PLOS | ONE

RESEARCH ARTICLE

Attack Vulnerability of Network Controllability

Zhe-Ming Lu*☯, Xin-Feng Li☯

School of Aeronautics and Astronautics, Zhejiang University, Hangzhou, P. R. China

☯ These authors contributed equally to this work.
* zheminglu@zju.edu.cn

Abstract

Controllability of complex networks has attracted much attention, and understanding the robustness of network controllability against potential attacks and failures is of practical significance. In this paper, we systematically investigate the attack vulnerability of network controllability for the canonical model networks as well as the real-world networks subject to attacks on nodes and edges. The attack strategies are selected based on degree and betweenness centralities calculated for either the initial network or the current network during the removal, among which random failure is as a comparison. It is found that the node-based strategies are often more harmful to the network controllability than the edge-based ones, and so are the recalculated strategies than their counterparts. The Barabási-Albert scale-free model, which has a highly biased structure, proves to be the most vulnerable of the tested model networks. In contrast, the Erdős-Rényi random model, which lacks structural bias, exhibits much better robustness to both node-based and edge-based attacks. We also survey the control robustness of 25 real-world networks, and the numerical results show that most real networks are control robust to random node failures, which has not been observed in the model networks. And the recalculated betweenness-based strategy is the most efficient way to harm the controllability of real-world networks. Besides, we find that the edge degree is not a good quantity to measure the importance of an edge in terms of network controllability.

Citation: Lu Z-M, Li X-F (2016) Attack Vulnerability of Network Controllability. PLoS ONE 11(9): e0162289. doi:10.1371/journal.pone.0162289

Editor: Wen-Bo Du, Beihang University, CHINA

Received: April 20, 2016

Accepted: August 19, 2016

Published: September 2, 2016

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Funding: The authors received no specific funding for this work.

Competing Interests: The authors have declared that no competing interests exist.

Introduction

Complex networks are ubiquitous in nature and society describing various systems [1–3], the dynamics taking place on them has become one of the most popular research fields in the past decades [4–6]. Most previous works have been focused on dynamics such as epidemic spreading [7, 8], synchronization [6, 9, 10], random walk [11, 12], opinion formation [13, 14], and so on. However, our ultimate goal of studying complex networks is to develop the capacity to control them [15]. That is, driving the network from any initial state to any desired final state in finite time [15–18]. Although great efforts [19–24] have been devoted to understanding the controllability of networks, the results are less than satisfactory. Recently,

Liu et al. [15] made a significant breakthrough, they applied the structural control theory [25] to directed networks and proved that the minimal number of nodes needed to fully control a network is determined by its 'maximum matching' [26]. Following this seminal work, some intensive and extensive issues have been carefully addressed, such as controlling linear edge dynamics [27], effects of topological characteristics on controllability [28], energy needed for control [29], control centrality [30] and controllability optimization [31]. Besides, given the limitation that the structural controllability framework is only applicable to directed networks, Yuan et al. [32] presented the exact controllability framework to explore the controllability of arbitrary networks, especially for undirected networks and networks with exact link weights.

Most of the existing works assume that networks are in a relatively safe environment. However, the real-world networks are always confronting with random or intentional node or edge attacks. For example, in the computer networks [33], node attacking can be interpreted as breakdowns of servers by malicious hackers while edge attacking may correspond to the cutting-off of communication cables. In power grids [34], attacks on nodes can be interpreted as substation failures while attacks on edges may correspond to the cases that the connections of subsections are cut off so the power cannot be transmitted from one substation to another. Previous studies have shown that random failures and intentional attacks can easily damage network functions such as connectivity [35] and synchronization [36]. Therefore, it is natural and interesting to ask how the attacks will affect the controllability of networks. Liu et al. [15] first addressed this problem using core percolation, pointing out that the robustness of network controllability is closely related to its core. Pu et al. [37] investigated the robustness of control under node-based cascade failures and found that even if a small range of node failures can trigger great harm to the network controllability. Inspired by Pu et al.'s work, Nie et al. [38] studied edge-based cascade failures and showed that the larger scale of cascades in scale-free networks does not mean that there will be more increments of driver nodes.

In addition to cascade attacks [39], prominence based attacks, especially degree and betweenness based attacks, are the most common attacks in practice [35]. For example, it has been shown that even if 1% of the most highly connected routers were incapacitated, the average performance of Internet would drop 50% [40]. Therefore, it is of practical significance to study the robustness of network controllability subject to this kind of attacks. In this paper, we systematically investigate the attack vulnerability of network controllability for the canonical model networks as well as the real-world networks subject to degree and betweenness based attacks on nodes and edges. For each case of attacks, five different strategies are employed and the network controllability is quantitatively measured by the fraction of driver nodes. Furthermore, we also investigate the control robustness of 25 real-world networks.

The rest of the paper is organized as follows. Section II gives a brief review of network controllability. Section III introduces our methods, including definitions, attack strategies and benchmark networks. Our main results and discussions are presented in Section IV. Finally, Section V concludes the whole paper.

Network Controllability

In this section, we briefly review the controllability of complex networks, both the structural controllability framework [15] and the exact controllability framework [32] are introduced. The former is only applicable to directed networks whereas the latter can treat arbitrary networks without any limitations [32].

Structural Controllability

Consider a network of N nodes governed by the following linear time invariant dynamics [15]:

$$\dot{\boldsymbol{x}}(t) = A\boldsymbol{x}(t) + B\boldsymbol{u}(t) \tag{1}$$

where $\boldsymbol{x}(t) = (x_1(t), \ldots, x_N(t))^T$ stands for the states of nodes at time t, the $N \times N$ matrix A denotes the coupling strength between N nodes, where the element a_{ij} gives the strength or weight that node-j can affect node-i. $\boldsymbol{u}(t) = (u_1(t), u_2(t), \ldots, u_M(t))^T$ is the vector of input signals and the $N \times M$ ($M \leq N$) matrix B is called input matrix which defines how input signals are connected to the nodes in the network, which are often called 'driver nodes'.

According to the classic Kalman rank condition [16, 41], the system described by Eq (1) is controllable if and only if

$$\mathrm{rank}(C) = \mathrm{rank}([B, AB, A^2B, \ldots, A^{N-1}B]) = N \tag{2}$$

where $C = [B, AB, A^2B, \ldots, A^{N-1}B]$ is called controllability matrix. Since the matrix A is fixed for a given network, in order to make the network fully controllable, one need to choose a suitable matrix B having the minimal number of driver nodes to satisfy the Kalman rank condition. However, the practical difficulty lies in that there are $2^N - 1$ possible combinations of selecting driver nodes. To say the least, even one can enumerate all the combinations efficiently, the link weights (a_{ij}) are often not known for real-world networks. In order to overcome the inherently incomplete information of link weights, Liu et al. [15] introduced the concept of 'structural controllability' [25] to complex networks. The structural controllability can ensure that a network is controllable for almost all weight combinations except for some pathological cases [15]. The authors also developed the minimum input theory to avoid brute-force searching, which states that the minimal number of driver nodes needed to fully control a network, N_D, is determined by the maximum matching [26] in the network [15], where the unmatched nodes are exactly driver nodes. This theory allows one to find the driver nodes within $O(\sqrt{N}L)$ rather than $O(2^N)$ time, where L denotes the number of links [15].

Liu et al.'s work also showed that a network's controllability is mainly determined by its underlying degree distribution, networks that are sparse and heterogeneous are more difficult to control [15]. Let $n_D = N_D/N$ denote the density of driver nodes, for Erdős-Rényi (ER) networks [42] with mean degree $\langle k \rangle$, Liu et al. give n_D analytically as [15]

$$n_D \approx e^{-\frac{\langle k \rangle}{2}} \tag{3}$$

While for scale-free networks [2] with mean degree $\langle k \rangle$ and degree exponent $\gamma = \gamma_{in} = \gamma_{out}$, n_D is given by [15]

$$n_D \approx \exp\left[-\frac{1}{2}\left(1 - \frac{1}{\gamma - 1}\right)\langle k \rangle\right] \tag{4}$$

Exact Controllability

The exact controllability framework was proposed by Yuan et al. [32] to overcome the limitation that the structural controllability framework can only treat directed networks. As an alternative, the new paradigm can be applied to arbitrary network structures and link weights without any limitations [32].

Instead of empolying the Kalman rank condition, the exact controllability framework is based on the equivalent Popov-Belevitch-Hautus (PBH) rank condition [43], which stipulates

that the system (Eq (1)) is controllable, if and only if

$$\mathrm{rank}(\psi I_N - A, B) = N \tag{5}$$

holds for any complex number ψ, where I_N is the $N \times N$ identify matrix. It can be further proved that full control can be guaranteed if and only if all the eigenvalues λ of A satisfy Eq (5) [32]. From the perspective of matrix, the minimal number of driver nodes, N_D, is defined by matrix B with $N_D = \min\{\mathrm{rank}(B)\}$. Equivalently, Yuan et al. proved that for any network with arbitrary matrix A, N_D is actually determined by the maximum geometric multiplicity $\mu(\lambda_i)$ of the eigenvalue λ_i of A [32], i.e.

$$N_D = \max_i\{\mu(\lambda_i)\} \tag{6}$$

where $\lambda_i (i = 1, \ldots, l)$ is the nonidentical eigenvalues of A and $\mu(\lambda_i) = N - \mathrm{rank}(\lambda_i I_N - A)$ is the geometric multiplicity of λ_i. In particular, for undirected networks, N_D is simply determined by the maximum algebraic multiplicity $\delta(\lambda_i)$ of λ_i [32], i.e.

$$N_D = \max_i\{\delta(\lambda_i)\} \tag{7}$$

For large sparse networks, N_D can be dramatically simplified in terms of $\mathrm{rank}(A)$ [32], i.e.

$$N_D = \max\{1, N - \mathrm{rank}(A)\} \tag{8}$$

Methods

In this section, we introduce our experimental methods, which include definitions, attack strategies and benchmark networks.

Definitions

In this paper, we allow the existence of self loops in the network but multiple (parallel) edges are not allowed. We also denote the number of nodes as N and the number of edges as L.

The measure of controllability n_D, or simply controllability, of a network is defined as the ratio of the minimum number of driver nodes N_D to the network size N [15], i.e.

$$n_D = N_D / N \tag{9}$$

The degree k_v of a node v is defined as the number of its direct connections to other nodes. The degree k_e of an edge e is defined as

$$k_e = k_u k_v \tag{10}$$

where e connects nodes u and v with degrees k_u and k_v.

The betweenness centrality of a node v is defined as

$$C_B(v) = \sum_{s \neq v \neq t \in V} \frac{\sigma_{st}(v)}{\sigma_{st}} \tag{11}$$

where σ_{st} is the total number of shortest paths from node s to node t and $\sigma_{st}(v)$ is the number of those paths that pass through v. Similarly, the edge betweenness centrality of an edge e is defined as

$$C_B(e) = \sum_{s \neq t \in V} \frac{\sigma_{st}(e)}{\sigma_{st}} \tag{12}$$

where $\sigma_{st}(e)$ is the number of shortest paths from node s to node t that include the edge e. Throughout the present paper, we call $C_B(v)$ and $C_B(e)$ the node betweenness and the edge betweenness for brevity.

Attack Strategies

Here we consider five different attack strategies for nodes and edges. The nodes are removed in the descending order of degree or betweenness centrality calculated for either the initial network or the current network during the removal procedure. Random attack (also called random failure) is taken as a comparison. The five node attack strategies are described as follows:

- Random Attack (RA): to remove node one by one randomly.
- Initial Degree Attack (ID): to remove node one by one in the descending order of degree using the degree distribution of the initial network.
- Initial Betweenness Attack (IB): to remove node one by one in the descending order of betweenness using the betweenness distribution of the initial network.
- Recalculated Degree Attack (RD): to remove node one by one in the descending order of degree using the recalculated degree distribution at every removal step.
- Recalculated Betweenness Attack (RB): to remove node one by one in the descending order of betweenness using the recalculated betweenness distribution at every removal step.

Note that the degree-based attacks (ID and RD) are local strategies, whereas the betweenness-based attacks (IB and RB) are global strategies. Another significant difference is that the former concentrate on reducing the total edges as fast as possible whereas the latter concentrate on destroying as many as the shortest paths as possible.

The above strategies can also be applied to edge attacks by simply replacing node degree and betweenness with edge degree and betweenness. The attack vulnerability of network controllability subject to such edge attacks is also investigated in this paper.

Networks

We select four canonical model networks as well as two real networks as benchmarks to study the robustness of network controllability. All the six networks are undirected and in the following experiments the exact controllability framework [32] will be adopted to calculate their controllability.

The ER model [42] is a classic random network with a Poisson-type degree distribution $P(k) = e^{-\langle k\rangle}\langle k\rangle^k/k!$, a logarithmically increasing average path length and a small clustering coefficient close to zero. The Watts-Strogatz (WS) model [1] and Newman-Watts (NW) model [44] are classic small world networks, which have short average path lengths ($\sim \ln N$) and high clustering coefficients. Both of them are constructed from a regular ring lattice with each node linking to its left and right $K/2$ nearest neighbors, the only difference is that the former rewires edges with probability p whereas the latter adds edges with probability p. The shape of their degree distribution is similar to that of ER, having a pronounced peak at $k = K$ and decaying exponentially for large $|k - K|$. The Barabási-Albert (BA) model [2] is a representative scale free network with the power-law degree distribution $P(k) \sim k^{-3}$.

