

SCIENTIFIC REPORTS



OPEN

Localized recovery of complex networks against failure

Yilun Shang

Received: 05 May 2016

Accepted: 04 July 2016

Published: 26 July 2016

Resilience of complex networks to failure has been an important issue in network research for decades, and recent studies have begun to focus on the inverse recovery of network functionality through strategically healing missing nodes or edges. However, the effect of network recovery is far from fully understood, and a general theory is still missing. Here we propose and study a general model of localized recovery, where a group of neighboring nodes are restored in an invasive way from a seed node. We develop a theoretical framework to compare the effect of random recovery (RR) and localized recovery (LR) in complex networks including Erdős-Rényi networks, random regular networks, and scale-free networks. We find detailed phase diagrams for the subnetwork of occupied nodes and the “complement network” of failed nodes under RR and LR. By identifying the two competitive forces behind LR, we present an analytical and numerical approach to guide us in choosing the appropriate recovery strategy and provide estimation on its effect by using the degree distribution of the original network as the only input. Our work therefore provides insight for quantitatively understanding recovery process and its implications in infrastructure protection in various complex systems.

Interconnection between network nodes has important consequences for stability and function of many complex networked systems ranging from Internet and social networks to electric power grids and transportation systems, where nodes are prone to random failure or malicious attack^{1–4}. Through these interconnections, failure of individual components may crucially affect the performance of the whole system. The Internet with a scale-free connectivity distribution, for example, is highly vulnerable to attacks targeted at hubs but extremely robust against random errors⁵. Due to practical importance for applications, robustness of complex networks to failures remains a topic of central interest in the network science literature^{6–12}.

While most efforts have been focused on probing the influence of network topology on robustness, failure mechanism of complex networks is very recently incorporated into the picture. It is revealed that an entire class of real-world networks, including human brains, cancer networks, polymers, financial systems, and bot-nets are capable of spontaneously recovery (or self-healing) after their collapse by restoring the failed nodes or edges^{13,14}. The work¹³ describes a phase-flipping phenomenon between “active” and “inactive” collective modes in a time-dependent Erdős-Rényi networks using a threshold rule, where each node becomes inactive (or removed) if it has less than a certain fraction of active neighbors and it then can be recovered after a fixed time period. An extension to dynamic scale-free networks is presented in ref. 15. Self-healing capability of nodes using a related threshold rule has been dealt within percolation theory¹⁴. Ref. 16 studies the targeted recovery algorithm in which the hub nodes in interdependent networks are iteratively recovered. A recovery strategy that repairs failed nodes and edges randomly in the boundary of the largest connected component is introduced in ref. 17, where the critical probability of recovery that halts the cascade is identified for interdependent networks.

Different from recovering missing nodes or edges to restore functionality, another approach to improving network robustness consists in adding new edges or smartly modifying the network topology^{18–22}. For instance, optimal recovery algorithms of adding minimum links so that every node can reach any other node in the resulting network is studied in ref. 18. However, as pointed out in refs 14,23, the possibility to create new links is normally not available in many real networks such as infrastructure networks, where physical links are fixed in advance and creating new ones incurs time and investments.

In this paper, we tackle the robustness of complex networks by considering *localized recovery* (LR) strategy, and comparing it with *random recovery* (RR) strategy. The localized recovery consists of a healing process of one node, then its neighbors, and then their neighbors, etc. (see Fig. 1). It restores only failed nodes and edges without assuming any particular healing mechanisms, e.g., the threshold rules^{13–15}, hence addressing the shortcomings associated with both approaches. Localized recovery strategy is widely employed in the real world because (a) it

Department of Mathematics, Tongji University, 200092 Shanghai, China. Correspondence and requests for materials should be addressed to Y.S. (email: shyl@tongji.edu.cn)

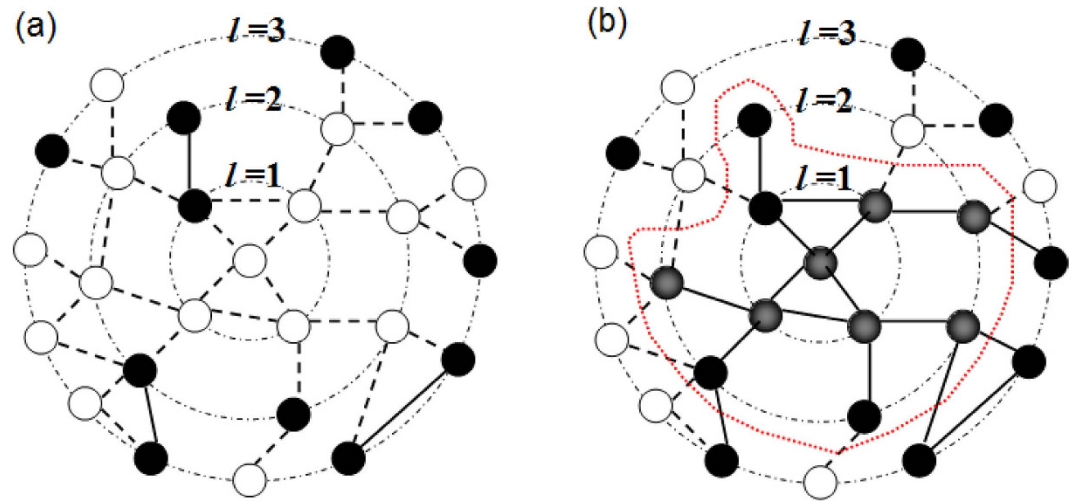


Figure 1. Schematic illustration of the localized recovery process. (a) A network suffers from random failures, where each node remains functional with probability q (black means functional nodes, white failed nodes); the root node is failed in this case. (b) A fraction r of the failed nodes are recovered starting from the root, its nearest neighbors, next nearest neighbors, and so forth (gradient black means recovered nodes); edges connecting two occupied (i.e., functional or recovered) nodes become active (solid lines) again. The recovery area is indicated by a red contour.

is usually easier and more effective to repair a node in the neighborhood than one far apart as functional nodes become connected after recovery and (b) it is sometimes the only way to repair when an external monitor/controller is not available to guide the recovery or limited resources fail to maintain simultaneous recovery in multiple unrelated areas.

