorithms optimize the same network, the network obtained by MA-R_{cf} is far less assortative than that obtained by MA-RSF_{MA}. In addition, the average shortest path length of the MA-R_{cf} is only half that of MA-RSF_{MA}. Moreover, the networks obtained by MA-R_{cf} have higher global communication efficiency.

Experiments on Real World Networks. In this section, MA-R_{cf} is applied to two real-world networks. One is an electrical power grid in Western Europe (mainly Portugal and Spain)[40], labelled the WE Power grid network. It has 217 nodes and 320 edges. The average degree of the WE Power grid network is 2.95. The other network is the US air network—the US air transportation system[41]—consisting of 332 airports and 2126 air routes, in which the nodes represent airports and the edges present flights. The average degree of the US air network is 12.81. These two real networks are well connected and without any separate component.

The hill climbing algorithm, simulated annealing algorithm and MA-R_{cf} are used to independently optimize the above two networks. The obtained robustness values are reported in Table 4. As we can see, MA-R_{cf} always performs better than the two other algorithms. The network topologies of these two real networks before and after optimized by MA-R_{cf} are shown in Fig. 6. Comparing the network structure before and after optimization, we can see that, in the optimized networks, nodes with similar degrees are more likely to connect with each other, and the hub nodes are more closely connected to each other than before. Consequently, even for an existing network, MA-R_{cf} can find a structure more robust against cascading failure. By comparing the robust structure with the