The two real-world networks are the USAir97 network and the Erdos971 network. The USAir97 [45] is a network formed by the direct air route between the American airports in 1997. It has 332 nodes and 2126 edges as shown in Fig 1(a). Each node stands for an airport and each directed edge represents an airline from one airport to another. The weight of an edge

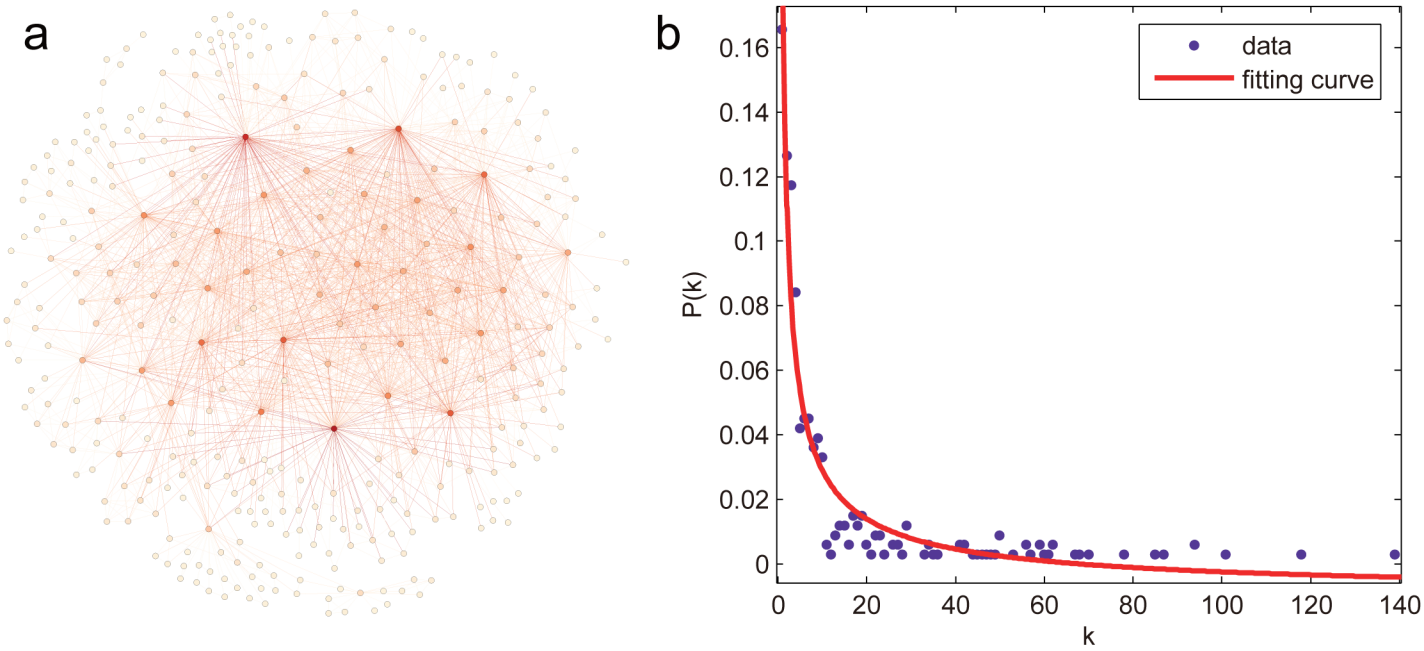

Fig 1. The USAir97 network. (a) Its topology, in which the darkness of node color is proportional to its degree. (b) Its degree distribution, where the points represent the frequency distribution of degree, the red line denotes the fitting curve with equation $P(k) = 0.1948k^{-0.6959} - 0.01029$, the goodness of fit is 0.9198.

doi:10.1371/journal.pone.0162289.g001

represents the number of seats available on the scheduled flights. USAir97 exhibits an approximate power-law degree distribution as shown in Fig 1(b). The airports with less connections to/from other airports follow more closely with the fitting curve whereas the airports with more connections show a slight deviation.

The Erdos971 [45] is a scientific collaboration network where each node represents an scientist who co-authored at least one paper with Paul Erdős [42], and two scientists are joined by an edge if they co-authored a paper. Note that the node corresponding to Paul Erdős himself and all the isolated components have been removed throughout the paper, the truncated network contains 429 nodes and 1312 edges. Its topology and degree distribution are displayed in Fig 2(a) and 2(b), respectively.

The parameter settings and summaries of the benchmark networks are shown in Table 1.

Results and Discussions

Node Attack

Firstly, we investigate the control robustness of the ER random network under node attacks and show the results in Fig 3. It can be seen that for all the attacks, n_D increases with the removal fraction f, indicating that we need constantly increase driver nodes to fully control the network, i.e., the network's controllability is decreasing. The intentional attacks (ID, RD, IB and RB) are much more harmful than RA as their n_D increases much faster. For the former attacks, the difference among them is not significant in the early stage of removals ($f \lesssim 0.07$), which can be attributed to the high degree-betweenness correlation (high degree nodes tend to have high betweenness) [35] and the negligible redistribution of degree and betweenness

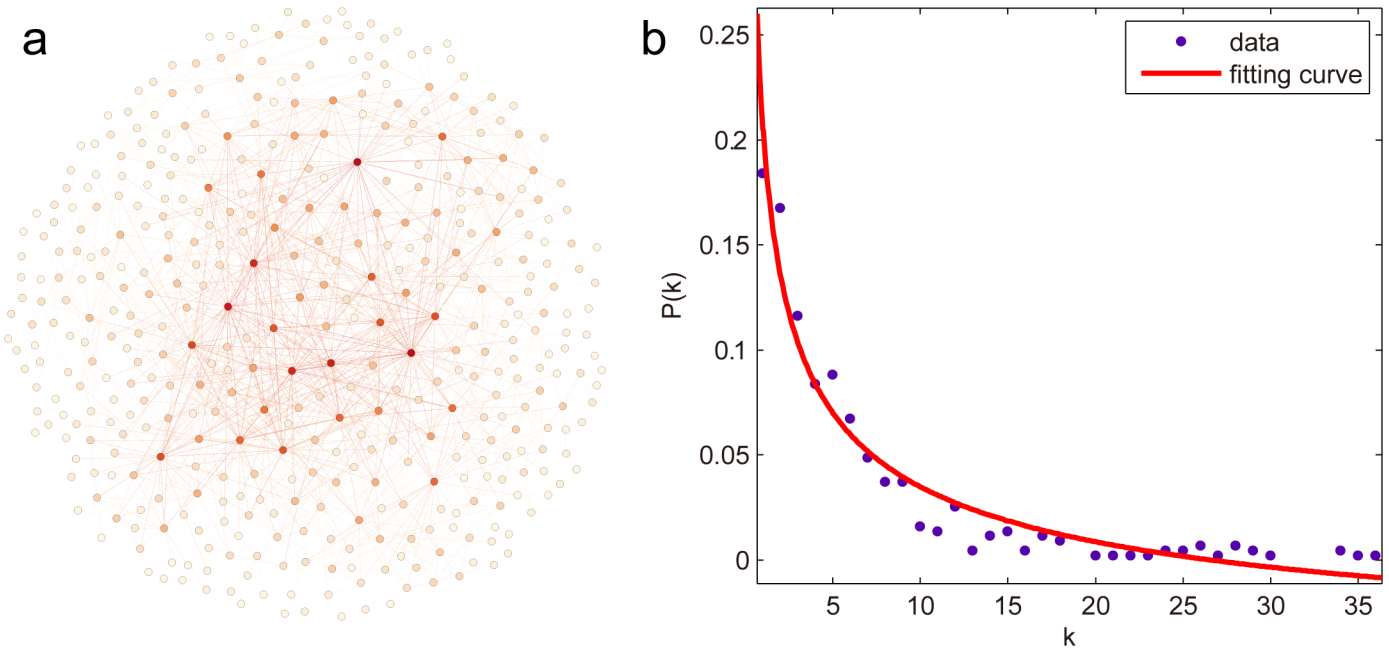

Fig 2. The Erdos971 network. (a) Its topology, in which the darkness of node color is proportional to its degree. (b) Its degree distribution, where the points represent the frequency distribution of degree, the red line denotes the fitting curve with equation $P(k) = 0.2727k^{-0.4256} - 0.06745$, the goodness of fit is 0.9434.

doi:10.1371/journal.pone.0162289.g002

Table 1. The summaries of benchmark networks used in this paper.

Type	Network	N	L	$\langle k \rangle$
Random Network	ER	200	400	4.0
Small World	WS	200	400	4.0
Small World	NW	200	440	4.4
Scale Free	BA	200	591	5.91
Transportation	USAir97	332	2126	6.4036
Scientific Collaboration	Erdos971	429	1312	6.1166

All the networks are undirected and for each network, we show its type, name, number of nodes (N), number of edges (L), and average degree ($\langle k \rangle$). Note that the original Erdos971 network has 472 nodes and 1314 edges, here we only keep its largest connected component.

doi:10.1371/journal.pone.0162289.t001

information for small f. However, as the removals proceed ($f > 0.07$), they gradually separate by harming the network controllability in the order RD $>$ RB $\simeq$ ID $>$ IB (the inequality RD $>$ RB means that RD is more harmful than RB), revealing that degree-based attacks are more harmful than betweenness-based attacks (RD $>$ RB and ID $>$ IB) and the attacks based on recalculated information are, as expected, more harmful than their counterparts based on initial information (RD $>$ ID and RB $>$ IB), which confirms the previous conclusions that network controllability is mainly determined by degree distribution [15] but distance based measures such as betweenness can also affect the controllability [46].

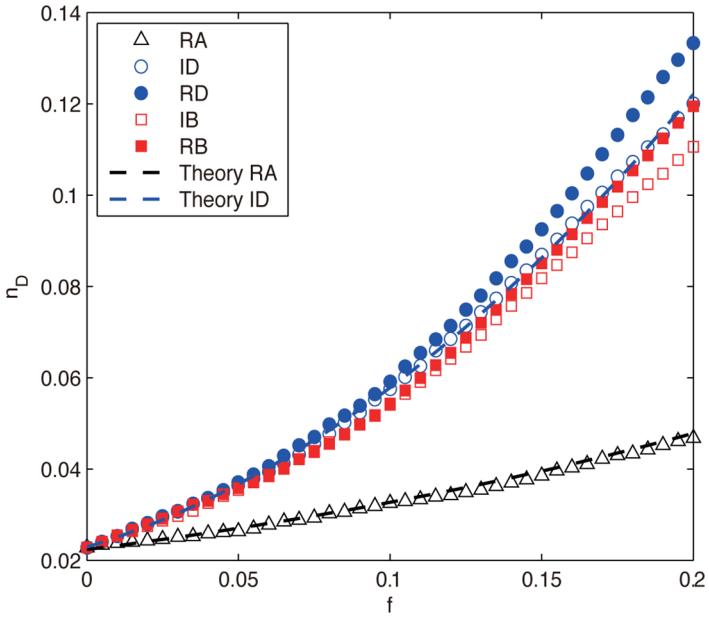

Fig 3. The density of driver nodes n_D as a function of removal fraction f for ER network under different node attacks. The black and blue dashed lines are theoretical results obtained from Eqs (24) and (26), respectively. The numerical results are averaged over 100 independent realizations.

doi:10.1371/journal.pone.0162289.g003

For RA and ID attacks, we can give the analytical results of n_D by employing the cavity method [15]. The ER network follows a Possion degree distribution $P(k) = e^{-\langle k\rangle}\langle k\rangle^k/k!$, indicating that most nodes have almost the same degree close to $\langle k\rangle^0$, the average degree of the original network. Therefore, for RA it is reasonable to assume that every removed node has degree $\langle k\rangle^0$, then the average degree after removals $\langle k\rangle'_{\mathrm{RA}}$ is

$$\langle k\rangle'_{\mathrm{RA}} = \langle k\rangle^0 - f\langle k\rangle^0 = (1-f)\langle k\rangle^0 \tag{13}$$

According to Liu et al.'s work [15], for networks where in-degrees and out-degrees share the similar distribution, n_D can be obtained by

$$n_D = G(w_2) + G(1-w_1) - 1 + \langle k\rangle w_1(1-w_2) \tag{14}$$

where $G(x) = \sum_{k=0}^{\infty} P(k)x^k$ is the predefined generating function, w_1 and w_2 can be derived from the following equations:

$$w_1 = H[1 - H(1-w_1)] \tag{15}$$

$$w_2 = 1 - H[1 - H(w_2)] \tag{16}$$

where $H(x) = \sum_{k=0}^{\infty} Q(k+1)x^k$ with $Q(k) = kP(k)/\langle k\rangle$ [31].