The analytical framework developed here for studying localized recovery of complex networks allows us to examine some percolation properties such as the critical recovery probability r_c , marking the threshold of existence/non-existence of giant component in the network of failed nodes, and the fraction P_∞ of failed nodes in the giant component of the network of failed nodes after recovery. Let r_0 be a second critical recovery probability signifying the threshold of the network of occupied (i.e., functional or recovered) nodes, at which a giant component of occupied nodes first forms. As we will see below, r_c is distinct from the threshold r_0 , which is essentially trivial under LR. We find detailed phase diagrams under RR and LR. By identifying the two competitive forces that influence the effect of localized recovery strategy, we develop an analytical tool that can guide us in choosing the appropriate recovery strategy and provide estimation on its effect by using the degree distribution of the original network as the only input. Real networks with degree-degree correlation are also investigated. Our theoretical calculations are confirmed by extensive numerical simulations.

We mention that the localized recovery strategy hinges on the interconnections in a network, especially the topology of its shells²⁴, which is defined as the set of nodes that are at some distance from a randomly chosen root node. Shell structure has been explored recently by some researchers to collapse an otherwise connected network^{25–28}.

Results

Description of the model. The interconnection topology of the random network considered here is characterized by an arbitrary degree distribution $P(k)$, which is the probability that a randomly chosen node has degree k . Starting from a randomly selected node (root), all nodes are listed in ascending order of their distances from the root. The set of nodes that are at distance l from the root is referred to as shell l^{th} . Nodes that are at the same distance from the root are deployed randomly in the same shell.

We assume that the network suffers from random failures after which a fraction q of nodes is functional (see Fig. 1(a)). Our localized recovery process is initiated by first checking the root node, then the nodes in the first shell, and so forth in the ascending order of their distance from the root according to the following rules:

1. Each failed node under check is recovered, and the edge connecting two occupied (i.e., functional or recovered) nodes become active (i.e., present) automatically.
2. Nodes in the same shell are checked randomly and we begin checking nodes in shell $l + 1$ only after all nodes in shell l are checked.

We continue the recovery process until a fraction r of failed nodes are recovered. Hence, a recovery area appears around the root node, and the remaining $1 - r$ fraction of failed nodes are at larger distance from the root (see Fig. 1(b)).

The key observation here is that the critical recovery fraction $r_0(\text{LR})$ (Here, $r_0(\text{LR})$ means r_0 under LR; similar notations will be used for r_c , P_∞ , and RR, etc.) signifying the emergence of a giant component of occupied nodes

is equal to zero for any $q \in [0, 1]$ when the original network is connected, i.e., having a single component. This observation motivates our interest in a closely related critical recovery fraction $r_c(\text{LR})$, pointing to the threshold of a giant component in the network of failed nodes after recovery. This sort of study has important practical relevance since it informs us when the failures (such as epidemics in a population or virus in cyber spaces) can be controlled/confined to local areas, as well as a prediction of full system restoration.

We also consider the random recovery strategy as a comparison, where each initially failed node is recovered with probability r independently of each other, and edges connecting two occupied (i.e., functional or recovered) nodes are present automatically. The random recovery process can be easily accommodated by the classical node percolation theory²⁹ with occupation probability $q + r - qr$. Following refs 29,30, the probability generating functions of node degree and excess node degree are defined as $G_0(x) = \sum_{k=0}^{\infty} P(k)x^k$ and $G_1(x) = G_0'(x)/\langle k \rangle$, respectively, where $\langle k \rangle = G_0'(1)$ is the average degree of the original network. We find that the critical recovery probability is (see Supplementary Note A)

$$r_0(\text{RR}) = \max \left\{ 0, \frac{1 - qG_1'(1)}{(1 - q)G_1'(1)} \right\}. \quad (1)$$

Hence, when $q < 1/G_1'(1)$, the initial random failure collapses the network, and the RR process plays a role in integrating the network.

Network of failed nodes under RR. Recall that the generating function for the degree distribution of the original network is $G_0(x) = \sum_{k=0}^{\infty} P(k)x^k$. Based upon the concept of induced subgraph on the set complement in graph theory, the induced network of failed nodes under RR can be readily viewed as the result of independent node percolation²⁹ with “occupation” probability $(1 - q)(1 - r)$. Recall that $(1 - q)(1 - r)$ is the fraction of failed nodes after recovery, and hence it is the probability that a node is in the complement network (composed of failed nodes). Under LR, for example, the complement network is the induced subnetwork of white nodes depicted in Fig. 1(b).

If we focus on the complement network, following ref. 29, the size distribution of the clusters that can be reached following a randomly chosen edge is generated in a self-consistent equation

$$H_1(x) = q + r - qr + (1 - q)(1 - r)xG_1(H_1(x)), \quad (2)$$

where $G_1(x) = G_0'(x)/G_0'(1)$. Here, the first term $q + r - qr$ represents the probability that a node is occupied, namely, not in the complement network. The second term $(1 - q)(1 - r)xG_1(H_1(x))$ addresses the situation where the edge lead to a failed node with k other edges leading out of it, distributed according to the generating function $(1 - q)(1 - r)G_1(x)$. Analogously, the size distribution of the cluster to which a randomly chosen node belongs is generated by

$$H_0(x) = q + r - qr + (1 - q)(1 - r)xG_0(H_1(x)), \quad (3)$$

where the second term $(1 - q)(1 - r)xG_0(H_1(x))$ addresses the situation where the chosen node is failed (hence included in the complement network), and it has k other edges leading out of it, distributed according to the generating function $(1 - q)(1 - r)G_0(x)$. Then the critical recovery probability r_c , at which a giant component of the network of failed nodes emerges, is determined by

$$r_c(\text{RR}) = \max \left\{ 0, \frac{1 + (q - 1)G_1'(1)}{(q - 1)G_1'(1)} \right\}. \quad (4)$$

When the initial random failure is not severe, i.e., $q \geq q_c = 1 - 1/G_1'(1)$, the network of failed nodes remains small components even without recovery (since $r_c = 0$). In general, a giant component of failed nodes first forms when $r < r_c$.