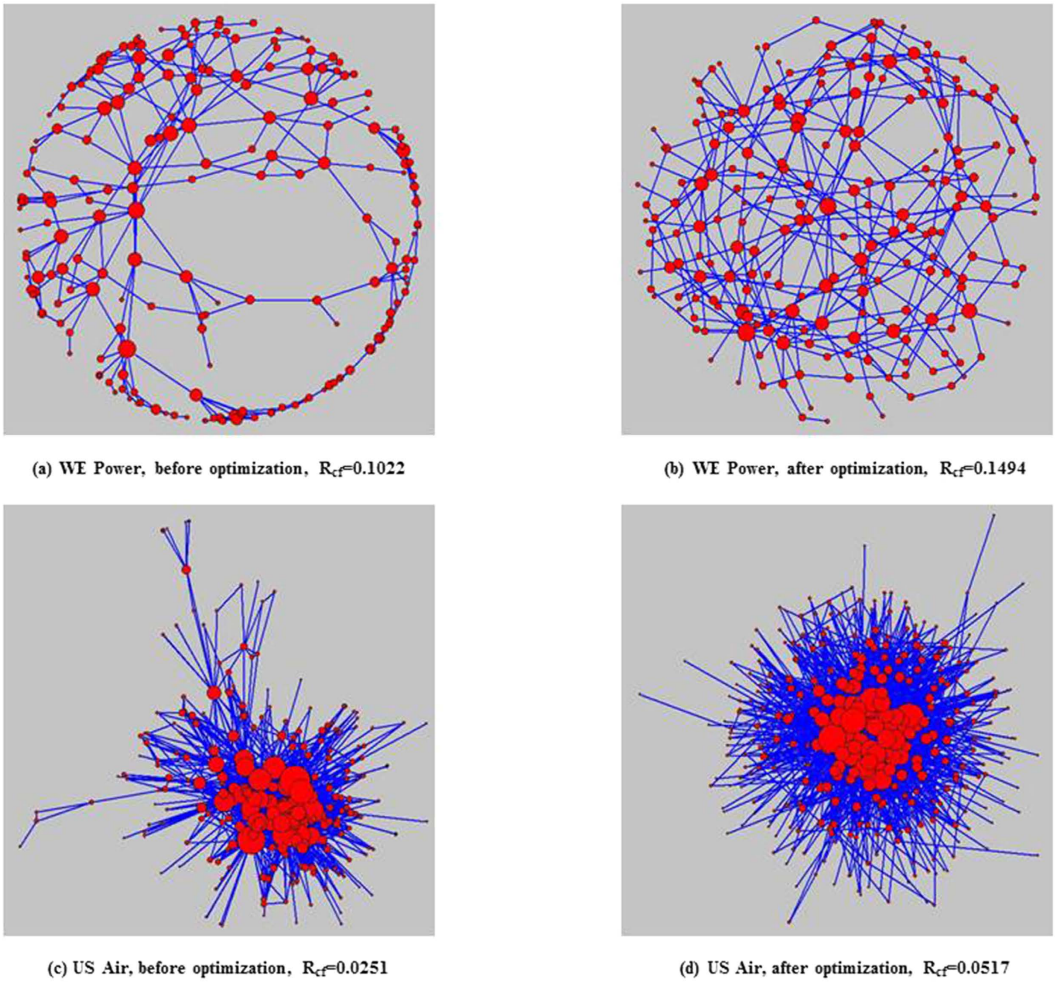

Figure 6. The network topology of two real world networks before and after optimization by MA-R_{cf}. The size of each node is proportional to its degree.

N		A (Average ± Standard Deviation)	L (Average ± Standard Deviation)	C (Average ± Standard Deviation)
WE Power	Before Optimization	0.1269 ± 0.0000	6.9381 ± 0.0000	0.1870 ± 0.0000
	After Optimization	0.2176 ± 0.0097	4.9511 ± 0.1903	0.2313 ± 0.0115
US Air	Before Optimization	−0.2078 ± 0.0000	2.7381 ± 0.0000	0.4059 ± 0.0000
	After Optimization	−0.0819 ± 0.0056	2.4980 ± 0.1410	0.4336 ± 0.0089

Table 5. The changes in some important network properties of real world networks before and after optimization by MA-R_{cf}, including the assortative coefficients (A), the average shortest path length (L) and the global communication efficiency (C). The values of optimized networks were averaged over 10 independent runs.

initial structure, we can find several key edges, which—if changed—would increase network robustness significantly. Considering the cost of optimization, there is no need to change all the edges of the real network; instead, MA-R_{cf} can help find the key edges.

The assortativity coefficient, average shortest path length and global communication efficiency of these two real networks both before and after optimization are reported in Table 5. As listed, the WE Power network has a positive assortativity coefficient while the US air network has a negative assortativity coefficient. After optimization by MA-R_{cf}, the WE Power network has stronger assortativeness, while the disassortativeness of the US air network gets weaker. This result is the same as the results of the experiments with synthetic networks, further verifying that networks with more hub nodes connected to each other have a stronger ability to resist cascading failures. Interestingly, after optimization by MA-R_{cf}, the average shortest path length of these two networks decreases while their global communication efficiency increases. In other words, MA-R_{cf} can not only increase a network's global robustness against cascading failures but can also increase its global communication efficiency—even when applied to real networks.

Discussions

Securing network infrastructure is critical in today's society. When studying networks subject to cascading failures, considering only the damage from one or even several nodes is insufficient. In this paper, we describe the design of a more comprehensive index that can evaluate the ability of networks to resist cascading failures. To enhance networks resistance to cascading failures, we propose a memetic algorithm, MA-R_{cf}. Then, to test the performance of MA-R_{cf}, we tested it on both synthetic and real networks. The topologies of the robust networks obtained by MA-R_{cf} are shown and some of their network properties are discussed. From experiments comparing with MA-R_{cf} other network optimization algorithms, we can conclude that MA-R_{cf} achieves a better performance, showing that MA-R_{cf} is an effective algorithm for enhancing the robustness of networks against cascading failures.

When dealing with complex networks, the large computational complexity of calculating shortest paths limits the algorithms that rely on such calculations from being applied to large-scale networks. For example, the computation of R_{cf} in Equation (2) needs to calculate the shortest path of the network under each cascading attack circle; consequently, MA-R_{cf} is unable to optimize large networks at a low computational cost. However, studying the effects of MA-R_{cf} on cascading failures is still meaningful. MA-R_{cf} provides an opportunity to explore network structures that are robust against cascading failures. In this paper, we apply MA-R_{cf} to networks with specific sizes and study the topological structure and network properties of robust networks. The experiments show connecting hub nodes to each other more closely would be a good strategy when designing networks that are robust against cascading failures. Moreover, extending this discovery to large-scale networks is not difficult. In contrast with other traditional algorithms, MA-R_{cf} is a competitive algorithm for optimizing networks against cascading failures.