For RA, since the removal fraction is small ($f \leq 0.2$), we assume that the degree distribution $P(k)$ is not significantly affected. By plugging $\langle k \rangle'_{\mathrm{RA}}$ into $P(k)$ and $P(k)$ into $G(x)$, one has

$$G(x) = \sum_{k=0}^{\infty} \frac{e^{-\langle k \rangle^0 (1-f)} [\langle k \rangle^0 (1-f)]^k}{k!} x^k = e^{-\langle k \rangle^0 (1-f)(1-x)} \tag{17}$$

Similarly, by plugging $P(k)$ into $Q(k)$ and $Q(k)$ into $H(x)$, one has

$$H(x) = \sum_{k=0}^{\infty} \frac{e^{-\langle k \rangle^0 (1-f)} [\langle k \rangle^0 (1-f)]^k}{k!} x^k = e^{-\langle k \rangle^0 (1-f)(1-x)} \tag{18}$$

From Eqs (15) and (16) and the monotonicity of $H(x)$, one can easily check that $w_1 = H(w_2)$ and $w_2 = 1 - H(1 - w_1)$. Substituting $H(x)$ gives

$$w_1 = e^{-\langle k \rangle^0 (1-f)[H(1-w_1)]} = e^{-\langle k \rangle^0 (1-f)[1-w_2]} \tag{19}$$

$$w_2 = 1 - e^{-\langle k \rangle^0 (1-f) H(w_2)} = 1 - e^{-\langle k \rangle^0 (1-f) w_1} \tag{20}$$

where w_1 can be further reduced by substituting w_2 into Eq (19), as follows

$$w_1 = \exp\left[-\langle k \rangle^0 (1-f) e^{-\langle k \rangle^0 (1-f) w_1}\right] \tag{21}$$

Now n_D can be simplified as

$$\begin{aligned} n_D &= G(w_2) + G(1-w_1) - 1 + \langle k \rangle w_1 (1-w_2) \\ &= e^{-\langle k \rangle^0 (1-f)[1-w_2]} + e^{-\langle k \rangle^0 (1-f) w_1} - 1 + \langle k \rangle^0 (1-f) w_1 (1-w_2) \\ &= w_1 - w_2 + \langle k \rangle^0 (1-f) w_1 (1-w_2) \end{aligned} \tag{22}$$

For $k \gg 1$, one has $w_1 \sim e^{-\langle k \rangle^0 (1-f)}$, $w_2 = 1 - e^{-\langle k \rangle^0 (1-f) w_1} \sim \langle k \rangle^0 (1-f) w_1$, and thus

$$n_D \sim \exp\left[-\langle k \rangle^0 (1-f)\right] - \left(\langle k \rangle^0 (1-f)\right)^2 \exp\left[-2\langle k \rangle^0 (1-f)\right] \tag{23}$$

Ignoring the higher order term in Eq (23) gives the final n_D as

$$n_D \sim \exp\left[-\langle k \rangle^0 (1-f)\right] \tag{24}$$

For ID attack, let $k^{\max}$ denote the maximum degree after removing Nf high degree nodes, then $f = \int_{k^{\max}}^{\infty} P(k)dk$. The average degree after removals, $\langle k \rangle'_{\mathrm{ID}}$, can be derived as follows

$$\begin{aligned}
\langle k \rangle'_{\mathrm{ID}} &= \int_0^{k^{\max}} kP(k)dk = \int_0^{k^{\max}} \frac{e^{-\langle k \rangle^0}(\langle k \rangle^0)^k}{(k-1)!}dk \\
&= \langle k \rangle^0 \int_0^{k^{\max}} \frac{e^{-\langle k \rangle^0}(\langle k \rangle^0)^{k-1}}{(k-1)!}d(k-1) \\
&= \langle k \rangle^0 \int_{-1}^{k^{\max}-1} \frac{e^{-\langle k \rangle^0}(\langle k \rangle^0)^k}{k!}dk = \langle k \rangle^0 \int_0^{k^{\max}-1} \frac{e^{-\langle k \rangle^0}(\langle k \rangle^0)^k}{k!}dk \\
&= \langle k \rangle^0 \left(\int_0^{k^{\max}} \frac{e^{-\langle k \rangle^0}(\langle k \rangle^0)^k}{k!}dk - \int_{k^{\max}-1}^{k^{\max}} \frac{e^{-\langle k \rangle^0}(\langle k \rangle^0)^k}{k!}dk \right) \\
&= \langle k \rangle^0 \left(\int_0^{k^{\max}} P(k)dk - \int_{k^{\max}-1}^{k^{\max}} P(k)dk \right) \\
&= \langle k \rangle^0 \left(1 - f - \int_{k^{\max}-1}^{k^{\max}} P(k)dk \right)
\end{aligned} \tag{25}$$

For simplicity, let $\int_{k^{\max}-1}^{k^{\max}} P(k)dk = \alpha f$ where α is a constant coefficient, we have $\langle k \rangle'_{\mathrm{ID}} = \langle k \rangle^0 (1 - (1+\alpha)f)$. Using the similar method, n_D can be approximated as

$$n_D \sim \exp\left[-\langle k \rangle^0 (1 - (1+\alpha)f)\right] \tag{26}$$

where the specific α can be obtained by curve fitting. From Fig 3 we can see that the theoretical results of Eqs (24) and (26) agree well with their corresponding numerical results.

Next, we compare the two small world model networks, WS and NW, both of which have exponential cutoffs in the degree distribution. From Fig 4, we can see that the two networks exhibit very similar vulnerability behavior: n_D increases as the removals proceed, RB harm the controllability most, followed by the degree-base attacks and IB, RA is the least harmful strategy (RB > ID, RD > IB > RA). This phenomenon is mainly due to their similar construction method that both are generated from the one dimensional regular ring lattice. The most interesting and unexpected behavior is that for both networks, RB proves to be the most harmful strategy and the superiority even becomes more obvious for large f ($f > 0.1$ for WS, $f \gtrsim 0.15$ for NW), which is beyond our expectation that degree-based attacks (at least RD) should be more harmful than betweenness-based ones due to the decisive role of degree distribution to network controllability.

To explain this behavior, we explore the correlation between the node betweenness C_B and the node degree k as shown in Fig 5a, from which we can see that the correlation is quite weak, which excludes the possibility that degree helps to contribute the effects of RB. Recall that the two prominent characteristics of small world networks are the short average path length (APL) and the high clustering. Since recent study [28] has shown that clustering has no discernible impact on network controllability, thus the effectiveness of RB can only be attributed to the small APL, which again confirms that betweenness also affects network's controllability [46]. Our results clearly show that RB is the most efficient way to destroy the controllability of small world networks rather than RD.

From Fig 4, we can also see that the degree-based attacks on WS and NW networks behave quite differently from those on the ER network. The n_D value of the former has a clear slow-down trend for large f whereas the latter does not. For WS network, both RD and ID harm the network controllability equally in the interval $0 \leq f \leq p$, after that ID prevails RD but with less

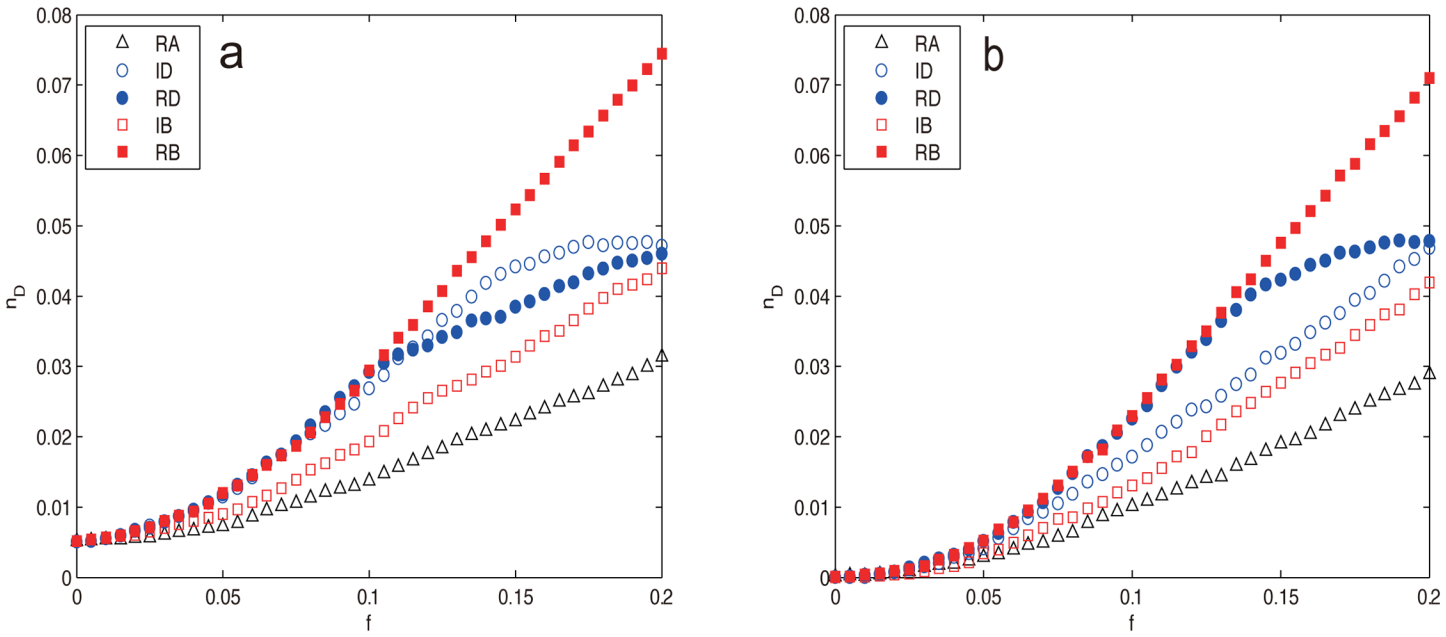

Fig 4. n_D as a function of the removal fraction f under different node attacks for (a) WS network and (b) NW network. The results are averaged over 100 independent realizations.

doi:10.1371/journal.pone.0162289.g004

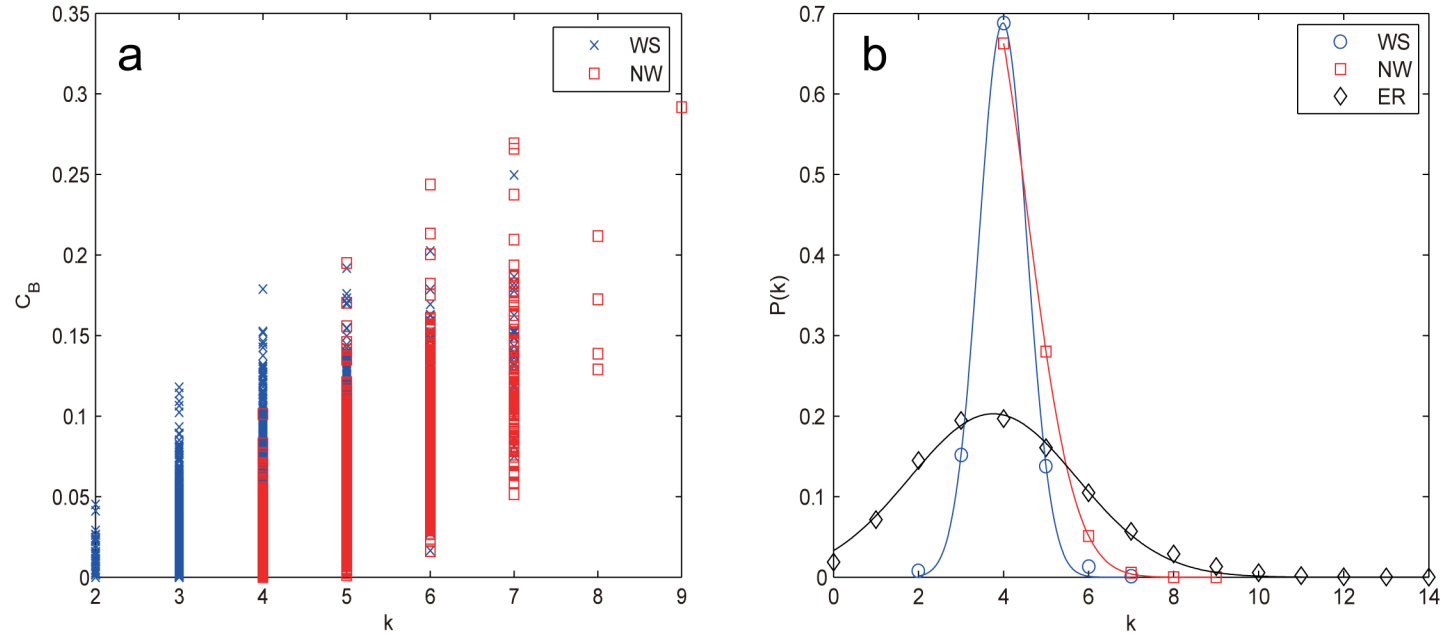

Fig 5. (a) Correlation between the node betweenness C_B and the node degree k for WS and NW networks. (b) The degree distribution of WS, NW and ER network.

doi:10.1371/journal.pone.0162289.g005

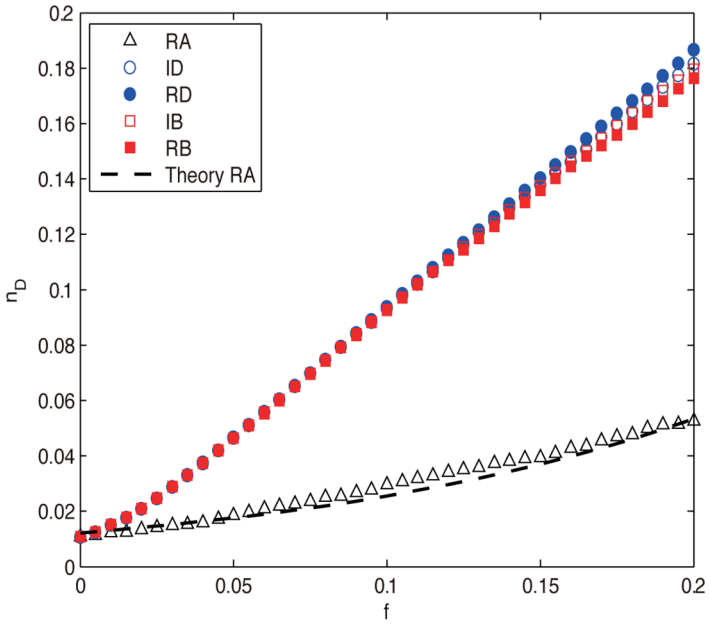

Fig 6. n_D as a function of removal fraction f under different node attacks for the BA scale-free network. The simulation results are averaged over 100 independent realizations, the analytical result of RA is obtained by Eq (27).

doi:10.1371/journal.pone.0162289.g006

and less superiority, finally coincides with the latter at $f = 0.2$. The emergence of watershed ($f \approx p = 0.1$) can be explained since when f exceeds p, the original WS topology is lost, the network degenerates into a regular ring lattice. It is worth noting that WS is the only case where a procedure based on recalculated information is less harmful than its counterpart based on the initial configuration (RD < ID). For NW network, RD retains as harmful as RB for $f \lesssim 0.14$, then slows down for $f > 0.14$, the emergence of the turning point ($f \approx 0.14$) occurs later than that of WS ($f \approx 0.1$) due to its longer tail of degree distribution as shown in Fig 5b resulting from adding instead of rewiring edges.