Comparing the two thresholds (1) and (4), we observe interestingly that the critical point $G_1'(1) = 2$ separates a regime of $r_0(\text{RR}) > r_c(\text{RR})$ and that of $r_0(\text{RR}) < r_c(\text{RR})$. The quantity $G_1'(1) = (\langle k^2 \rangle - \langle k \rangle)/\langle k \rangle$ is called the branching factor of the original network, specifying the average number of outgoing links of the underlying branching process^{2,30}. In a sparse original network with $G_1'(1) < 2$, we have $r_0(\text{RR}) > r_c(\text{RR})$. This can be intuitively explained as follows. With the increase of the recovery fraction r , the giant component in the network of failed nodes first collapses and then a giant component in the network of occupied nodes forms. In other words, there is a coexistence phase where both networks have only small components. This is due to the sparseness of the network. While in a dense original network with $G_1'(1) > 2$, we have $r_0(\text{RR}) < r_c(\text{RR})$ indicating a coexistence of giant components in both subnetworks of occupied nodes and failed nodes. (See Fig. 2(a) and Supplementary Note C for the evolution of phase diagrams).

The fraction S of the giant component in the network of failed nodes satisfies

$$S(\text{RR}) = 1 - H_0(1) = (1 - q)(1 - r)(1 - G_0(H_1(1))), \quad (5)$$

where $H_1(1)$ satisfies $H_1(1) = (q + r - qr) + (1 - q)(1 - r)G_1(H_1(1))$. We denote by P_∞ the relative size of the giant component as a fraction of the original network. Clearly, $P_\infty(\text{RR}) = S(\text{RR})$.

In what follows we instantiate the above formula (4) and (5) in the special cases of ER and REG networks (see Methods section for definitions). For ER networks, $G_0(x) = G_1(x) = e^{\lambda(x-1)}$, and the critical recovery probability is

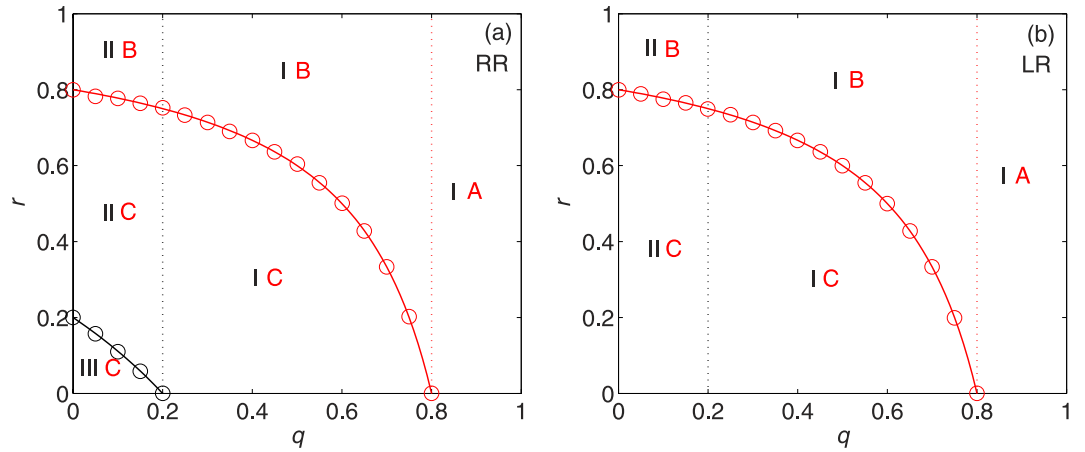


Figure 2. Phase diagrams in the q - r plane for ER networks with $\lambda=5$ under (a) RR and (b) LR strategies. The meanings of the combined phases are described in Tab. 2. Solid lines are analytical results, from (1) for r_0 (black line) and (6) for r_c (red lines). Data points (black and red circles) correspond to the simulation results averaged over 30 random graphs with 20 independent realizations for each.

$$r_c(\text{RR}) = \max\left\{0, \frac{1 + (q - 1)\lambda}{(q - 1)\lambda}\right\}. \tag{6}$$

The relative size of the giant component satisfies

$$P_\infty(\text{RR}) = (1 - q)(1 - r)(1 - e^{\lambda(H_1(1)-1)}), \tag{7}$$

where $H_1(1)$ satisfies $H_1(1) = (q + r - qr) + (1 - q)(1 - r)e^{\lambda(H_1(1)-1)}$. For REG networks, $G_0(x) = xG_1(x) = x^{k_0}$, and the critical recovery probability is

$$r_c(\text{RR}) = \max\left\{0, \frac{1 + (q - 1)(k_0 - 1)}{(q - 1)(k_0 - 1)}\right\}. \tag{8}$$

Similarly, the relative size of the giant component satisfies

$$P_\infty(\text{RR}) = (1 - q)(1 - r)(1 - H_1(1)^{k_0}), \tag{9}$$

where $H_1(1)$ satisfies $H_1(1) = (q + r - qr) + (1 - q)(1 - r)H_1(1)^{k_0-1}$.

Network of failed nodes under LR. As in the RR scenario above, the critical recovery probability $r_c(\text{LR})=0$ (i.e., no recovery is virtually needed to maintain the functionality of the original network) when the initial random damage is below a threshold, namely, $q \geq q_c = 1 - 1/G_1'(1)$. We find that $r_c(\text{LR})$ satisfies (see Supplementary Note B)

$$G_0'(1) = (1 - q)G_0''(G_0^{-1}(1 - r_c)), \tag{10}$$

when $q \leq 1 - 1/G_1'(1)$. Recall that $r_0(\text{LR}) \equiv 0$, and hence the inequality $r_0(\text{LR}) \leq r_c(\text{LR})$ always holds, which is in sharp contrast to the RR case. Clearly, whenever the branching factor $G_1'(1) > 1$, there is a co-existence phase where both subnetworks of occupied nodes and failed nodes have giant components (see the figures below for ER, REG and SF networks, and Methods section for the definition of these networks).