References

1. Barabási, A. L. & Albert, R. Emergence of scaling in random networks. *Science* **286,** 509 (1999).
2. Gao, J., Buldyrev, S. V., Havlin, S. & Stanley, H. E. Robustness of a network of networks. *Phys. Rev. Lett.* **107,** 195701 (2011).
3. Gao, J., Buldyrev, S. V., Havlin, S. & Stanley H. E. Robustness of a network formed by n interdependent networks with a one-to-one correspondence of dependent nodes. *Phys. Rev. E.* **85,** 066134 (2012).
4. Liu, X., Stanley, H. E. & Gao, J. Breakdown of interdependent directed networks. *Proceedings of the National Academy of Sciences*. **113,** 1138–1143 (2016).
5. Albert, R. & Barabási, A. L. Statistical mechanics of complex networks. *Rev. Mod. Phys*. **74,** 47 (2002).
6. Newman, M. E. J. The structure and function of networks. *SIAM Review* **45,** 167–256 (2003).
7. Jeong, H., Mason, S. P., Barabási, A. L. & Oltvai, Z. N. Lethality and centrality in protein networks. *Nature* **411,** 41 (2001).
8. Holme, P., Kim, B. J., Yoon, C. N. & Han, S. K. Attack vulnerability of complex networks. *Phys. Rev. E*. **65,** 056109 (2002).
9. Crucitti, P., Latora, V. & Marchiori, M. A model for cascading failures in complex networks. *Phys. Rev. E*. **69,** 045104 (2004).
10. Arenas, A., Dias-Guilera, A. & Guimera, R. Communication in networks with hierarchical branching. *Phys. Rev. Lett*. **86,** 3196–3199 (2001).
11. Jacobson, V. Congestion avoidance and control. *Comput. Commun. Rev*. **18,** 314 (1988).
12. Glanz, J. & Perez-Pena, R. *90 seconds that left tens of millions of people in the dark*. New York Times. (26th August 2003).
13. Sachtjen, M. L., Carreras, B. A. & Lynch, V. E. Disturbances in a power transmission system. *Phys. Rev. E*. **61,** 4877 (2000).
14. Carreras, B. A., Newman, D. E., Dolrou, I. & Poole, A. B. Initial evidence for self-organized criticality in electric power system blackouts. In: *Proceeding of Hawaii International Conference on System Sciences*. January 4-7, Maui, Hawaii (2000).
15. Di Muro, M. A. *et al.* Recovery of interdependent networks. *Sci. Rep*. **6,** 22834 (2016).
16. Wang, J., Xu, B. & Wu, Y. Ability paradox of cascading model based on betweenness. *Sci. Rep*. **5,** 13939 (2015).
17. Li, D., Jiang, Y. & Kang, R. Spatial correlation analysis of cascading failures: congestions and blackouts. *Sci. Rep*. **4,** 5381 (2014).
18. Motter, A. E. & Lai, Y. C. Cascade-based attacks on complex networks. *Phys. Rev. E*. **66,** 065102 (2002).
19. Crucitti, P., Latora, V. & Marchiori, M. Model for cascading failures in complex networks. *Phys. Rev. E*. **69,** 266–289 (2003).
20. Zhao, L., Park, K. & Lai, Y. C. Attack vulnerability of scale-free networks due to cascading breakdown. *Phys. Rev. E*. **70,** 035101 (2004).
21. Kinney, R. & Albert, R. Modeling cascading failures in the north American power grid. In: *2005 APS March Meeting American Physical Society. USA*. 101–107 (2005).
22. Wang, J. W. & Rong, L. L. Cascade-based attack vulnerability on the US power grid. *Safety Science* **47,** 1332–1336 (2009).
23. Feng, L., Monterola, C. P. & Hu, Y. The simplified self-consistent probabilities method for percolation and its application to interdependent networks. *New Journal of Physics* **17,** 063025 (2015).
24. Hu, Y., Ksherim, B., Cohen, R. & Havlin, S. Percolation in interdependent and interconnected networks: abrupt change from second- to first-order transitions. *Phys. Rev. E.* **84,** 066116 (2011).
25. Koç, Y., Warnier, M., Brazier, F. M. T. & Kooij, R. E. A robustness metric for cascading failures by targeted attacks in power networks. In: *Proceedings of the 10th IEEE International Conference on Networking, Sensing and Control* (*ICNSC'13*), Piscataway, NJ, USA, 48–53 (2013).
26. Koç, Y., Warnier, M., Kooij, R. E. & Brazier, F. M. T. An entropy-based metric to quantify the robustness of power grids against cascading failures. *Safety Science* **59,** 126–134 (2013).
27. Wang, J. W., Jiang, C. & Qian, J. Robustness of Internet under targeted attack: a cascading failure perspective. *Journal of Network and Computer Applications* **40,** 97–104 (2014).
28. Dawkins, R. *The Selfish Gene*, Oxford University Press. Oxford. (1989).
29. Ong, Y. S. & Keane, A. J. Meta-Lamarckian learning in memetic algorithms. *IEEE Trans. Evol. Comput*. **8,** 99–100 (2004).
30. Ong, Y. S., Lim, M. H. & Chen, X. S. Research frontier: memetic computation-past, present & future. *IEEE Comput. Intell. Mag*. **5,** 24–36 (2010).
31. Zhou, M. & Liu, J. A memetic algorithm for enhancing the robustness of scale-free networks against malicious attacks. *Phys. A*. **410,** 131 (2014).
32. Newman, M. E. J. Scientific collaboration networks. II. Shortest paths, weighted networks, and centrality. *Phys. Rev. E.* **64,** 016132 (2001).
33. Holme, P. & Kim, B. J. Vertex overload breakdown in evolving networks. *Phys. Rev. E*. **65,** 066109 (2002).
34. Schneider, C. M., Moreira, A. A., Andrade, J. S., Havlin, S. & Herrmann, H. J. Mitigation of malicious attacks on networks. *Proc. Natl. Acad. Sci. USA* **108,** 3838–3841 (2011).
35. Tanizawa, T., Havlin, S. & Stanley, H. E. Robustness of onion-like correlated networks against targeted attacks. *Phys. Rev. E*. **85,** 046109 (2012).
36. Buesser, P., Daolin, F. & Tomassini, M. Optimizing the robustness of scale-free networks with simulated annealing. *ICANNGA*, Part II, LACS **6594,** 167–176 (2011).
37. Kirkpatrick, S., Gelatt, C. D. & Vecchi, M. P. Optimization by simulated annealing. *Science* **220,** 671–680 (1983).