The BA model is in focus in the first study of the control vulnerability of scale-free networks. From Fig 6, we can see that all the attacks except RA harm the network controllability equally for $f \lesssim 0.15$, after that the degree-based strategies prevail the betweenness-based ones with negligible advantages. This coincidence should not be attributed to the short APL like the WS and NW network but the high correlation between the node betweenness and degree as shown in Fig 7a. It can also be seen that n_D rises from 0.01 to about 0.19 ($\Delta n_D \approx 0.18$), which is much more significant than that of the ER ($\Delta n_D \approx 0.10$), WS ($\Delta n_D \approx 0.07$) and NW ($\Delta n_D \approx 0.07$) network, indicating that the scale free networks are more control vulnerable than both the ER random and small world networks, which is due to the existence of hub nodes resulting from the power-law degree distribution as shown in Fig 7b. Another notable finding is that though Liu et al. have shown that driver nodes tend to avoid to be hub nodes [15], here we show that attacking the hub nodes is still the most efficient way to harm the controllability of scale free networks.

The analytical result of n_D for the RA attack can be obtained by employing the cavity method [15]. The BA network follows a power-law degree distribution with $P(k) \sim 2m^2k^{-3}$ [2], since the removal fraction f is small, we assume that the degree distribution $P(k)$ is not

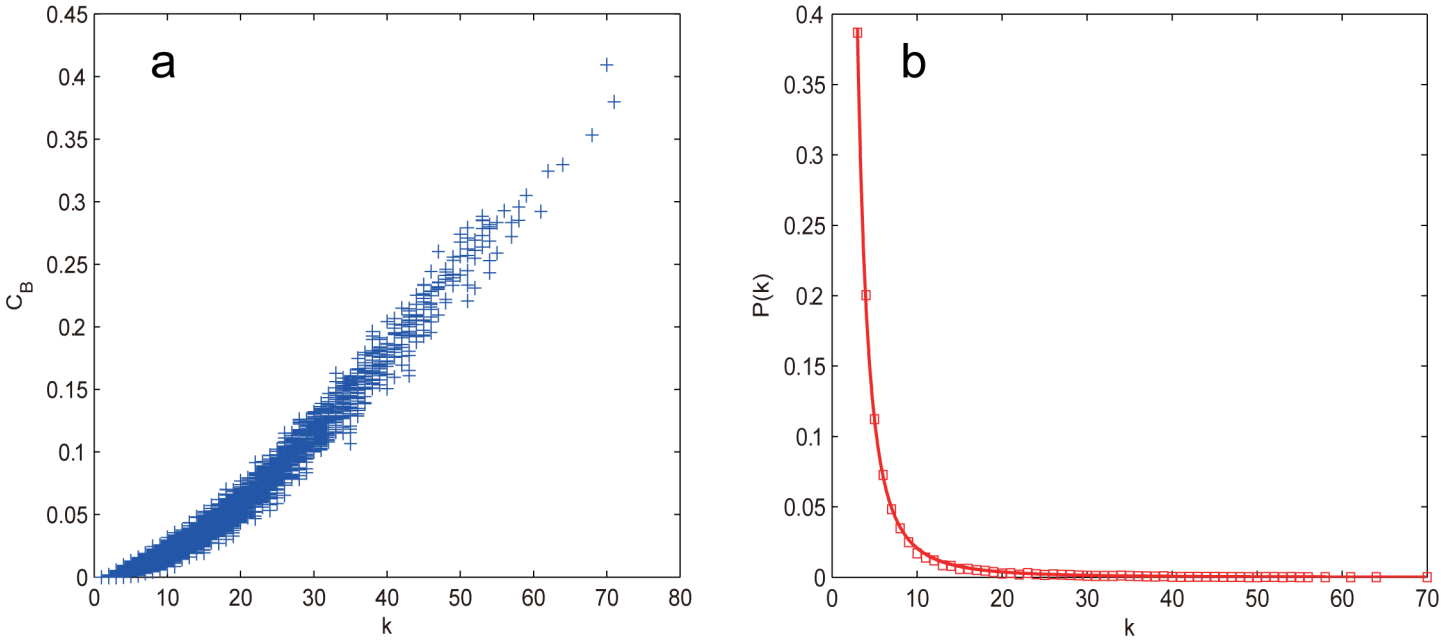

Fig 7. (a) Node betweenness-degree correlation of the BA network. (b) The degree distribution of the BA network.

doi:10.1371/journal.pone.0162289.g007

significantly affected, which gives us [15]

$$n_D \sim \exp\left[-\langle k\rangle^0(1-f)\left(1-\frac{1}{\gamma^0-1}\right)\right] \tag{27}$$

where $\langle k\rangle^0$ and $\gamma^0 = -3$ are the average degree and degree exponent of the original network, respectively. From Fig 6 we can see that the theoretical prediction agrees well with the numerical result with negligible difference.

The two real-world networks, USAir97 and Erdos971, display very similar vulnerable behaviors like the BA scale-free network for the intentional attacks as shown Fig 8, which is mainly due to their power-law degree distributions as shown Figs 1b and 2b. Nevertheless, the differences between them are significant. For USAir97, it can be seen that the intentional attacks harm the network controllability almost equally in the early stage of removal ($f \lesssim 0.12$), which is due to the strong betweenness-degree correlation in the region of high degrees as shown in Fig 9a, after that the strategies based on recalculated information prevail those based on initial configuration (RB, RD > IB, ID) owing to the redistribution of degree and betweenness information. Besides, the betweenness-based strategies outperform those degree-based ones (RB > RD and IB > ID) with very slight superiorities. The most confusing and unexpected behavior is that n_D decreases with f for RA, which means that we need less and less driver nodes to maintain the full control of network, i.e., the network controllability is indeed increasing, this abnormal phenomenon has not been observed in other cases and the reason behind is not clear and needs further research.

Erdos971 exhibits much similar behaviors like USAir97 as shown in Fig 8b. The intentional attacks ID, IB, RD and RB coincide in the early stage of removal ($f \gtrsim 0.08$) due to the strong betweenness-degree correlation in the region of high degrees as shown in Fig 9b; these attacks

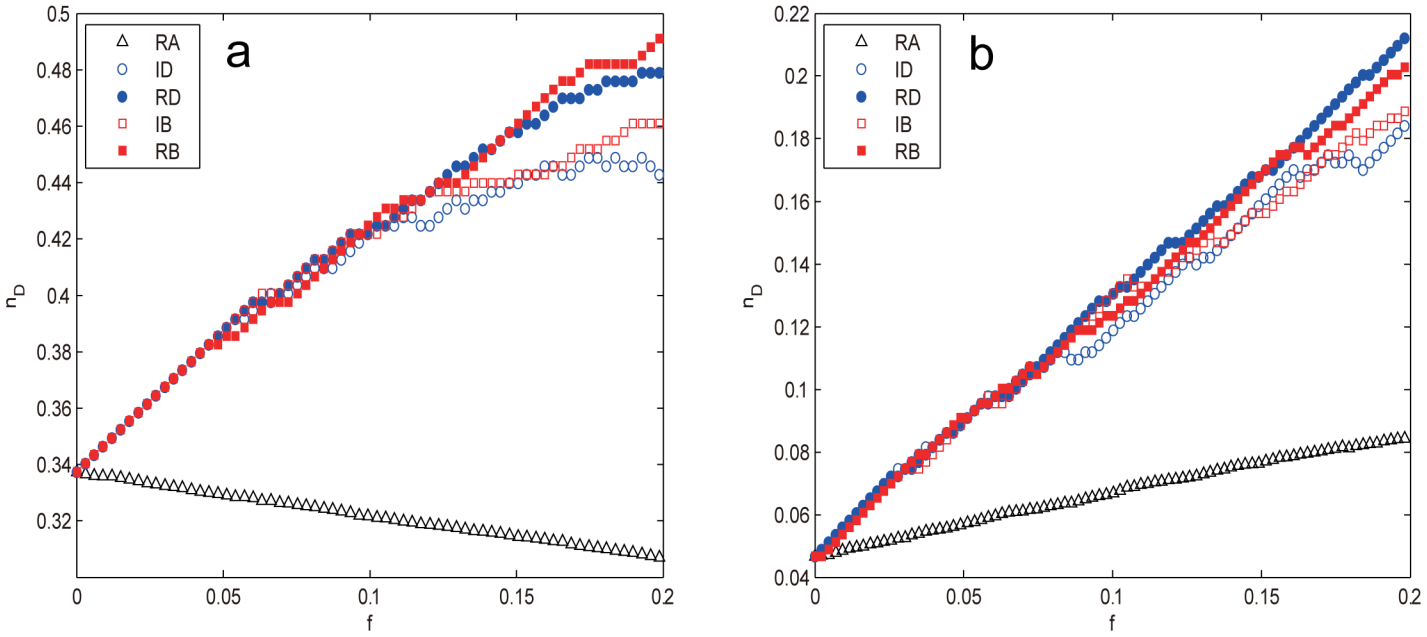

Fig 8. n_D as a function of removal fraction f under different node attacks for the (a) USAir97 and (b) Erdos971 network. The results are averaged over 100 independent runs.

doi:10.1371/journal.pone.0162289.g008

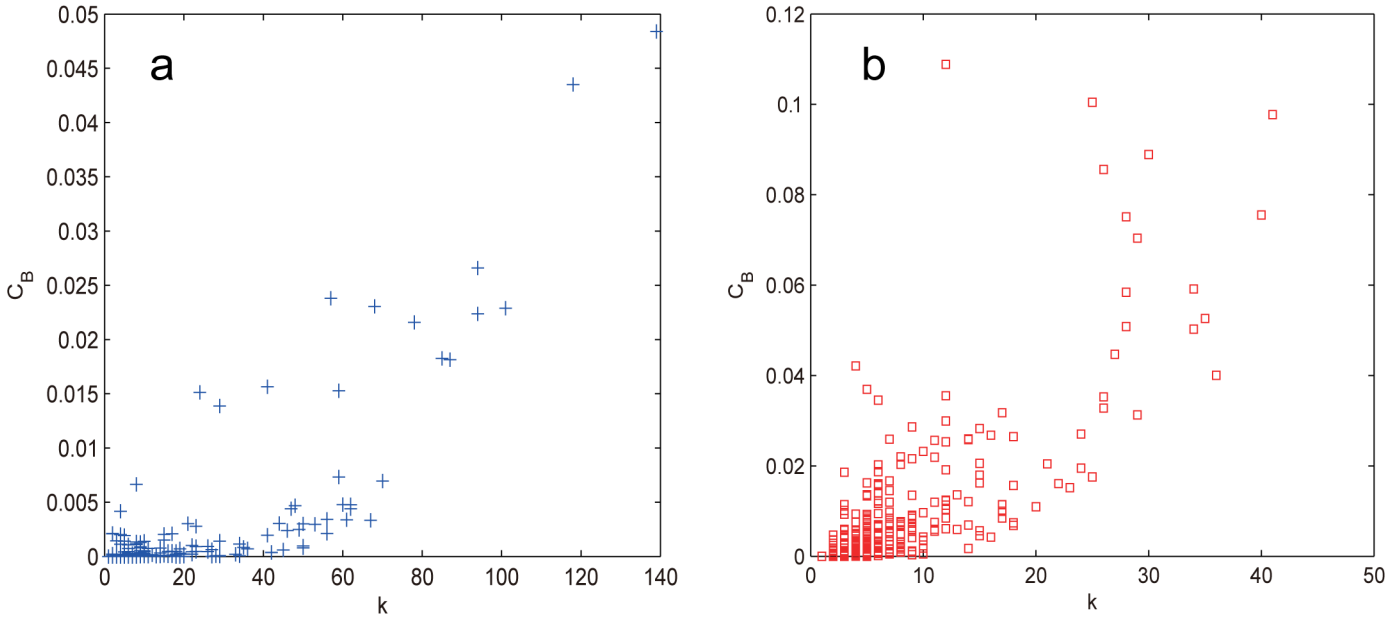

Fig 9. The node betweenness-degree correlation for the (a) USAir97 and (b) Erdos971 network.

doi:10.1371/journal.pone.0162289.g009

also deviate in the last stage of removal ($f \gtrsim 0.17$) in the approximate order RD > RB > IB > ID. The remarkable difference occurs in the interval $0.08 \lesssim f \lesssim 0.17$, where RD prevails RB with less and less advantages and finally coincides with the latter at $f \approx 0.17$, so does IB with ID. The recalculated strategies RD and RB are, as expected, more harmful than their counterparts ID and IB. Like the BA scale-free network, RD proves to be the most efficient way to harm the network's controllability.

Edge Attack

In this subsection, we study the attack vulnerability of network controllability against various edge attacks (see Sec. III for details of the edge attack strategies). Generally speaking, the edge attacks may be not as efficient as the node attacks since after removing Ef edges (Nf nodes), the former only removes $N\langle k\rangle f/2$ edges whereas the latter can at most remove $Nf\langle k\rangle$ edges. It should also be noted that the study presented here is different from that of Ref. [38], in which cascading edge failure is assumed and the removal of one edge may trigger cascading removals of other edges, here we do not make such assumption and the removal of one edge does not affect other edges' removals at all.