Furthermore, we find that the relative size of the giant component as a fraction of the original network is (see Supplementary Note B)

$$P_\infty(\text{LR}) = (1 - q)(1 - r)(1 - \tilde{G}_0(u)), \tag{11}$$

where u satisfies $u = q + (1 - q)\tilde{G}_0'(u)/\tilde{G}_0'(1)$, and

$$\tilde{G}_0(x) = \frac{1}{1 - r}G_0\left[G_0^{-1}(1 - r) + \frac{G_0'(G_0^{-1}(1 - r))}{G_0'(1)}(x - 1)\right]. \tag{12}$$

In what follows we derive explicit analytical expressions for the formula (10) and (11) in the special cases of ER and REG networks. For ER networks, $\tilde{G}_0(x) = e^{\lambda(1-r)(x-1)}$ using (12). The critical recovery probability $r_c(\text{LR}) = r_c(\text{RR})$ which is given by (6). The relative size of the giant component satisfies

Phase	Description
Network of occupied nodes	
I (Non Collapse)	Existence of giant component w/o recovery
II (Recovery)	Existence of giant component w/recovery
III (Collapse)	Non-existence of giant component w/recovery
Network of failed nodes	
A (Non Prevalence)	Non-existence of giant component w/o recovery
B (Restriction)	Non-existence of giant component w/recovery
C (Prevalence)	Existence of giant component w/recovery

Table 1. Three phases for the network composed of occupied (namely, functional and recovered) nodes and three for the “complement network” composed of failed nodes.

$$P_\infty(\text{LR}) = (1 - q)(1 - r)(1 - e^{\lambda(1-r)(u-1)}), \tag{13}$$

where u satisfies $u = q + (1 - q)e^{\lambda(1-r)(u-1)}$. For REG networks, $\tilde{G}_0(x) = (1 + (1 - r)^{\frac{k_0-2}{k_0}}(x - 1))^{k_0}$, and the critical recovery probability is

$$r_c(\text{LR}) = \max\left\{0, 1 - ((1 - q)(k_0 - 1))^{-\frac{k_0}{k_0-2}}\right\}, \tag{14}$$

for $k_0 > 2$ and $r_c(\text{LR}) = 0$ for $k_0 = 2$. The relative size of the giant component satisfies

$$P_\infty(\text{LR}) = (1 - q)(1 - r)(1 - (1 + (1 - r)^{\frac{k_0-2}{k_0}}(u - 1))^{k_0}), \tag{15}$$

where u satisfies $u = q + (1 - q)(1 + (1 - r)^{\frac{k_0-2}{k_0}}(u - 1))^{k_0-1}$.

ER networks under RR and LR. Figure 2 shows the (combined) phase diagrams for ER networks with average degree $\langle k \rangle = 5$ under RR and LR (see also Table 1). The lines of the critical recovery probabilities r_0 and r_c , in good agreement with the simulation results, separate the regimes of (non)existence of giant component in the network of occupied nodes and in the network of failed nodes, respectively. As expected from the theoretical prediction, under RR there is a coexistence phase (i.e., the broken ring-shaped part between the red and black lines in Fig. 2(a)) where both networks have giant components since $5 = G_1'(1) > 2$. This coexistence phase disappears when $\langle k \rangle = 2$ and a different coexistence phase where both networks are absent of giant components emerges when $\langle k \rangle < 2$ (see Supplementary Note C).

Comparing Fig. 2(b) with Fig. 2(a), we observe that the combined phase IIIC only exists under RR. This means that the network of occupied nodes possesses a giant component as long as $r > 0$ under LR, which constitutes a unique feature of LR. On the other hand, the two curves $r_c(\text{RR})$ and $r_c(\text{LR})$ coincide, which indicates that RR and LR strategies have the same healing capability (in terms of r_c) for ER networks.

The effect of RR and LR is better appreciated when turning to the results reported in Fig. 3. Several interesting comments can be drawn. First, in all cases the simulation results confirm the theoretical predictions. The equalities $r_c(\text{RR}) = r_c(\text{LR})$ (arrows in Fig. 3) hold, and these critical points agree with those displayed in Fig. 2. For example, recall that q_c is the critical value at which P_∞ vanishes, and we have $q_c = 0.8$ for $r_c = 0$ and $q_c \approx 0.6$ for $r_c = 0.5$. Second, the fraction of giant component P_∞ in the network of failed nodes displays identical behavior under RR and LR for $r = 0, 0.5$, and 0.7 . In fact, such equivalence holds for all $q, r \in [0, 1]$ (see Supplementary Note D for a direct proof), meaning that RR and LR has exactly the same effect in containing the damage of an ER network. This can be interpreted as a detailed balance between two competitive forces: (i) the nature of the LR that nodes within the recovery area are connected to each other while only exterior nodes linked to the surface of the recovery area contribute to the network of occupied nodes (namely, the “recovering resistance”), and (ii) the level of degree heterogeneity, where a larger-degree node is more likely to be recovered under LR, leading even more likely to an occupied giant component (namely, the “recovering impetus”).

It is worth mentioning that a balance phenomenon for ER networks under random and localized attacks has been observed recently by Shao *et al.*²⁶. Their results correspond to the coincidence of $P_\infty(\text{RR})$ and $P_\infty(\text{LR})$ at the point $q = 0$ in Fig. 3. Note that if all nodes are removed initially (i.e., $q = 0$), the giant component in the network of failed nodes after random (or resp., localized) recovery considered here is exactly that in the network of remaining nodes under random (or resp., localized) attack in an original intact network.

