University of Konstanz | Network Science of Socio-Economic Systems

Lea Froitzheim – Winter 2024/2025

Error and Attack Tolerance of Complex Networks

by Réka Albert, Hawoong Jeong & Albert-László Barabási

Abstract

This study investigates the robustness of complex networks under two types of node failures: random failures and targeted attacks. Building on the work of Réka Albert and Barabási (2000), I replicate and extend their analysis by examining how different network topologies—specifically random networks and scale-free networks—respond to node removals. The networks' responses are evaluated using two key metrics: the p-diameter, defined as the average shortest path length in the largest connected component, and fragmentation, which measures tracking the size and distribution of disconnected clusters. The results confirm the original study's findings that scale-free networks are highly resilient to random failures but extremely vulnerable to targeted attacks. However, discrepancies in the behavior of random networks—particularly under random failure—highlight the sensitivity of such simulations to implementation details, including graph generation methods and statistical averaging. These findings emphasize the importance of methodological transparency in network robustness research.

1 Introduction

In today's interconnected world, complex networks form the backbone of critical infrastructure—from the power grid and transportation systems to communication networks and the internet. The malfunction or failure of individual components in these systems can lead to cascading effects and large-scale disruptions, such as blackouts or communication breakdowns. Understanding how these networks behave under failure conditions is essential for designing systems that are robust and resilient.

This replication study focuses on two primary types of node failure, random failures and targeted attacks. Random failures occur unpredictably, such as hardware malfunctions or accidental outages affecting arbitrary parts of the network. Targeted attacks, on the other hand, deliberately aim to disrupt the system by removing key nodes—those with the highest number of connections. Such attacks might resemble coordinated cyberattacks or strategic strikes on critical infrastructure.

To investigate how different network structures respond to these types of failures, this study replicates the analysis by Réka Albert and Barabási (2000), who studied the robustness of two well-known network models, random networks and scale-free networks. Random networks, often modeled using the Erdős–Rényi (ER) model, have a more uniform degree distribution, meaning most nodes have a more number of connections. Scale-free networks are characterized by a power-law degree distribution—most nodes have few connections, while a small number of nodes are highly connected hubs.

Despite both network types have redundancy in their connectivity, they respond very differently to failures. Scale-free networks are resilient to random node removals but highly vulnerable to targeted attacks. Random networks, in contrast, display a more symmetric—but less extreme—response to both types of failure.

This replication study not only reproduces key aspects of the original study but also highlights where results diverge, potentially due to differences in graph generation, statistical averaging, or the definition of metrics. These findings highlight the importance of transparency in methodologies and offer additional perspectives on how structural features influence network resilience.

2 Methods

Malfunction tolerance of networks will be tested on two different types of network models.

The first model is the *random network*, which the authors obtained with the Erdős–Rényi model. There, a graph is generated with a predefined number of nodes n and an edge probability p [Erdős and Rényi 1959].

However, this approach is not well-suited for generating large, sparse, and connected graphs, which is necessary for certain network metrics such as the graph diameter. To address this limitation, I propose an alternative method, the *two-phase model*, which generates graphs that closely resemble those produced by the ER model while ensuring connectivity. Here, the graph generation process consists of two phases:

1. *Edge Initialization:* First, an potentially unconnected graph with a given number of nodes is generated. Pairs of nodes are selected at random, and an edge is added if it does not already exist. This process continues until 95% of the edges have been added.

2. *Connectivity Enforcement:* If the graph remains disconnected, two nodes from different unconnected components are selected and an edge is added between them. This process is repeated until the graph becomes fully connected. Once connectivity is achieved, the model returns to Phase 1 to complete the edge generation process until the desired number of edges is reached.

To verify that the two-phase model produces graphs with structural properties similar to those of the Erdős–Rényi model, I conducted a series of comparative tests. The observed deviations between the two models are minimal, as illustrated in Table 1. Therefore, I consider graphs resulting from the two-phase model to be approximately equivalent to the Erdős–Rényi model in terms of overall network characteristics.

The second model is the scale-free network, where the graph starts with an initial set of m_0 nodes. At each time step, a new node is introduced and connects to m of the already existing nodes. The probability Π_i that the new node connects to an existing node i depends on the number of connections k_i that node i already has, following the rule [Barabási and Albert 1999]

$$\Pi_i = \frac{k_i}{\sum_j k_j}.$$

The key characteristic relevant to this study is the connectivity distribution of the networks. In a random network, the connectivity distribution has a peak at the average degree $\langle k \rangle$ and decays exponentially for large k. In contrast, the connectivity distribution of a scale-free network follows a power-law distribution [Barabási and Albert 1999]

$$P(k) \sim k^{-\gamma}.$$

For the empirical simulations, it is important that all networks—regardless of the generation model—contain the same number of nodes and edges. To achieve this for the scale-free model, I generated graphs with a fixed number of nodes and an initial connection probability.