Fig 10 shows the numerical results of the control robustness of the ER random network subject to edge strategies. We can see that the edge strategies are indeed less efficient than those node ones: $\Delta n_D \approx 0.03$ for the former compared with $\Delta n_D \approx 0.11$ for the latter (compare Figs 10 and 3). The RB procedure turns out to be the most destructive strategy, even so, it is just a little bit better than RA with very slight superiority (n_D(RBA) − n_D(RA) ≈ 0.003). n_Ds of others attacks (IB, ID and RD) are even inferior to that of RA with growing gaps, indicating that these strategies are even less efficient than random failures. Moreover, we can see that the degree-

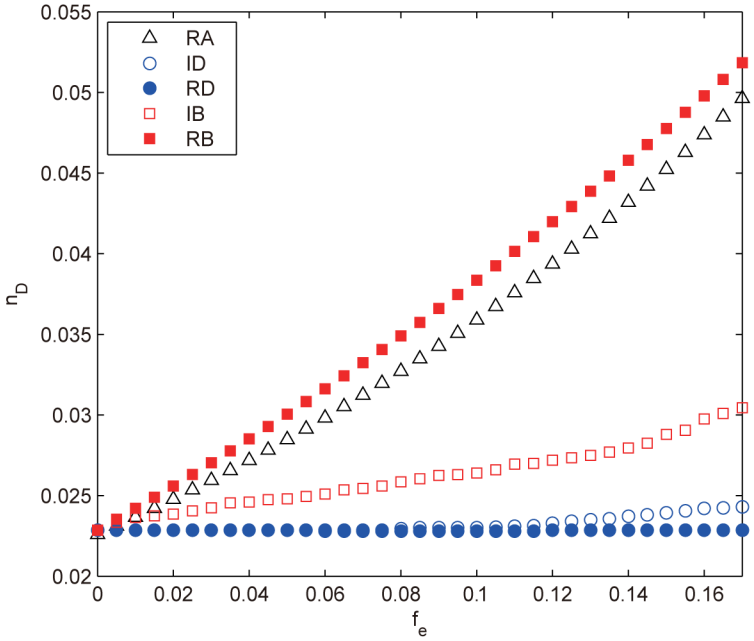

Fig 10. n_D as a function of the removal fraction f_e under different edge attacks for the ER random network. The results are averaged over 100 independent realizations.

doi:10.1371/journal.pone.0162289.g010

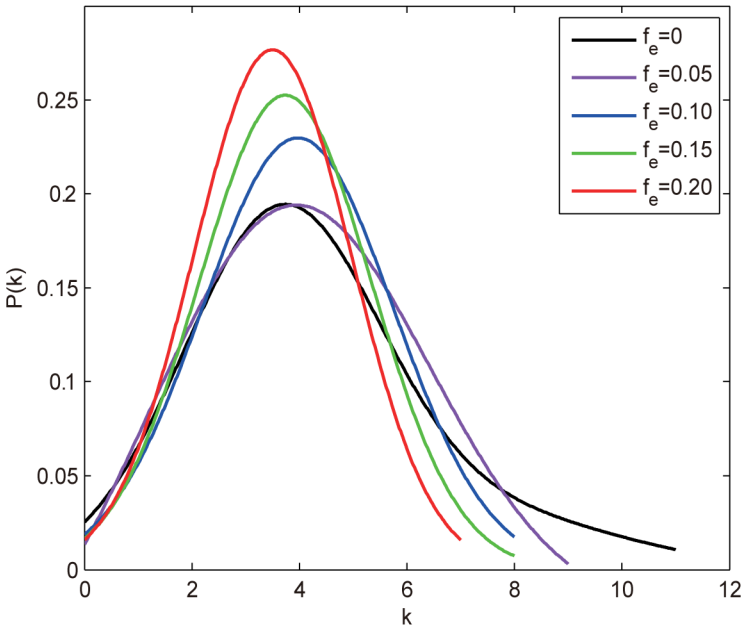

Fig 11. The degree distribution *P*(*k*) of the ER random network under different edge removal fraction f_e subject to the ID attack. f_e = 0 denotes the initial degree distribution.

doi:10.1371/journal.pone.0162289.g011

based attacks, ID and RD, can hardly harm the controllability at all as evident from the almost constant n_D shown in Fig 10.

These phenomena may seem difficult to understand at first glance, however, they can be explained by carefully analyzing the structural characteristics of ER. The ER random network has a short APL ($\sim \ln N/\ln\langle k\rangle$) due to its high proportion of shortcuts, such edges are removed by RB for their high betweenness and also have a great chance to be removed by RA. Therefore both RB and RA remove almost the same set of edges, resulting in almost the same damage to network controllability. This result may suggest that the distance measures like edge betweenness also affect the networks controllability. Given the importance of shortcuts in small-world networks, we infer that the edge betweenness-based strategies IB and RB, especially RB, will be more harmful in WS and NW. In addition, the difference between RB and IB is caused by the redistribution of edge betweenness, which is easy to understand, since as the removals proceed, the less shortcuts have to bear the same amount of shortest paths.

The degree-based attacks, ID and RD, tend to preferentially remove edges that connect two high degree nodes (Eq (10)), which will weaken the degree of hub nodes and thus increase the number of medium degree nodes, resulting in a more centralized degree distribution as confirmed in Fig 11. This implies that the more homogeneous network may contribute to the results. To confirm this observation, we define the network heterogeneity as the standard deviation of degree distribution, i.e., $H = [\Sigma(k_i - \langle k\rangle)^2/N]^{1/2}$. We then calculate the average degree $\langle k\rangle$, the degree heterogeneity H and the average betweenness centrality $\langle B\rangle$ under different f_e subject to ID attack and IB attack (as a comparison), the results are shown in Table 2. As $\langle k\rangle$ behaves identically for all the edge attacks, it does not require much elaboration. From Table 2, we can see that H decays with f_e whereas $\langle B\rangle$ increases with f_e for ID attack. Note that $\langle B\rangle$ exhibits almost the same behaviors for IB attack, which excludes the impact of $\langle B\rangle$ on the experimental results. In contrast, the declining speed of H for ID attack is much faster than

Table 2. The structural characteristics of ER random network, including average degree $\langle k \rangle$, degree heterogeneity H and average betweenness centrality $\langle B \rangle$, vary with f_e subject to ID and IB attacks.

f_e	0%	5%	10%	15%	20%
$\langle k \rangle_{ID}$	4.0	3.8	3.6	3.4	3.2
H_{ID}	2.24	1.87	1.64	1.52	1.38
$\langle B \rangle_{ID}$	0.013	0.014	0.015	0.016	0.017
$\langle k \rangle_{IB}$	4.0	3.8	3.6	3.4	3.2
H_{IB}	2.24	2.02	1.84	1.73	1.61
$\langle B \rangle_{IB}$	0.013	0.014	0.014	0.016	0.017

doi:10.1371/journal.pone.0162289.t002

that for IB attack, suggesting the homogeneous degree distribution is the key factor that contributes to such results. Liu et al. have pointed out that the degree heterogeneity affects network controllability and the homogeneous networks are easier to control [15]. As the degree-based attacks always make the network more homogeneous and thus easier to control (smaller n_D), we infer that the two strategies are not efficient to attack the network controllability, especially for scale-free networks with biased degree structure. One thing to note is that the edge-based attacks also make $\langle k \rangle$ smaller, resulting in larger n_D according to Eq (3), which compensates for the decrease of n_D, thus n_D stays almost unchanged.

The controllability of two small-world networks, WS and NW, again displays quite similar vulnerability against edge attacks as shown in Fig 12. n_D retains almost constant in the early stage of attack ($f_e \lesssim 0.05$ for WS and $f_e \lesssim 0.07$ for NW), indicating that the network controllability is almost not affected, however, after that the attacks harm the network controllability in the order RB $>$ IB $>$ ID $\approx$ RD $>$ RA. The betweenness-based strategies prove to be the most

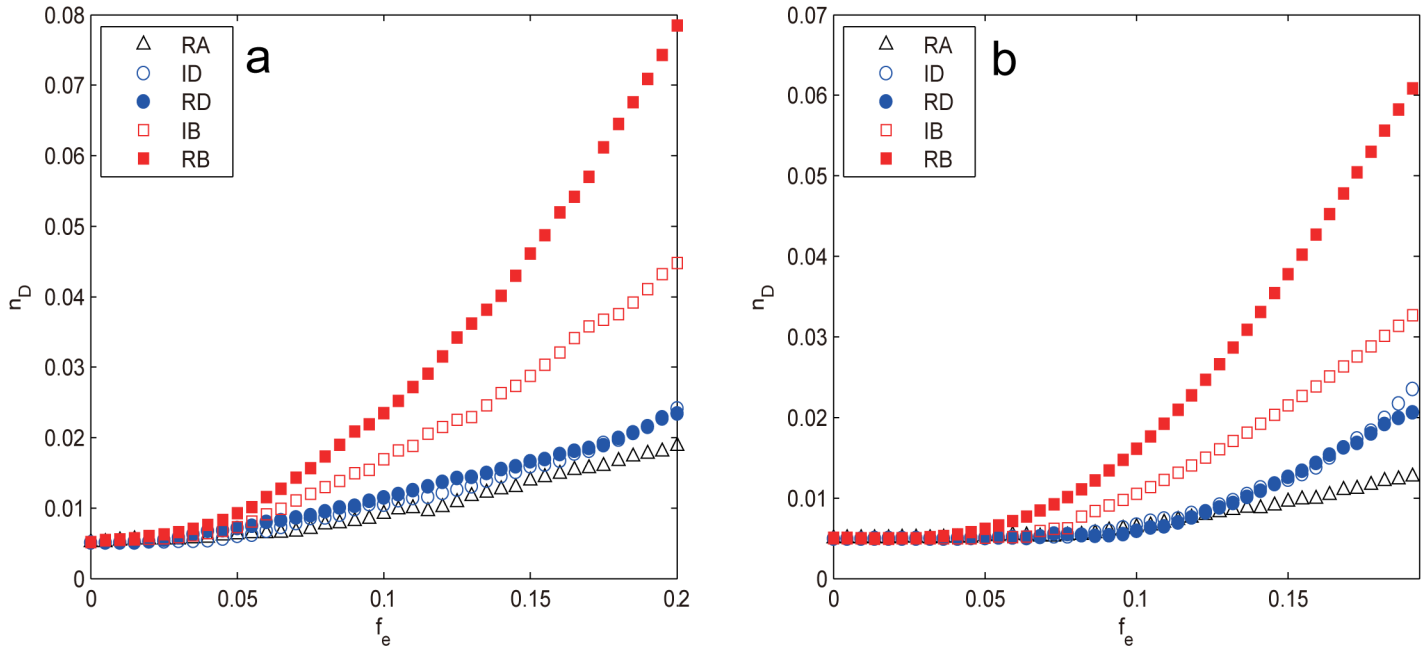

Fig 12. n_D as a function of removal fraction f_e subject to different edge attacks for (a) WS network and (b) NW network. The results are averaged over 100 independent realizations.

doi:10.1371/journal.pone.0162289.g012

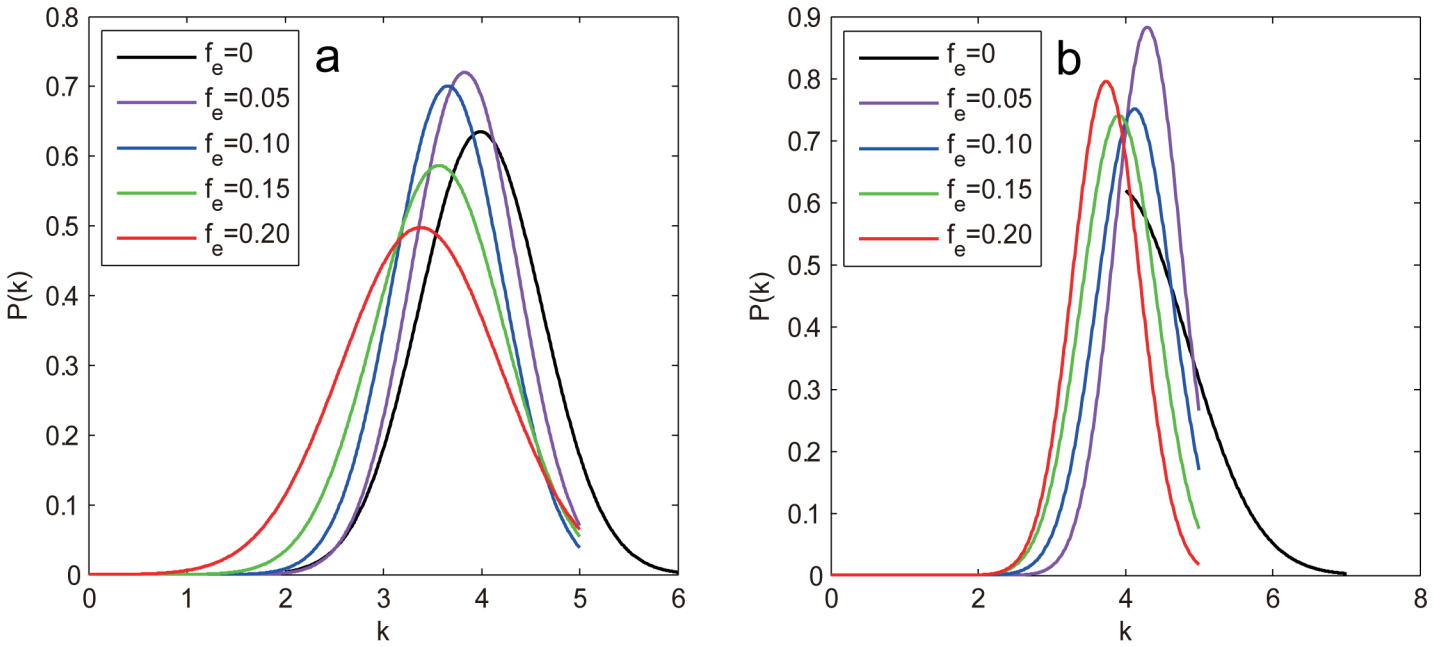

Fig 13. The degree distribution under different f_e subject to the ID attack for (a) WS network and (b) NW network.

doi:10.1371/journal.pone.0162289.g013

efficient way to damage the network controllability, which lives up to our previous conjecture. The reason for this is that the shortcuts in small-world networks play an important role in forming the so-called "small world" phenomenon, which shorten the distance between nodes that would otherwise be much farther. As such, these shortcuts usually have higher edge betweenness than ordinary edges and become the prior targets of the betweenness-based attacks. The removals of such shortcuts make the network lose most resemblance to the original topology and quickly degenerate into the one dimensional regular lattice with much larger APL, making the network harder to control. The difference between RB and IB is due to the redistribution of edge betweenness information: the less shortcuts have to bear the same number of shortest paths as the removals proceed. It should be noted that though RB and IB attack the same set of shortcuts, RB is still more efficient than IB, suggesting that the distance-based measures like edge betweenness indeed affect network's controllability.