REG networks under RR and LR. Figure 4 shows the (combined) phase diagrams for REG networks with constant node degree $k_0 = 5$ under RR and LR. The lines of the critical recovery probabilities r_0 and r_c agree very well with the simulation results. As anticipated, the coexistence phase (phases IC and IIC) of giant components in both networks appears under RR since $G_1'(1) = k_0 - 1 > 2$ (see Fig. 4(a)). Fig. 4(b) again reflects the fact that $r_0(\text{LR}) \equiv 0$. Comparing Fig. 4(a) with Fig. 4(b), we find that $r_c(\text{RR}) < r_c(\text{LR})$ for all $q < q_c = 0.75$. This means that

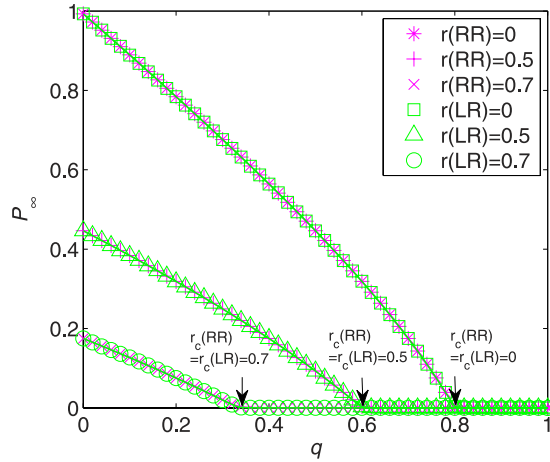


Figure 3. Relative sizes of giant component in the network of failed nodes, P_∞ , as a function of q for ER networks with $\lambda = 5$ and $r(\text{RR}) = 0$ (stars), $r(\text{RR}) = 0.5$ (pluses), $r(\text{RR}) = 0.7$ (crosses), $r(\text{LR}) = 0$ (squares), $r(\text{LR}) = 0.5$ (triangles), and $r(\text{LR}) = 0.7$ (circles). Solid lines are analytical results, from (7) for RR (magnet lines) and (13) for LR (green lines). Symbols correspond to the simulation results averaged over 30 random graphs with 20 independent realizations for each. The critical recovery fractions are indicated by arrows.

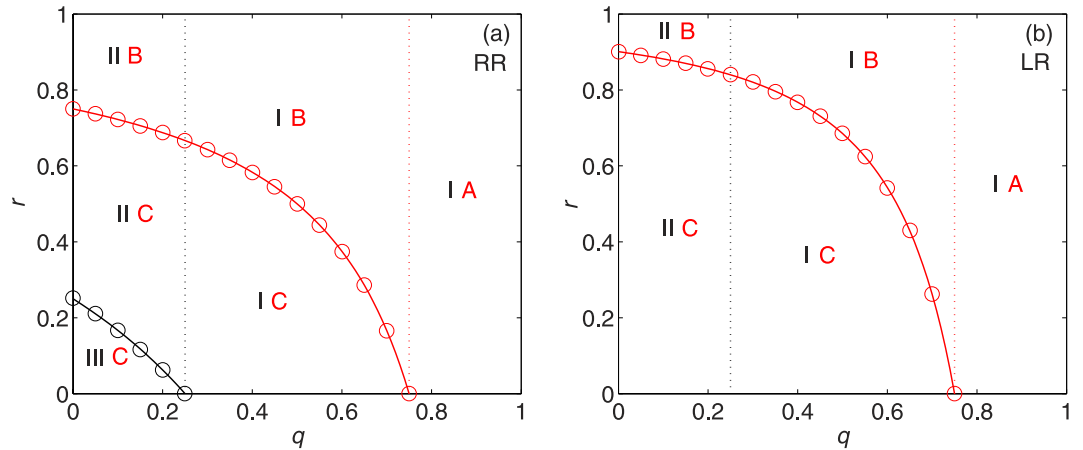


Figure 4. Phase diagrams in the q - r plane for REG networks with $k_0 = 5$ under (a) RR and (b) LR strategies. The meanings of the combined phases are described in Table 1. Solid lines are analytical results, from (1) for r_0 (black line) and (8), (14) for r_c (red lines). Data points (black and red circles) correspond to the simulation results averaged over 30 random graphs with 20 independent realizations for each.

LR is always less powerful than RR, namely, more nodes have to be recovered under LR in order to confine the damage to small/scattered areas.

The insight that RR is more effective for healing an REG network is better fathomed when it comes to Fig. 5. Similarly as in ER networks, the critical points of percolation, $r_c(\text{RR})$ and $r_c(\text{LR})$ (indicated by arrows), agree with those displayed in Fig. 4. In addition to these discrete points, the whole curve of P_∞ follows the same pattern that $P_\infty(\text{RR}) < P_\infty(\text{LR})$ for $r > 0$. This can be easily understood as there is no “recovering impetus” (the degrees are all the same in REG networks); the “recovering resistance” renders LR inferior for healing REG networks after any degree of damage.

SF networks under RR and LR. In Fig. 6 we plot the (combined) phase diagrams for SF networks with $\gamma = 2.4$, $k_{\min} = 2$, and $\langle k \rangle = 5.08$ under RR and LR. The simulations for the critical recovery probabilities r_0 and r_c are consistent with theoretical results. Similarly as in ER and REG networks, the coexistence phase (phases IC and IIC) of giant components in both networks appears under RR (see Fig. 6(a)) since $G_1'(1) > 2$ (which is actually more than 50 here).