I then adjusted the resulting graph to match the desired total number of edges: if the graph contains too few edges, additional edges are added between randomly selected node pairs; if it has too many, some edges are randomly removed. To ensure that these modifications did not significantly alter the fundamental properties of the scale-free model, I compared various structural metrics between the adjusted and unmodified versions of the graph (Table 2. Since the observed discrepancies were minimal, I assume that the modified graphs are valid approximations of the original scale-free model.

Two types of malfunctions were tested on both network models. The first one is random failure in which nodes are removed from the network at random. The second one is a targeted attack, where the nodes with the highest degrees are systematically removed.

The network's response to these failures is evaluated using two key metrics. The first one is the network diameter, defined by Réka Albert and Barabási (2000) as the average shortest path length

between all pairs of nodes in the largest connected component (additional comments on that can be found in 4). In the following, I will refer to this definition as the *paper diameter*, or p-diameter. This serves as a proxy for communication efficiency within the network.

As more nodes are removed, the network may fragment into multiple disconnected components, making the p-diameter undefined for the overall graph.

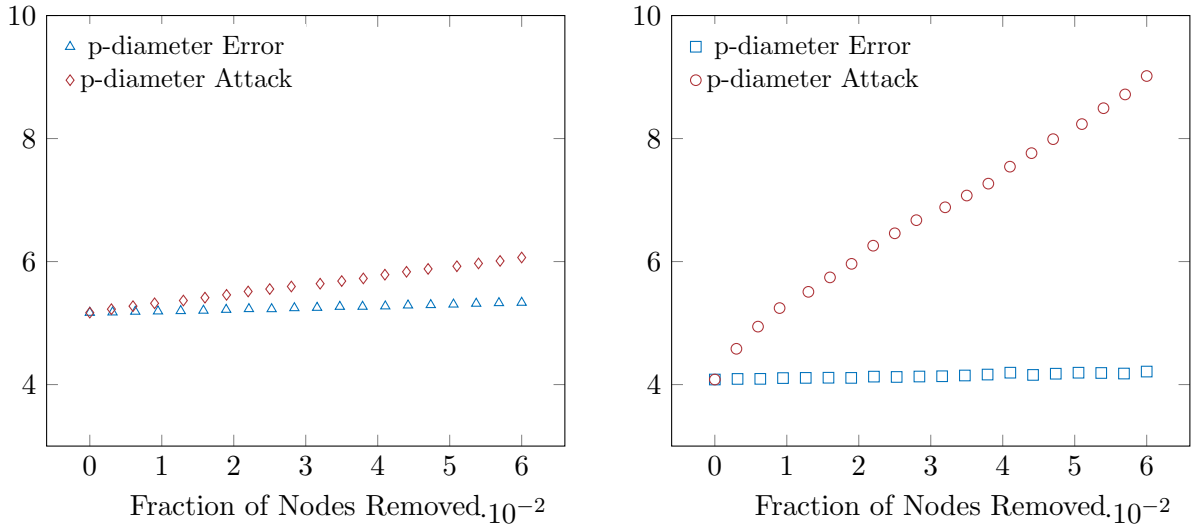
The second metric is a pair of values that track this fragmentation process. The first, S, represents the relative size of the largest connected component—that is, the number of nodes in the largest fragment divided by the total number of nodes in the original network. The second, $\langle s \rangle$, is the average size of all remaining components, excluding the largest one. Together, these metrics provide a detailed view of how the network disintegrates under failure conditions, offering insights into its structural robustness and resilience.

3 Results

3.1 Changes in the P-Diameter with Errors and Attacks

The first set of experiments examines how errors and targeted attacks affect the p-diameter of random and scale-free networks. Since network connectivity is a crucial characteristic, the random network is generated using the two-phase Model (see Section 2) to ensure the computation of p-diameter is feasible.

To obtain statistically reliable results, 30 instances of each network model—random and scale-free—are generated, each containing $1,000$ nodes and $2,000$ edges. The simulation results are shown in Figure 1. Initially, the p-diameter of the random graph starts around 5.16, while the scale-free graph begins at about 4.08. This discrepancy was expected, as the scale-free graph inherently has a smaller p-diameter.