In contrast, the two edge degree-based attacks are much less harmful than the betweenness-based ones and just comparable to random attack as seen in Fig 12. This is mainly because the degree-based strategies make the network more homogeneous and thus easier to control as has been analyzed previously. However, the real situation is slightly different. In Fig 13, we plot the degree distribution under different f_e subject to IDA for both networks, it can be seen that the peak of the curve continues to shift to left as the removal goes on, revealing that the network average degree is decreasing, while the network heterogeneity seems to be increasing judging from the more dispersed degree distribution. The more accurate numerical results in Table 3 show that the network heterogeneity, in fact, first decreases for small f_e ($f_e \lesssim 0.05$ for WS and $f_e \lesssim 0.07$ for NW) then increases until the end. This explains the behavior of n_D: for small f_e it stays constant due to the opposite effects of $\langle k \rangle$ and H, after that it starts to increase quickly as both factors make the network more difficult to control.

Table 3. The average degree $\langle k \rangle$ and degree heterogeneity H under different f_e subject to IDA for WS and NW network.

f	0%	5%	10%	15%	20%
$\langle k \rangle$(WS)	4.00	3.80	3.60	3.40	3.20
H(WS)	0.640	0.600	0.671	0.762	0.819
$\langle k \rangle$(NW)	4.40	4.20	4.00	3.80	3.60
H(NW)	0.662	0.492	0.671	0.755	0.809

doi:10.1371/journal.pone.0162289.t003

Compared with the node attacks, the BA scale-free network for edge attacks exhibits strikingly different behaviors as shown in Fig 14. The random attack proves to be the most harmful strategy whereas the other attacks are far less efficient with broad range coincidence: n_D for ID, RD and IB stays constant throughout the removal; n_D for RB also changes little for $f_e \lesssim 0.1$, thereafter it starts to increase, but the growth is quite insignificant ($\Delta n_D \approx 0.012$) compared with that of RA. Since all the strategies affect the network degree equally, the difference must result from other factors. In Fig 15, we explore the variations of related structural characteristics of the BA scale-free network as the removal procedure proceeds. It can be seen that the network heterogeneity H under the deliberate attacks decays much faster than that of RA, which is easy to understand since both the degree-based and betweenness-based strategies preferentially remove edges that connect hub nodes, which weakens the high degree nodes and makes the network become more homogeneous and thus easier to control (smaller n_D), resulting in the inferior performance of the deliberate attacks compared with that of RA. The difference between RB and the other deliberate attacks can be attributed to the distance-based measures as shown in Fig 15b and 15c, where both the APL and the average betweenness centrality $\langle B \rangle$

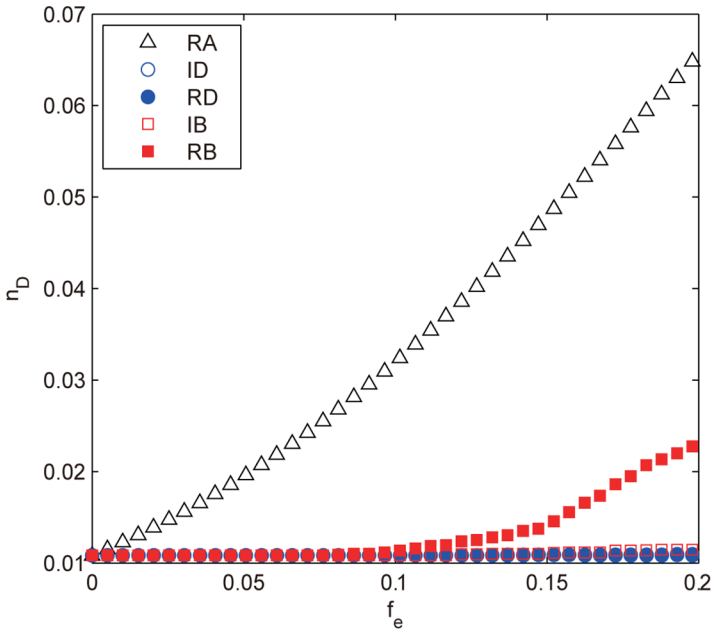

Fig 14. n_D as a function of f_e under different edge attacks for the BA scale-free network. The results are averaged over 100 independent realizations.

doi:10.1371/journal.pone.0162289.g014

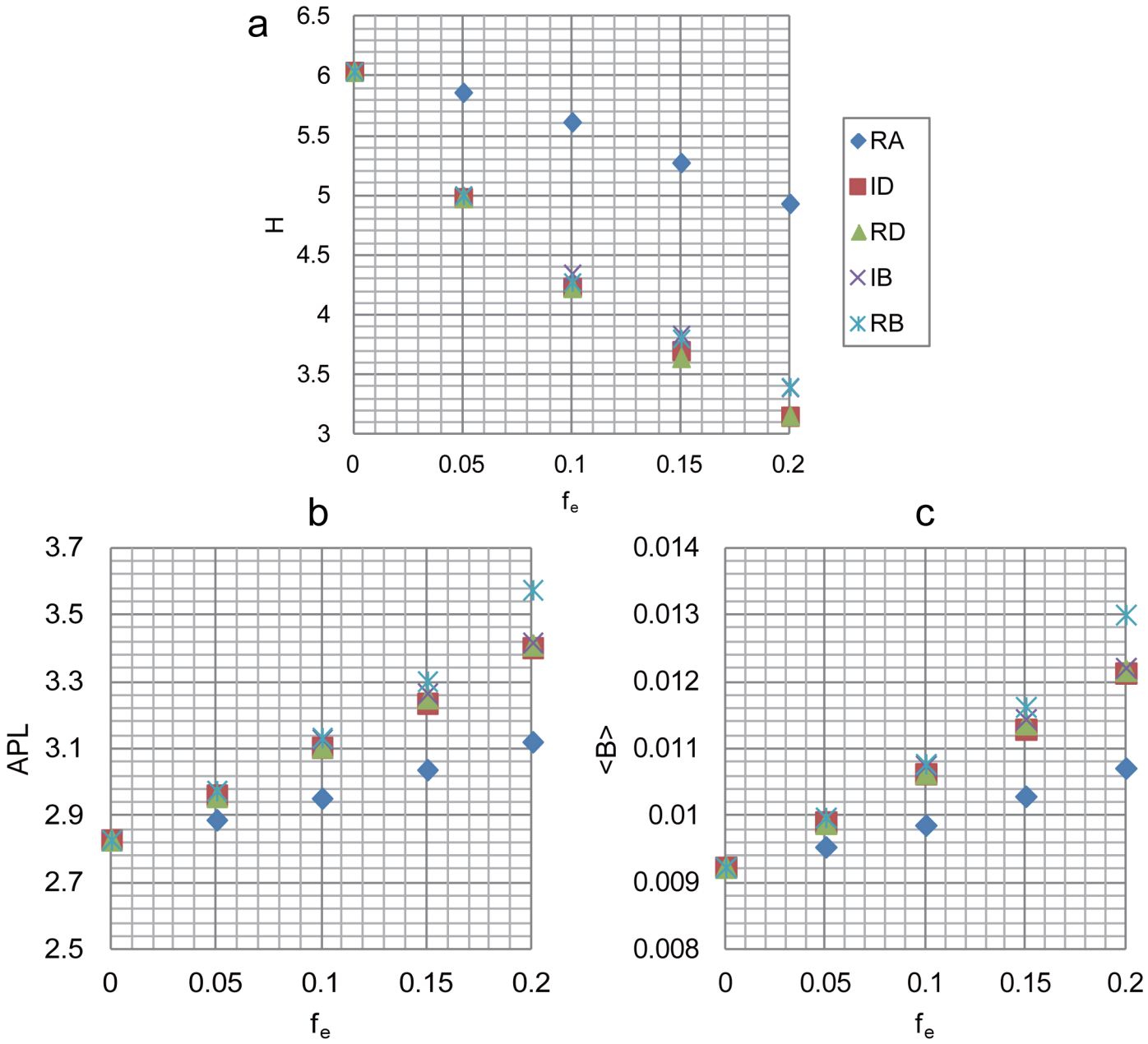

Fig 15. The structural characteristics of BA network vary with the edge removal fraction f_e. The characteristics include (**a**) the degree heterogeneity H, (**b**) APL and (**c**) the average betweenness centrality $\langle B \rangle$.

doi:10.1371/journal.pone.0162289.g015

for RB prevail other attacks with increasing advantages for $f_e \gtrsim 0.1$, which also explains the sudden increase of its n_D at $f_e \approx 0.1$.

The two real-world networks, USAir97 and Erdos971, display again quite similar control vulnerability under edge attacks as shown in Fig 16. The degree-based strategies still have little effects on n_D whereas RA continues to harm the controllability steadily as usual. The notable difference is the behaviors of two betweenness-based strategies. For USAir97, both IB and RB prevail RA with more and more advantages; while for Erdos971, the order is reversed—RA

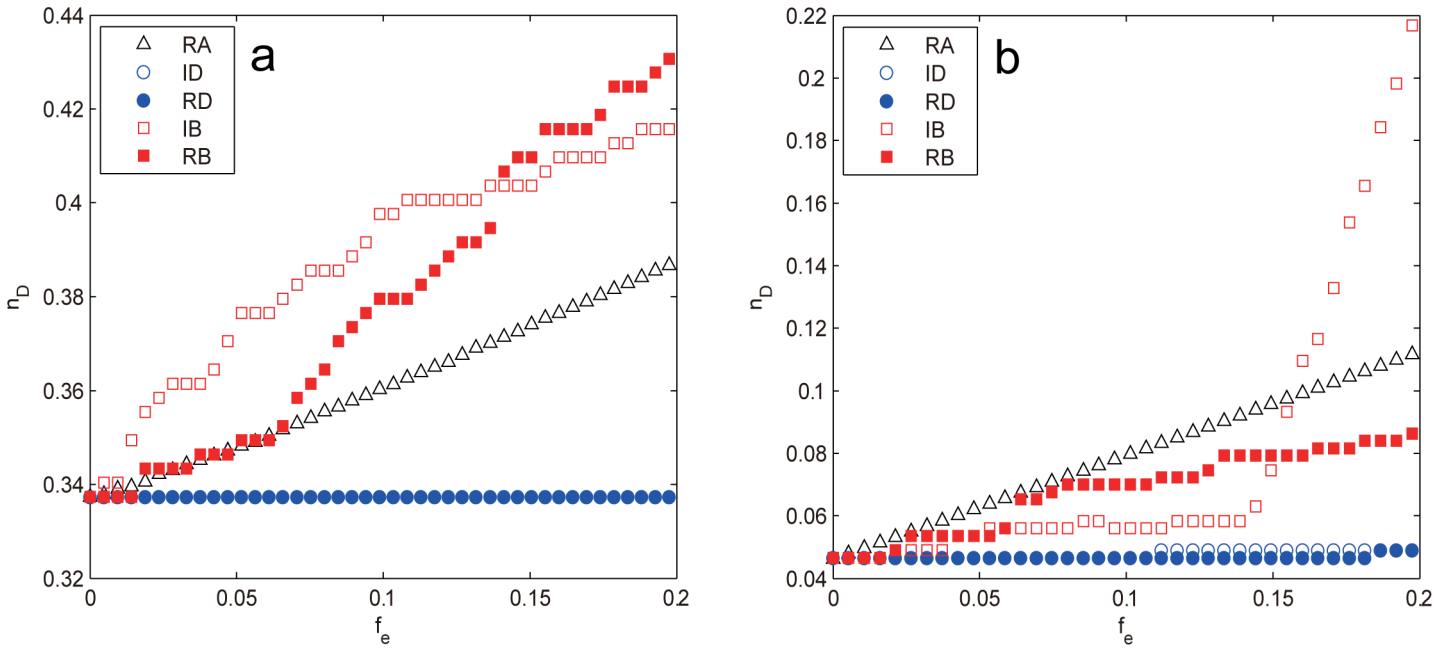

Fig 16. n_D as a function of f_e subject to different edge attacks for (a) USAir97 network and (b) Erdos971 network. The results are averaged over 100 independent realizations.

doi:10.1371/journal.pone.0162289.g016

remains prevailing over IB and RB for $f_e \lesssim 0.15$, thereafter IB starts to damage the controllability drastically and becomes the most harmful strategy. Once again these behaviors can be mainly attributed to the distanced-based measures like betweenness centrality. Nevertheless, the behaviors cannot match any of the model networks, implying that there may be other factors contributing to the network controllability.