Remarkably, the phases IA, IIB, IIC, IIIC under RR (Fig. 6(a)) and the phases IA, IIB, IIC under LR (Fig. 6(b)) are very small yet not negligible. This phenomenon is known to be the finite-size effect^{2,31}. For an infinite SF network with power-law distribution $P(k) \sim k^{-\gamma}$, the ratio $\langle k^2 \rangle / \langle k \rangle = G_1'(1) + 1$ diverges when $1 \leq \gamma < 3$. So, in an infinite SF network, we have $q_c = 1 - 1/G_1'(1) = 1$, meaning that for any $q > 0$ the network of failed nodes always has a giant component (hence is extremely robust^{2,5}) even without recovery (i.e., $r = 0$). When the recovery

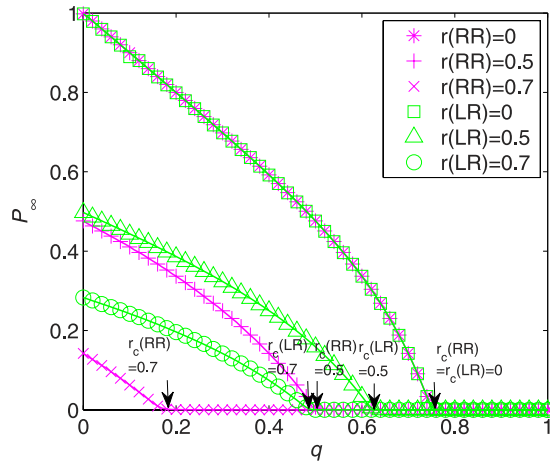


Figure 5. Relative sizes of giant component in the network of failed nodes, P_∞ , as a function of q for REG networks with $k_0 = 5$ and $r(\text{RR}) = 0$ (stars), $r(\text{RR}) = 0.5$ (pluses), $r(\text{RR}) = 0.7$ (crosses), $r(\text{LR}) = 0$ (squares), $r(\text{LR}) = 0.5$ (triangles), and $r(\text{LR}) = 0.7$ (circles). Solid lines are analytical results, from (9) for RR (magenta lines) and (15) for LR (green lines). Symbols correspond to the simulation results averaged over 30 random graphs with 20 independent realizations for each. The critical recovery fractions are indicated by arrows.

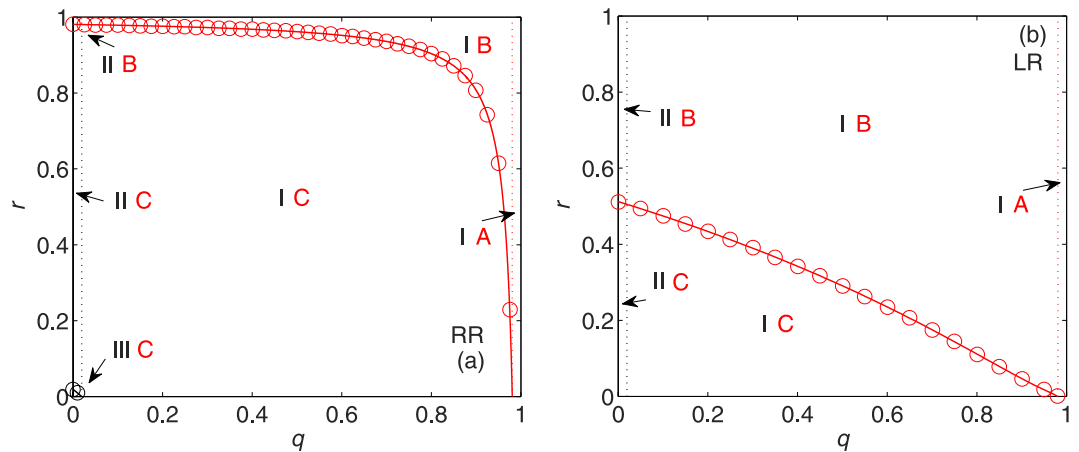


Figure 6. Phase diagrams in the q - r plane for SF networks with $\gamma = 2.4$, $k_{\min} = 2$, and $\langle k \rangle = 5.08$ under (a) RR and (b) LR strategies. The meanings of the combined phases are described in Tab. 2. Solid lines are analytical results, from (1) for r_0 (black line) and (4), (10) for r_c (red lines). Data points (black and red circles) correspond to the simulation results averaged over 30 random graphs with 20 independent realizations for each.

process is taken into consideration, the red curve r_c divides the whole phase plane into two regions: IB, where the network of failed nodes has no giant component and IC, where it has a giant component. Since $r_c(\text{LR}) < r_c(\text{RR})$ for all q , we conclude that LR is more powerful than RR for healing SF networks, opposite to the best practice for healing REG networks.

In addition to r_c , the entire curve of P_∞ follows the same pattern that $P_\infty(\text{LR}) < P_\infty(\text{RR})$ for $r > 0$ (see Fig. 7), meaning that the superiority of LR consists not only in the threshold but also in the size of giant component. Clearly, the level of degree heterogeneity (i.e., the “recovering impetus”) is dominant in SF networks: recovered hub nodes accelerate the recovery of the original network (hence rapidly collapse the network of failed nodes).

When the scaling exponent γ in an SF network is getting larger, the network becomes less heterogeneous. A phenomenon similar to that in REG networks, e.g., $r_c(\text{LR}) > r_c(\text{RR})$, is expected for a large γ (see Supplementary Note E). Then it is possible to determine a critical exponent $\gamma_c = \gamma_c(q)$ so that $r_c(\text{RR}) = r_c(\text{LR})$ at q . Indeed, if a given network has a power-law degree distribution with exponent γ , we can identify $\gamma_c(q)$ for any $q \in [0, 1]$ by using Eqs (4),(10) and parameters in $P(k)$ such as k_{\min} and k_{\max} . Likewise, we can identify another critical exponent $\gamma_\infty = \gamma_\infty(q, r)$ so that $P_\infty(\text{RR}) = P_\infty(\text{LR})$ at any desired values of q and r . By comparing γ , $\gamma_c(q)$, and $\gamma_\infty(q, r)$, we are able to determine the best recovery strategy (RR or LR) and acquire a good estimation on its effect under any level of initial failure as well as recovery in the network by using the degree distribution of the original network as the only input. Applications to some real-world networks with different sizes can be found in Supplementary Note F.