(a) Random Graph

(b) Scale-Free Graph

Figure 1: Fragmentation processes of complex graphs.

Figure 1a illustrates the effect of errors and attacks on the random graphs. With random errors, the p-diameter increases by about 3%, from approximately 5.16 to around 5.33 after 6% of the nodes are randomly removed. However, under targeted attack, the p-diameter grows by 17%, from roughly 5.16 to 6.06 when the same percentage of nodes is removed. Since targeted attacks remove the most connected nodes, and random graphs do not have uniform degree distributions, the steeper increase in p-diameter under attack is to be expected.

In Figure 1b, the effects of errors and attacks on the scale-free graph are shown. Under random errors, the p-diameter of the scale-free network, which starts at about 4.8, decreases slightly to around 4.21 after 6% of the nodes are removed. The change in p-diameter is far more drastic under targeted

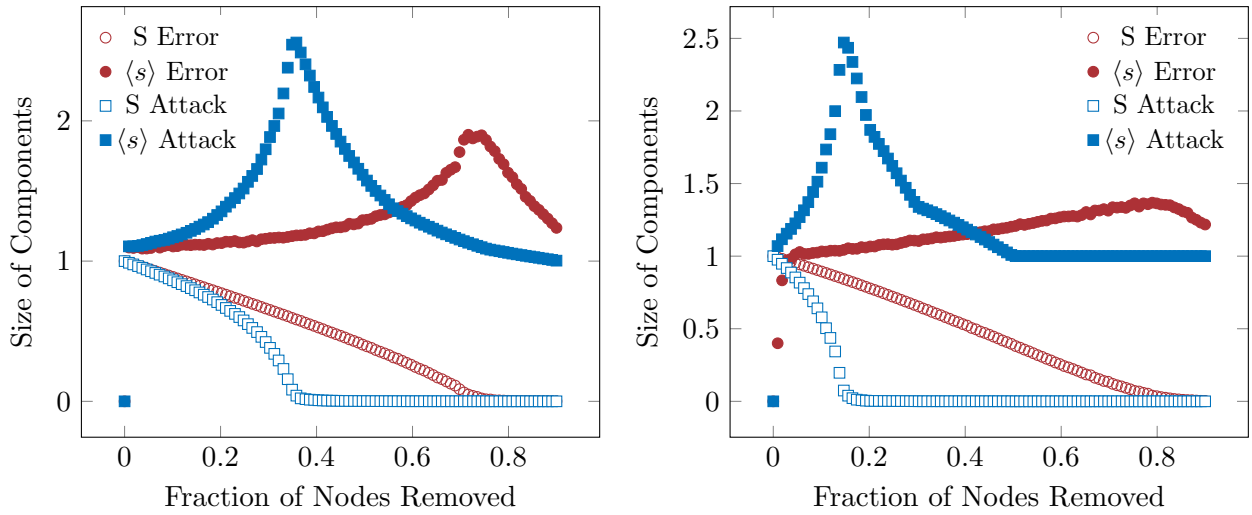
attacks. After 6% of the nodes are removed, the p-diameter increases by an average of approximately 221%, which is significantly higher than the 3% increase under random errors. This increase can be explained by the degree distribution of scale-free networks. A small number of nodes have extremely high degrees and are thus essential for maintaining a low p-diameter. The majority of nodes have relatively low degrees—at least 2 (CHECK). With random failures, it is far more likely to remove one of these low-degree nodes, resulting in only minor increases in p-diameter. In contrast, targeted attacks systematically eliminate the highly connected nodes, significantly reducing overall connectivity and thereby causing the p-diameter to increase substantially.

3.2 Fragmentation Process after Errors and Attacks

When nodes are removed from a network, clusters of nodes that lose their connections to the main system can become disconnected, resulting in fragmentation. To better understand the effects of errors and attacks on the network structure, I investigate this fragmentation process. Specifically, I measure the size of the largest isolated cluster, S, as a fraction of the total system size when a fraction f of the nodes are removed. Additionally, we compute the average size of the smaller clusters, $\langle s \rangle$, which represents the number of nodes in each of those isolated clusters. To get robust results, the measures were taken on 30 graphs, each with $10,000$ nodes and $20,000$ edges.

In the case of random network errors, $\langle s \rangle$ initially starts at 0, as there is only one intact graph. However, as nodes are removed, $\langle s \rangle$ increases, first jumping to 1 and then growing exponentially for $f < 0.75$, as shown in blue in Figure 2a. After this point, $\langle s \rangle$ begins to decrease and eventually stabilizes at 1. The opposite behavior is observed for S, which declines exponentially until it approaches 0 when $f > 0.75$. This pattern is expected as initially node removal mainly isolates individual nodes, shrinking the largest cluster. Once the threshold is crossed, the largest cluster becomes comparable in size to the smaller clusters, and the other clusters continue to shrink.