More Results

Despite that the canonical model networks can reveal general concepts of control robustness, our ultimate goal is to understand the control robustness of real-world networks. Therefore, we further investigate the attack vulnerability of network controllability for 25 real networks subject to different node and edge attacks using the same strategies, the results are shown in Tables 4 and 5.

From Table 4, we can see that the regulatory networks such as the transcriptional regulatory network of *Saccharomyces cerevisiae* (TRN-Yeast-1 [48], TRN-Yeast-2 [49]) and *Echerichia coli* (TRN-EC-2 [49]) and the ownership network of US telecommunications and media corporations (Ownership-USCrop [50]) exhibit strikingly control robust under attacks: n_D decreases instead of increases throughout the removal for all the five strategies, meaning that we need ever less and less driver nodes to control the whole network and the network controllability is, in fact, increasing. This phenomenon has not been observed before our work and clearly shows that the control robustness of regulatory networks are more than good enough. The other class of networks with great control robustness is the case of the intra-organizational networks [51, 52], in which n_D stays constant for all five attacks, indicating that the attacking cannot affect the network controllability at all. One exception is the university network [52], which can only resist some attacks (ID, RD, RB) for small $f_e \leq 10\%$. The control robustness of other networks

Table 4. The node attack vulnerability of the real networks analyzed in this paper.

Type	Name	N	L	n_D	n_D under node attack				
					RA	IDA	RDA	IBA	RBA
Regulatory	TRN-Yeast-1	4,441	12,873	0.965	0.869 [0.771]	0.897 [0.798]	0.900 [0.800]	0.894 [0.794]	0.886 [0.787]
	TRN-Yeast-2	688	1,079	0.821	0.738 [0.660]	0.815 [0.765]	0.823 [0.801]	0.773 [0.692]	0.769 [0.689]
	TRN-EC-2	418	519	0.751	0.677 [0.612]	0.739 [0.691]	0.754 [0.754]	0.722 [0.639]	0.718 [0.636]
	Ownership-USCrop	7,253	6,726	0.820	0.750 [0.680]	0.812 [0.731]	0.832 [0.800]	0.788 [0.688]	0.775 [0.675]
Trust	College student	32	96	0.188	0.188 [0.250]	0.188 [0.219]	0.219 [0.250]	0.219 [0.250]	0.219 [0.281]
	Prison inmate	67	182	0.134	0.164 [0.164]	0.149 [0.179]	0.164 [0.194]	0.164 [0.179]	0.164 [0.224]
	WikiVote	7,115	103,689	0.666	0.600 [0.534]	0.668 [0.663]	0.664 [0.662]	0.677 [0.672]	0.678 [0.658]
Food web	Ythan	135	601	0.511	0.467 [0.430]	0.563 [0.570]	0.556 [0.570]	0.548 [0.563]	0.563 [0.570]
	Little Rock	183	2,494	0.541	0.497 [0.454]	0.541 [0.541]	0.541 [0.541]	0.541 [0.541]	0.546 [0.546]
	Grassland	88	137	0.523	0.477 [0.409]	0.534 [0.580]	0.534 [0.580]	0.545 [0.545]	0.557 [0.602]
	Seagrass	49	226	0.265	0.265 [0.245]	0.286 [0.265]	0.265 [0.265]	0.286 [0.265]	0.265 [0.265]
Metabolic	*E. coli*	2,275	5,763	0.382	0.367 [0.346]	0.457 [0.424]	0.480 [0.581]	0.378 [0.350]	0.481 [0.578]
	S. cerevisiae	1,511	3,833	0.329	0.325 [0.312]	0.406 [0.365]	0.428 [0.527]	0.326 [0.311]	0.428 [0.525]
	C. elegans	1,809	1,303	0.657	0.603 [0.545]	0.641 [0.612]	0.655 [0.654]	0.655 [0.655]	0.655 [0.655]
Electronic	s838	512	819	0.232	0.246 [0.244]	0.313 [0.305]	0.314 [0.354]	0.254 [0.275]	0.311 [0.367]
circuits	s420	252	399	0.234	0.262 [0.274]	0.310 [0.313]	0.310 [0.353]	0.262 [0.274]	0.310 [0.369]
	s208	122	189	0.238	0.230 [0.262]	0.303 [0.328]	0.303 [0.352]	0.279 [0.287]	0.311 [0.361]
Neuronal	*C. elegans*	297	2,345	0.165	0.158 [0.152]	0.192 [0.219]	0.185 [0.222]	0.178 [0.205]	0.195 [0.242]
WWW	Political blogs	1,224	19,025	0.356	0.334 [0.305]	0.408 [0.456]	0.408 [0.454]	0.417 [0.470]	0.425 [0.475]
Internet	p2p-1	10,876	39,994	0.552	0.503 [0.454]	0.571 [0.571]	0.576 [0.581]	0.573 [0.576]	0.575 [0.590]
	p2p-2	8,846	31,839	0.578	0.525 [0.473]	0.585 [0.579]	0.590 [0.595]	0.589 [0.591]	0.595 [0.605]
	p2p-3	8,717	31,525	0.577	0.527 [0.474]	0.585 [0.583]	0.588 [0.592]	0.591 [0.593]	0.596 [0.604]
Organizational	Consulting	46	879	0.043	0.043 [0.043]	0.043 [0.043]	0.043 [0.043]	0.043 [0.043]	0.043 [0.043]
	Manufacturing	77	2,228	0.013	0.013 [0.013]	0.013 [0.013]	0.013 [0.013]	0.013 [0.013]	0.013 [0.013]
	University	81	817	0.012	0.025 [0.025]	0.012 [0.025]	0.012 [0.025]	0.025 [0.037]	0.012 [0.025]

For each network, we show its type, name, number of nodes (N), number of edges (L), the initial density of driver nodes (n_D) and n_D after 10% [20%] nodes are removed. All the networks are directed and n_D is calculated with the structural controllability framework [15]. For networks with $N \geq 3000$, nodal betweenness centrality is estimated with approximate algorithm [47] to speed up the computation. For data sources and references, see S1 Table.

doi:10.1371/journal.pone.0162289.t004

can be, more or less, grasped from the model networks. The betweenness-based attacks and the degree-based attacks are almost equal harmful for the food web networks [53–55] and the metabolic networks [56]. However, for most other networks, the former are usually more harmful than the latter and so are the strategies based on recalculated information than those based on initial information. The last interesting observation is that most of the real-world networks seem to have a decreasing or constant n_D for RA, indicating that the real-world networks tend to be control robust against random node failures.

In Table 5, we display the attack vulnerability of real-world networks subject to edge attacks. It can be seen that the edge degree-based attacks (both ID and RD) still cannot damage the network controllability at all as observed in the model networks, which, as has been explained, is because of the fact that such strategies make networks become more homogenous and thus easier to control. Our conclusion is that the edge degree defined as Eq (10) is not a good quantity to measure the importance of an edge in terms of network controllability. Another significant difference is the behaviors of the random edge failures. Here only the intra-organizational

Table 5. The edge attack vulnerability of real networks analyzed in this paper.

Type	Name	N	L	n_D	n_D under edge attack				
					RA	IDA	RDA	IBA	RBA
Regulatory	TRN-Yeast-1	4,441	12,873	0.965	0.965 [0.965]	0.965 [0.965]	0.965 [0.965]	0.965 [0.965]	0.965 [0.965]
	TRN-Yeast-2	688	1,079	0.821	0.826 [0.830]	0.821 [0.821]	0.821 [0.823]	0.830 [0.846]	0.830 [0.833]
	TRN-EC-2	418	519	0.751	0.758 [0.778]	0.751 [0.751]	0.751 [0.751]	0.768 [0.778]	0.773 [0.785]
	Ownership-USCrop	7,253	6,726	0.820	0.833 [0.843]	0.820 [0.821]	0.820 [0.820]	0.832 [0.849]	0.842 [0.852]
Trust	College student	32	96	0.188	0.188 [0.188]	0.188 [0.188]	0.188 [0.188]	0.219 [0.313]	0.250 [0.281]
	Prison inmate	67	182	0.134	0.164 [0.209]	0.134 [0.149]	0.134 [0.134]	0.179 [0.224]	0.164 [0.224]
	WikiVote	7,115	103,689	0.666	0.667 [0.669]	0.666 [0.666]	0.666 [0.666]	0.674 [0.678]	0.666 [0.676]
Food web	Ythan	135	601	0.511	0.533 [0.533]	0.511 [0.511]	0.511 [0.511]	0.541 [0.548]	0.533 [0.600]
	Little Rock	183	2,494	0.541	0.541 [0.541]	0.541 [0.541]	0.541 [0.541]	0.557 [0.579]	0.568 [0.601]
	Grassland	88	137	0.523	0.545 [0.591]	0.523 [0.523]	0.523 [0.523]	0.557 [0.602]	0.580 [0.659]
	Seagrass	49	226	0.265	0.265 [0.286]	0.265 [0.286]	0.265 [0.265]	0.306 [0.347]	0.327 [0.327]
Metabolic	*E. coli*	2,275	5,763	0.382	0.406 [0.436]	0.383 [0.385]	0.382 [0.383]	0.390 [0.410]	0.394 [0.410]
	S. cerevisiae	1,511	3,833	0.329	0.355 [0.392]	0.330 [0.333]	0.329 [0.330]	0.343 [0.351]	0.343 [0.361]
	C. elegans	1,809	1,303	0.657	0.666 [0.689]	0.657 [0.657]	0.657 [0.657]	0.669 [0.684]	0.669 [0.685]
Electronic	s838	512	819	0.232	0.285 [0.340]	0.246 [0.264]	0.232 [0.246]	0.299 [0.334]	0.346 [0.422]
circuits	s420	252	399	0.234	0.294 [0.333]	0.246 [0.274]	0.234 [0.242]	0.302 [0.341]	0.353 [0.425]
	s208	122	189	0.238	0.262 [0.320]	0.238 [0.262]	0.238 [0.238]	0.295 [0.361]	0.352 [0.434]
Neuronal	*C. elegans*	297	2,345	0.165	0.172 [0.178]	0.165 [0.165]	0.165 [0.165]	0.256 [0.269]	0.205 [0.273]
WWW	Political blogs	1,224	19,025	0.356	0.373 [0.389]	0.356 [0.356]	0.356 [0.356]	0.431 [0.468]	0.404 [0.494]
Internet	p2p-1	10,876	39,994	0.552	0.558 [0.566]	0.556 [0.562]	0.552 [0.553]	0.558 [0.564]	0.565 [0.573]
	p2p-2	8,846	31,839	0.578	0.585 [0.591]	0.582 [0.585]	0.578 [0.578]	0.583 [0.588]	0.588 [0.597]
	p2p-3	8,717	31,525	0.577	0.582 [0.589]	0.583 [0.586]	0.578 [0.578]	0.583 [0.587]	0.589 [0.597]
Organizational	Consulting	46	879	0.043	0.043 [0.043]	0.043 [0.043]	0.043 [0.043]	0.065 [0.087]	0.130 [0.174]
	Manufacturing	77	2,228	0.013	0.013 [0.013]	0.013 [0.013]	0.013 [0.013]	0.013 [0.026]	0.052 [0.065]
	University	81	817	0.012	0.012 [0.012]	0.012 [0.012]	0.012 [0.012]	0.062 [0.074]	0.074 [0.099]

For each network, we show its type, name, number of nodes (N), number of edges (L), the initial density of driver nodes (n_D) and n_D after 10% [20%] edges are removed. All the networks are directed and n_D is calculated with the structural controllability framework [15]. For networks with $N \geq 3000$, nodal betweenness centrality is estimated with approximate algorithm [47] to speed up the computation. For data sources and references, see S1 Table.

doi:10.1371/journal.pone.0162289.t005

networks [51, 52] show good control robustness against random edge failures while other networks do not. The rest cases are very similar. In general, RB is the most harmful strategy, followed by IB, and RA is the least one. Two exceptions are the trust networks [57–59] in which IB are more efficient than RB and the metabolic networks [56] in which RA suppresses both IB and RB with slight advantage.

In summary, the real world networks display rather diverse control robust behaviors, which may suggest that one should really choose the attack strategies carefully before harming the controllability of real networks as it may affect the results to a very large extent.

Summary and Conclusions

In this paper, we have systematically investigate the attack vulnerability of network controllability for the canonical model networks as well as the real-world networks subject to five different strategies on the basis of nodes and edges. The strategies are chosen based on degree and betweenness centralities evaluated with the initial information as well as the recalculated

information, among which random failure is as a comparison. We found that for node attacks, the ER random network is more control vulnerable to the degree-based attacks (RD and ID) with $\Delta n_D \approx 0.12$ while the small-world networks (WS and NW) are more vulnerable to the betweenness-based attack (RB) with $\Delta n_D \approx 0.07$. The BA scale-free model network with an exponential degree distribution, which is one of the most important characteristics in real-world networks [60–63], turns out to be the most vulnerable network ($\Delta n_D \approx 0.19$) due to the existence of hub nodes and the high correlation between node degree and betweenness. The similar vulnerability is also observed in the two real-world networks, USAir97 and Erdos971. For edge attacks, the strategies are not as that efficient as the node-based ones. For example, both RD and ID can hardly harm the controllability at all for all the tested networks, and even the most harmful strategy RB only damages the ER random network with $\Delta n_D \approx 0.033$ and the small-world networks with $\Delta n_D \approx 0.065$. The notable difference is the BA scale-free network, which, out of our expectation, exhibits great control robustness to all the intentional attacks with almost constant n_D. The two real-world networks, USAir97 and Erdos971, display complicated behaviors different from any of the model networks with $\Delta n_D \approx 0.09$ for the former and $\Delta n_D \approx 0.18$ for the latter, which suggests that there may be other factors contributing to the controllability.