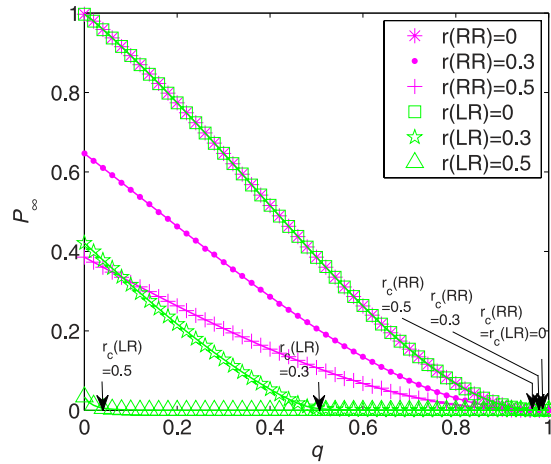


Figure 7. Relative sizes of giant component in the network of failed nodes, P_∞ , as a function of q for SF networks with $\gamma = 2.4$, $k_{\min} = 2$, $\langle k \rangle = 5.08$ and $r(\text{RR}) = 0$ (stars), $r(\text{RR}) = 0.3$ (dots), $r(\text{RR}) = 0.5$ (pluses), $r(\text{LR}) = 0$ (squares), $r(\text{LR}) = 0.3$ (pentagrams), and $r(\text{LR}) = 0.5$ (triangles). Solid lines are analytical results, from (5) for RR (magenta lines) and (11) for LR (green lines). Symbols correspond to the simulation results averaged over 30 random graphs with 20 independent realizations for each. The critical recovery fractions are indicated by arrows.

Discussion

We have proposed and studied the localized recovery strategy for complex networks under random failures. Based on the generating function formalism, we tackle analytically the percolation properties associated with both the network of occupied nodes and the “complement network” of failed nodes. Detailed phase diagrams under RR and LR are obtained. We identify two competitive forces behind LR strategy, which exactly compensate each other in ER networks while yield an interesting crossover phenomenon in SF networks with respect to the scaling exponent γ . Our developed analytical framework successfully guides us in choosing more powerful recovery strategy and provides estimation on the desired recovery fraction by using the degree distribution of the original network as the only input.

Note that the analytical and numerical results presented in this work are based on uncorrelated graphs. However, correlations between nodes of similar degree are often found in real-world networks. The tendency to be connected with other nodes with similar degree is referred to as assortative mixing, while disassortative mixing is a bias that high degree nodes tend to attach to low degree nodes^{32–34}. Mathematically, the assortativity/disassortativity property can be measured by the Pearson correlation coefficient ρ averaged for all pairs of adjacent nodes in the network. In general, $\rho \in [-1, 1]$ with $\rho > 0$ indicating an assortative network while $\rho < 0$ a disassortative network. In Supplementary Note G, we compare RR and LR strategies on two real social and biological networks. It is revealed that LR seems to be more powerful than RR for healing assortative networks, while the opposite is true for disassortative networks.

This phenomenon can be explained by the dominance of “recovering impetus” in assortative networks in that groups of hub nodes are more likely to be recovered under LR accelerating the recovery process. On the other hand, the disassortative mixing tend to impede the recovery process adding to the “recovering resistance”. We mention that there are other topological properties, such as motif and clustering, that are likely to affect the LR strategy. In addition to numerical study, closed-form results are highly desirable. It is hoped that this work could stimulate further research efforts on the related problems in localized network recovery against damages.

Methods

Based on the generating function formalism^{24,26,29,30}, we are able to solve our theoretical model for the (induced) network composed of failed nodes under RR and LR. We apply our theoretical results to three different types of complex networks: Erdős-Rényi (ER) networks, random regular (REG) networks, and scale-free (SF) networks. An ER network follows a Poisson degree distribution $P(k) = e^{-\lambda} \lambda^k / k!$ ($k \geq 0$) with average degree $\langle k \rangle = \lambda$. An REG network has a degenerated degree distribution $P(k) = \delta_{k,k_0}$, meaning that each node is connected to the same number k_0 of neighbors. A SF network follows a power-law degree distribution $P(k) \sim k^{-\gamma}$ ($k_{\min} \leq k \leq k_{\max}$), where $\gamma > 0$ is the scaling exponent, k_{\min} and k_{\max} mean the minimum and maximum degrees, respectively. For all the simulations, we use networks of size $N = 10^6$. Moreover, three real-life networks with power-law degree distributions from biological and technological fields are investigated (see Supplementary Note F).

To obtain the r_c curve in the q - r plane, we first label each node as failed independently with probability $1 - q$ for any given $q \in [0, 1]$ and inactivate its incident edges. For RR strategy, we begin with $r = 0$ and take a node from the list of failed nodes, change its label to functional with probability r , and reactivate each of its incident edges if the other end is connected to a functional node. For LR strategy, we begin with $r = 0$ and take a node from the list of failed nodes in an increasing order according to the distance from a root node, change its label to functional, and reactivate each of its incident edges if the other end is connected to a functional node. After checking

the whole list for RR strategy (or respectively, after checking r fraction of failed nodes for LR strategy), we then calculate the fraction P_∞ of the giant cluster in the network of failed nodes. Increase r and repeat the process until $P_\infty < 10^{-3}$. The r_0 curve is obtained similarly by checking the fraction of the giant cluster in the network of functional nodes.