When the network undergoes a targeted attack, a similar process is observed (red curve in Figure 2a). However, the threshold for significant fragmentation is lower, at approximately $f \approx 0.38$. At this threshold, $\langle s \rangle$ is smaller than in the case of random errors. This is because, in an attack, nodes with higher degrees are preferentially targeted, leading to faster fragmentation of the network.



(a) Random Graph

(b) Scale-Free Graph

Figure 2: Fragmentation processes of complex graphs.

In the case of the scale-free network under attack, as shown in red in Figure 2b, a similar pattern emerges. The largest cluster size S decreases exponentially, while $\langle s \rangle$ grows exponentially until $f \approx 0.18$, after which it shrinks again, ultimately stabilizing at 1. The shift in the threshold compared to the random graph can be easily explained by the unequal distribution of degrees in the scale-free network. With fewer highly connected nodes, the removal of these key nodes causes the network to fragment rapidly.

For random errors in the scale-free network, the same fragmentation pattern is observed (blue curve in Figure 2b), but with a higher threshold of around $f \approx 0.8$. This is due to the inhomogeneity of the degree distribution: the likelihood of randomly removing highly connected nodes is low, so the destruction of the network takes longer.

4 Discussion

Results of my simulations largely confirm the central findings of Réka Albert and Barabási (2000), scale-free networks are highly resilient to random failures but extremely vulnerable to targeted attacks. In contrast, random networks exhibit a more symmetric response to both failure types. However, some notable discrepancies emerged between my replication and the original paper.

One of the most apparent differences lies in the behavior of random networks under error versus attack. According to Réka Albert and Barabási (2000), there is little to no difference between the two scenarios, particularly in terms of the p-diameter and the fragmentation metrics S and $\langle s \rangle$. My results, however, show a clear divergence, while the p-diameter increases slowly under random errors, it increases rapid under targeted attacks. Similarly, the threshold for fragmentation is significantly lower in the attack scenario. My results are also supported by Crucitti et al. (2004). This lack of divergence in Réka Albert and Barabási (2000) is somewhat counterintuitive. If nodes are removed at random in one case and the highest-degree nodes are removed in another, one would expect the network to respond differently—unless all nodes have approximately the same degree and the graph therefore being regular. This condition is possible in certain random graphs. However, given the entire space of Erdős–Rényi graphs with n nodes and e edges, the probability of obtaining such a perfectly regular graph is exceedingly low. It is plausible that Réka Albert and Barabási (2000) either used—or inadvertently obtained—random graphs with unusually homogeneous degree distributions. Unfortunately, without access to their experiment code, it is impossible to verify this hypothesis. This lack of transparency limits the reproducibility of their results and may explain the discrepancy.

Another contributing factor to the differences between my results and the original study may lie in the network generation methods. As described in Section 2, I used a two-phase model to generate sparse yet connected random graphs, which is a departure from the traditional Erdős–Rényi approach. Similarly, for the scale-free model, I generated graphs with a fixed number of nodes and then adjusted the edge count to match the target. These modifications may have introduced minor structural variations, although I verified their overall similarity through comparative metrics. Moreover, while the original authors reported that their findings held for graphs of varying sizes, they did not specify the extent of replication or whether their results were averaged over multiple trials. In my case, each data point represents an average over 30 independent simulations. Additionally, due to computational constraints, particularly concerning the p-diameter calculations, I limited my main simulations to networks with $1,000$ nodes and $2,000$ edges, compared to the $10,000$-node networks used for fragmentation metrics.

An interesting methodological detail in Réka Albert and Barabási (2000) is their definition of network diameter. As mentioned in Section 2, they define the diameter as the average shortest path length between all node pairs in the largest connected component, which I refer to as the p-diameter. This contrasts with the standard graph-theoretical definition of the diameter as the maximum shortest path between any two nodes in the graph Pennycuff and Weninger (2015). While the average shortest path length is a well-established metric in network science, it is not synonymous with the diameter. Average path lengths are more robust to small perturbations, whereas the diameter is highly sensitive to outliers. To explore this further, I plotted both the p-diameter and the traditional diameter for each type of graph under random and targeted node removal (see Figure 3).

These findings suggest that while both metrics capture aspects of network degradation, the p-diameter is more stable and may better reflect overall communication efficiency, especially as networks fragment.