We also investigate the control robustness of 25 real-world networks and find that the control vulnerability of real-world networks is much different from that of model networks. For example, most of the real-world networks exhibit good control robustness against random node failures, which are not observed in the model networks. The regulatory networks are even excessively robust to all the node attacks with decreasing n_Ds ($\Delta n_D \leq 0$), followed by the intra-organizational networks with constant n_Ds ($\Delta n_D = 0$). For other networks subject to node attacks, the betweenness-based strategies are more harmful than the degree-based ones, and so are the strategies based recalculated information than their counterparts. In contrast, for edge attacks, only the intra-organizational networks are robust to random edge failures whereas other networks are harmed by the attacks in the order RB $>$ IB $>$ RA. In addition, we find that edge degree is not a good quantity to measure the importance of an edge in terms of network controllability.

Our results raise several questions to answer to help us better understand the attack vulnerability of network controllability. For example, we have shown that besides the degree distribution, there are other factors affecting the network controllability, what are the factors and how do they affect the controllability? Besides, the real-world networks usually have certain defence against attacks and failures, so how about the attack vulnerability of network controllability with defence?

Supporting Information

S1 Table. Real networks analyzed in this paper.
(PDF)

Acknowledgments

The authors would like to thank Prof. Tamas Vicsek and Dr. Tamas Nepusz for providing some of the network data and an anonymous referee for insightful suggestions.

Author Contributions

Conceived and designed the experiments: ZML XFL.

Performed the experiments: XFL.

Analyzed the data: ZML XFL.

Contributed reagents/materials/analysis tools: ZML XFL.

Wrote the paper: ZML XFL.

References

1. Watts DJ, Strogatz SH. Collective dynamics of 'small-world'networks. nature. 1998; 393(6684):440–442. doi: 10.1038/30918 PMID: 9623998
2. Barabási AL, Albert R. Emergence of scaling in random networks. science. 1999; 286(5439):509–512. doi: 10.1126/science.286.5439.509 PMID: 10521342
3. Boccaletti S, Latora V, Moreno Y, Chavez M, Hwang DU. Complex networks: Structure and dynamics. Physics reports. 2006; 424(4):175–308. doi: 10.1016/j.physrep.2005.10.009
4. May RM, Lloyd AL. Infection dynamics on scale-free networks. Physical Review E. 2001; 64 (6):066112. doi: 10.1103/PhysRevE.64.066112
5. Moreno Y, Nekovee M, Pacheco AF. Dynamics of rumor spreading in complex networks. Physical Review E. 2004; 69(6):066130. doi: 10.1103/PhysRevE.69.066130
6. Arenas A, Díaz-Guilera A, Kurths J, Moreno Y, Zhou C. Synchronization in complex networks. Physics Reports. 2008; 469(3):93–153. doi: 10.1016/j.physrep.2008.09.002
7. Pastor-Satorras R, Vespignani A. Epidemic dynamics and endemic states in complex networks. Physical Review E. 2001; 63(6):066117. doi: 10.1103/PhysRevE.63.066117
8. Boguná M, Pastor-Satorras R. Epidemic spreading in correlated complex networks. Physical Review E. 2002; 66(4):047104. doi: 10.1103/PhysRevE.66.047104
9. Motter AE, Zhou C, Kurths J. Network synchronization, diffusion, and the paradox of heterogeneity. Physical Review E. 2005; 71(1):016116. doi: 10.1103/PhysRevE.71.016116
10. Motter AE, Zhou C, Kurths J. Enhancing complex-network synchronization. EPL (Europhysics Letters). 2005; 69(3):334. doi: 10.1209/epl/i2004-10365-4
11. Noh JD, Rieger H. Random walks on complex networks. Physical review letters. 2004; 92(11):118701. doi: 10.1103/PhysRevLett.92.118701 PMID: 15089179
12. Rosvall M, Bergstrom CT. Maps of random walks on complex networks reveal community structure. Proceedings of the National Academy of Sciences. 2008; 105(4):1118–1123. doi: 10.1073/pnas.0706851105
13. Sznajd-Weron K, Sznajd J. Opinion evolution in closed community. International Journal of Modern Physics C. 2000; 11(06):1157–1165. doi: 10.1142/S0129183100000936
14. Gonzalez M, Sousa A, Herrmann H. Opinion formation on a deterministic pseudo-fractal network. International journal of modern physics C. 2004; 15(01):45–57. doi: 10.1142/S0129183104005577
15. Liu YY, Slotine JJ, Barabási AL. Controllability of complex networks. Nature. 2011; 473(7346):167–173. doi: 10.1038/nature10011 PMID: 21562557
16. Kalman RE. Mathematical description of linear dynamical systems. Journal of the Society for Industrial & Applied Mathematics, Series A: Control. 1963; 1(2):152–192. doi: 10.1137/0301010
17. Luenberger D. Introduction to dynamic systems: theory, models, and applications. 1979;.
18. Slotine JJE, Li W, et al. Applied nonlinear control. vol. 199. Prentice-hall Englewood Cliffs, NJ; 1991.
19. Bechhoefer J. Feedback for physicists: A tutorial essay on control. Reviews of Modern Physics. 2005; 77(3):783. doi: 10.1103/RevModPhys.77.783
20. Lombardi A, Hörnquist M. Controllability analysis of networks. Physical Review E. 2007; 75(5):056110. doi: 10.1103/PhysRevE.75.056110
21. Porfiri M, Fiorilli F. Experiments on node-to-node pinning control of Chua's circuits. Physica D: Nonlinear Phenomena. 2010; 239(8):454–464. doi: 10.1016/j.physd.2010.01.012
22. Shmulevich I, Dougherty ER. Probabilistic Boolean networks: the modeling and control of gene regulatory networks. siam; 2010.
23. Liu B, Chu T, Wang L, Xie G. Controllability of a leader–follower dynamic network with switching topology. Automatic Control, IEEE Transactions on. 2008; 53(4):1009–1013. doi: 10.1109/TAC.2008.919548
24. Rahmani A, Ji M, Mesbahi M, Egerstedt M. Controllability of multi-agent systems from a graph-theoretic perspective. SIAM Journal on Control and Optimization. 2009; 48(1):162–186. doi: 10.1137/060674909

25. Lin CT. Structural controllability. Automatic Control, IEEE Transactions on. 1974; 19(3):201–208. doi: 10.1109/TAC.1974.1100557
26. Gabow HN. An efficient implementation of Edmonds' algorithm for maximum matching on graphs. Journal of the ACM (JACM). 1976; 23(2):221–234. doi: 10.1145/321941.321942
27. Nepusz T, Vicsek T. Controlling edge dynamics in complex networks. Nature Physics. 2012; 8(7):568–573. doi: 10.1038/nphys2327
28. Pósfai M, Liu YY, Slotine JJ, Barabási AL. Effect of correlations on network controllability. Scientific reports. 2013; 3. doi: 10.1038/srep01067 PMID: 23323210
29. Yan G, Ren J, Lai YC, Lai CH, Li B. Controlling complex networks: How much energy is needed? Physical review letters. 2012; 108(21):218703. doi: 10.1103/PhysRevLett.108.218703 PMID: 23003312
30. Liu Y, Slotine J, Barabási A. Control centrality and hierarchical structure in complex networks. PloS one. 2011; 7(9):e44459–e44459. doi: 10.1371/journal.pone.0044459
31. Wang WX, Ni X, Lai YC, Grebogi C. Optimizing controllability of complex networks by minimum structural perturbations. Physical Review E. 2012; 85(2):026115. doi: 10.1103/PhysRevE.85.026115
32. Yuan Z, Zhao C, Di Z, Wang WX, Lai YC. Exact controllability of complex networks. Nature communications. 2013; 4. doi: 10.1038/ncomms3447
33. Barabási AL, Albert R, Jeong H. Mean-field theory for scale-free random networks. Physica A: Statistical Mechanics and its Applications. 1999; 272(1):173–187.
34. Solé RV, Rosas-Casals M, Corominas-Murtra B, Valverde S. Robustness of the European power grids under intentional attack. Physical Review E. 2008; 77(2):026102. doi: 10.1103/PhysRevE.77.026102
35. Holme P, Kim BJ, Yoon CN, Han SK. Attack vulnerability of complex networks. Physical Review E. 2002; 65(5):056109. doi: 10.1103/PhysRevE.65.056109
36. Wang XF, Chen G. Synchronization in scale-free dynamical networks: robustness and fragility. Circuits and Systems I: Fundamental Theory and Applications, IEEE Transactions on. 2002; 49(1):54–62. doi: 10.1109/81.974874
37. Pu CL, Pei WJ, Michaelson A. Robustness analysis of network controllability. Physica A: Statistical Mechanics and its Applications. 2012; 391(18):4420–4425. doi: 10.1016/j.physa.2012.04.019
38. Nie S, Wang X, Zhang H, Li Q, Wang B. Robustness of controllability for networks based on edge-attack. PloS one. 2014; 9(2):e89066. doi: 10.1371/journal.pone.0089066 PMID: 24586507
39. Motter AE, Lai YC. Cascade-based attacks on complex networks. Physical Review E. 2002; 66 (6):065102.
40. Tu Y. How robust is the Internet? Nature. 2000; 406(6794):353–354. doi: 10.1038/35019222 PMID: 10935616
41. Rugh WJ. Linear system theory. vol. 2. prentice hall Upper Saddle River, NJ; 1996.
42. Erd6s P, Rényi A. On the evolution of random graphs. Publ Math Inst Hungar Acad Sci. 1960; 5:17–61.
43. Hautus M. Controllability and observability conditions of linear autonomous systems. PROCEEDINGS OF THE KONINKLIJKE NEDERLANDSE AKADEMIE VAN WETENSCHAPPEN SERIES A-MATHEMATICAL SCIENCES. 1969; 72(5):443.
44. Newman ME, Watts DJ. Renormalization group analysis of the small-world network model. Physics Letters A. 1999; 263(4):341–346. doi: 10.1016/S0375-9601(99)00757-4
45. Batagelj V, Mrvar A. Pajek datasets; 2006.
46. Banerjee SJ, Roy S. Key to network controllability. arXiv preprint arXiv:12093737. 2012;.
47. Brandes U, Pich C. Centrality estimation in large networks. International Journal of Bifurcation and Chaos. 2007; 17(07):2303–2318. doi: 10.1142/S0218127407018403
48. Balaji S, Babu MM, Iyer LM, Luscombe NM, Aravind L. Comprehensive analysis of combinatorial regulation using the transcriptional regulatory network of yeast. Journal of molecular biology. 2006; 360 (1):213–227. doi: 10.1016/j.jmb.2006.04.029 PMID: 16762362
49. Milo R, Shen-Orr S, Itzkovitz S, Kashtan N, Chklovskii D, Alon U. Network motifs: simple building blocks of complex networks. Science. 2002; 298(5594):824–827. doi: 10.1126/science.298.5594.824 PMID: 12399590
50. Norlen K, Lucas G, Gebbie M, Chuang J. EVA: Extraction, visualization and analysis of the telecommunications and media ownership network. In: Proceedings of International Telecommunications Society 14th Biennial Conference (ITS2002), Seoul Korea. Citeseer; 2002.
51. Cross R, Parker A, Christensen CM, Anthony SD, Roth EA. The hidden power of social networks. Audio-Tech Business Book Summaries, Incorporated; 2004.

52. Nepusz T, Petróczi A, Négyessy L, Bazsó F. Fuzzy communities and the concept of bridgeness in complex networks. Physical Review E. 2008; 77(1):016107. doi: 10.1103/PhysRevE.77.016107

53. Dunne JA, Williams RJ, Martinez ND. Food-web structure and network theory: the role of connectance and size. Proceedings of the National Academy of Sciences. 2002; 99(20):12917–12922. doi: 10.1073/pnas.192407699

54. Martinez ND. Artifacts or attributes? Effects of resolution on the Little Rock Lake food web. Ecological Monographs. 1991; p. 367–392. doi: 10.2307/2937047

55. Christian RR, Luczkovich JJ. Organizing and understanding a winter's seagrass foodweb network through effective trophic levels. Ecological Modelling. 1999; 117(1):99–124. doi: 10.1016/S0304-3800(99)00022-8

56. Jeong H, Tombor B, Albert R, Oltvai ZN, Barabási AL. The large-scale organization of metabolic networks. Nature. 2000; 407(6804):651–654. doi: 10.1038/35036627 PMID: 11034217

57. Milo R, Itzkovitz S, Kashtan N, Levitt R, Shen-Orr S, Ayzenshtat I, et al. Superfamilies of evolved and designed networks. Science. 2004; 303(5663):1538–1542. doi: 10.1126/science.1089167 PMID: 15001784

58. Van Duijn MA, Zeggelink EP, Huisman M, Stokman FN, Wasseur FW. Evolution of sociology freshmen into a friendship network. Journal of Mathematical Sociology. 2003; 27(2-3):153–191. doi: 10.1080/00222500305889

59. Leskovec J, Huttenlocher D, Kleinberg J. Signed networks in social media. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. ACM; 2010. p. 1361–1370.

60. Dorogovtsev SN, Mendes JF. Evolution of networks: From biological nets to the Internet and WWW. Oxford University Press; 2013.

61. Newman M, Barabasi AL, Watts DJ. The structure and dynamics of networks. Princeton University Press; 2006.

62. Caldarelli G. Scale-free networks: complex webs in nature and technology. OUP Catalogue. 2007;.

63. Albert R, Barabási AL. Statistical mechanics of complex networks. Reviews of modern physics. 2002; 74(1):47. doi: 10.1103/RevModPhys.74.47