References

1. Albert, R., Jeong, H. & Barabási, A.-L. Error and attack tolerance of complex networks. *Nature*, 406, 378–382 (2000).
2. Cohen R. & Havlin, S. *Complex Networks: Structure, Robustness and Function*. Cambridge University Press, Cambridge, (2010).
3. Buldyrev, S. V., Parshani, R., Paul, G., Stanley, H. E. & Havlin, S. Catastrophic cascade of failures in interdependent networks. *Nature* **464**, 1025–1028 (2010).
4. Dorogovtsev, S. N., Goltsev, A. V. & Mendes, J. F. F. Critical phenomena in complex networks. *Rev. Mod. Phys.* **80**, 1275–1335 (2008).
5. Cohen, R., Erez, K., ben-Avraham, D. & Havlin, S. Breakdown of the internet under intentional attack. *Phys. Rev. Lett.* **86**, 3682–3685 (2001).
6. Iyer, S., Killingback, T., Sundaram, B. & Wang, Z. Attack robustness and centrality of complex networks. *PLoS ONE* **8**, e59613 (2013).
7. Gao, J., Buldyrev, S. V., Havlin, S. & Stanley, H. E. Robustness of a network of networks. *Phys. Rev. Lett.* **107**, 195701 (2011).
8. Dong, G. *et al.* Robustness of network of networks under targeted attack. *Phys. Rev. E* **87**, 052804 (2013).
9. Pahwa, S., Scoglio, C. & Scala, A. Abruptness of cascade failures in power grids. *Sci. Rep.* **4**, 3694 (2014).
10. Shekhtman, L. M., Shai, S. & Havlin, S. Resilience of networks formed of interdependent modular networks. *New J. Phys.* **17**, 123007 (2015).
11. Peixoto, T. P. & Bornholdt, S. Evolution of robust network topologies: emergence of central backbones. *Phys. Rev. Lett.* **109**, 118703 (2012).
12. Min, B., Yi, S. D., Lee, K.-M. & Goh, K.-I. Network robustness of multiplex networks with interlayer degree correlations. *Phys. Rev. E* **89**, 042811 (2014).
13. Majdandzic, A. *et al.* Spontaneous recovery in dynamical networks. *Nat. Phys.* **10**, 34–38 (2014).
14. Shang, Y. Impact of self-healing capability on network robustness. *Phys. Rev. E* **91**, 042804 (2015).
15. Podobnik, B. *et al.* Network risk and forecasting power in phase-flipping dynamical networks. *Phys. Rev. E* **89**, 042807 (2014).
16. Gong, M., Ma, L., Cai, Q. & Jiao, L. Enhancing robustness of coupled networks under targeted recoveries. *Sci. Rep.* **5**, 8439 (2015).
17. Di Muro, M. A., La Rocca, C. E., Stanley, H. E., Havlin, S. & Braunstein, L. A. Recovery of interdependent networks. *Sci. Rep.* **6**, 22834 (2016).
18. Gallos, L. K. & Fefferman, N. H. Simple and efficient self-healing strategy for damaged complex networks. *Phys. Rev. E* **92**, 052806 (2015).
19. Schneider, C. M., Moreira, A. A., Andrade, Jr. J. S., Havlin, S. & Herrmann, H. J. Mitigation of malicious attacks on networks. *Proc. Natl. Acad. Sci. USA* **108**, 3838–3841 (2011).
20. Schneider, C. M., Yazdani, N., Araújo, N. A. M., Havlin, S. & Herrmann, H. J. Towards designing robust coupled networks. *Sci. Rep.* **3**, 1969 (2013).
21. Parshani, R., Buldyrev, S. V. & Havlin, S. Interdependent networks: reducing the coupling strength leads to a change from a first to second order percolation transition. *Phys. Rev. Lett.* **105**, 048701 (2010).
22. Cao, X. B., Hong, C., Du, W. B. & Zhang, J. Improving the network robustness against cascading failures by adding links. *Chaos Solitons Fractals* **57**, 35–40 (2013).
23. Quattrociocchi, W., Caldarelli, G. & Scala, A. Self-healing networks: redundancy and structure. *PLoS ONE* **9**, e87986 (2014).
24. Shao, J., Buldyrev, S. V., Braunstein, L. A., Havlin, S. & Stanley, H. E. Structure of shells in complex networks. *Phys. Rev. E* **80**, 036105 (2009).
25. Shang, Y., Luo, W. & Xu, S. I -hop percolation on networks with arbitrary degree distributions and its applications. *Phys. Rev. E* **84**, 031113 (2011).
26. Shao, S., Huang, X., Stanley, H. E. & Havlin, S. Percolation of localized attack on complex networks. *New J. Phys.* **17**, 023049 (2015).
27. Berezin, Y., Bashan, A., Danziger, M. M., Li, D. & Havlin, S. Localized attacks on spatially embedded networks with dependencies. *Sci. Rep.* **5**, 8934 (2015).
28. Yuan, X., Shao, S., Stanley, H. E. & Havlin, S. How breadth of degree distribution influences network robustness: comparing localized and random attacks. *Phys. Rev. E* **92**, 032122 (2015).
29. Callaway, D. S., Newman, M. E. J., Strogatz, S. H. & Watts, D. J. Network robustness and fragility: Percolation on random graphs. *Phys. Rev. Lett.* **85**(25), 5468–5471 (Dec 2000).
30. Newman, M. E. J., Strogatz, S. H. & Watts, D. J. Random graphs with arbitrary degree distributions and their applications. *Phys. Rev. E* **64**(2), 026118 (Jul 2001).
31. Cohen, R., Erez, K., ben-Avraham, D. & Havlin, S. Resilience of the internet to random breakdowns. *Phys. Rev. Lett.* **85**, 4626–4628 (2000).
32. Newman, M. Assortative mixing in networks. *Phys. Rev. Lett.* **89**, 208701 (2002).
33. Vázquez, A., Boguñá, M., Moreno, Y., Pastor-Satorras, R. & Vespignani, A. Topology and correlations in structured scale-free networks. *Phys. Rev. E* **67**, 046111 (2003).
34. Sendiña-Nadal, I., Danziger, M. M., Wang, Z., Havlin, S. & Boccaletti, S. Assortativity and leadership emerge from anti-preferential attachment in heterogeneous networks. *Sci. Rep.* **6**, 21297 (2016).

Acknowledgements

The author wishes to thank the anonymous referees for their constructive suggestions for improving the manuscript. The author acknowledges support from the National Natural Science Foundation of China (11505127), the Shanghai Pujiang Program (15PJ1408300), and the Program for Young Excellent Talents in Tongji University (2014KJ036).

Author Contributions

Y.S. wrote the manuscript.

Additional Information

Supplementary information accompanies this paper at <http://www.nature.com/srep>

Competing financial interests: The author declares no competing financial interests.

How to cite this article: Shang, Y. Localized recovery of complex networks against failure. *Sci. Rep.* **6**, 30521; doi: 10.1038/srep30521 (2016).



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>

© The Author(s) 2016