References

Barabási, A.-L. and R. Albert (1999). 'Emergence of scaling in random networks'. In: *science* 286.5439, pp. 509–512.

Crucitti, P., V. Latora, M. Marchiori and A. Rapisarda (2004). 'Error and attack tolerance of complex networks'. In: *Physica A: Statistical mechanics and its applications* 340.1-3, pp. 388–394.

Erdős, P. and A. Rényi (1959). 'On random graphs I'. In: *Publicationes Mathematicae Debrecen* 6.290-297, p. 18.

Pennycuff, C. and T. Weninger (2015). 'Fast, exact graph diameter computation with vertex programming'. In: *1st High Performance Graph Mining workshop, Sydney, 10 August 2015*. Barcelona Supercomputing Center.

Réka Albert, H. J. and A.-L. Barabási (2000). 'Error and attack tolerance of complex networks'. In: *nature* 406.6794, pp. 378–382.

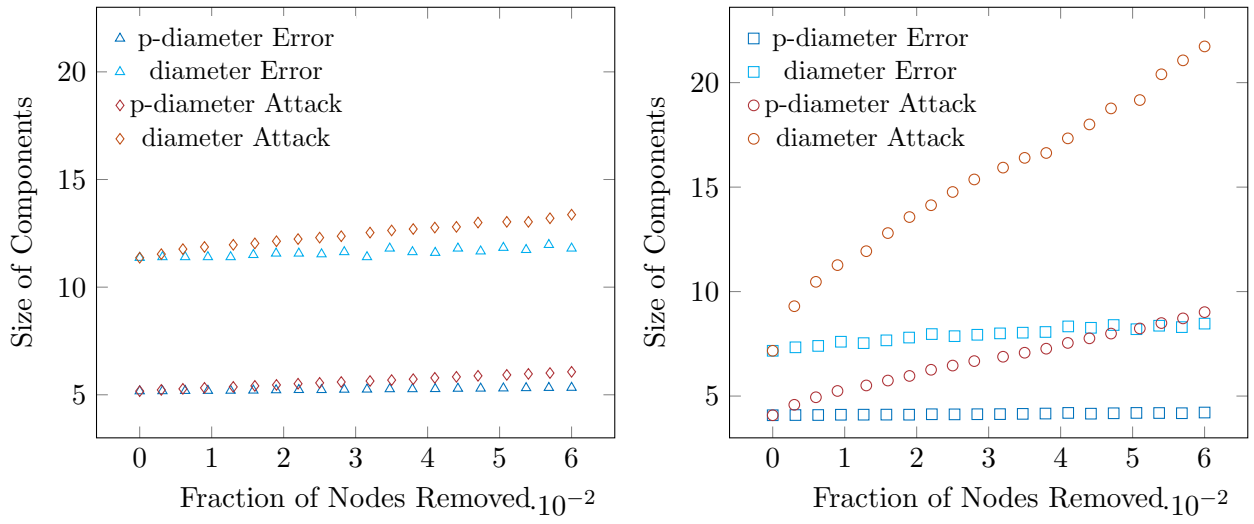
Appendix

Metric	Erdős–Rényi Model	Two-Phase Model	Ratio
Average Degree	5.085	5.080	1.001
Average Shortest Path Length	2.991	2.994	0.999
Average Clustering Coefficient	0.049	0.050	0.971
Average Diameter	5.867	6.080	0.965
Edge Density	0.051	0.051	1.001
Spectral Gap	1.840	1.843	0.998
Shannon Entropy	3.011	3.045	0.989
Graph Entropy	6.487	6.486	1.000

Table 1: Comparison of characteristics of graphs between the ER Graph Model and the Two Phase Model. The graphs had 100 nodes and on average 254 edges.

Metric	Barabási-Albert-Model	modified Barabási-Albert-Model	Ratio
Average Degree	3.992	4.000	0.998
Average Shortest Path Length	4.085	4.081	1.001
Average Clustering Coefficient	0.027	0.026	1.019
Average Diameter	7.220	7.240	0.997
Edge Density	0.004	0.004	0.998
Spectral Gap	2.794	2.822	0.990
Shannon Entropy	2.461	2.473	0.995
Graph Entropy	9.801	9.801	1.000

Table 2: Comparison of characteristics of graphs between the Barabási-Albert-Model and the modified Barabási-Albert-Model. The graphs had 1000 nodes and 1996 or 2000 edges.



(a) Random Graph

(b) Scale-Free Graph

Figure 3: Fragmentation processes of complex graphs.