38. Newman, M. E. J. Assortative mixing in networks. *Phys. Rev. Lett.* **89,** 208701 (2002).
39. Latora, V. & Marchiori, M. Efficient behavior of small-world networks. *Phys. Rev. Lett.* **87,** 198701 (2001).
40. Zhou, Q. & Bialek, J. W. Approximate model of European interconnected system as a benchmark system to study effects of cross-border trades. *IEEE Trans. Power Syst.* **20,** 782–788 (2005).
41. Batagelj, V. & Mrvar, A. Pajek – program for large network analysis. *Connections* **21,** 47–57 (1998), available at http://vlado.fmf.uni-lj.si/pub/networks/data.

Acknowledgements

This work is partially supported by the Outstanding Young Scholar Programme of the National Natural Science Foundation of China (NSFC) under Grant 61522311, the General Programme of NSFC under Grant 61271301, the Overseas, Hong Kong & Macao Scholars Collaborated Research Programme of NSFC under Grant 61528205, the Research Fund for the Doctoral Programme of Higher Education of China under Grant 20130203110010, and the Fundamental Research Funds for the Central Universities under Grant K5051202052.

Author Contributions

X. T. and J. L. designed the study, X. T. and X. H. performed the research, X. T., J. L. and X. H. interpreted the results, and X. T. and J. L. wrote the manuscript. All authors reviewed the manuscript.

Additional Information

Competing financial interests: The authors declare no competing financial interests.

How to cite this article: Tang, X. *et al.* Mitigate Cascading Failures on Networks using a Memetic Algorithm. *Sci. Rep.* **6**, 38713; doi: 10.1038/srep38713 (2016).

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.