mathematics

MDPI

Article

Quantifying the Robustness of Complex Networks with Heterogeneous Nodes

Prasan Ratnayake [1], Sugandima Weragoda [1], Janaka Wansapura [1], Dharshana Kasthurirathna [2] and Mahendra Piraveenan [3,*]

[1] Department of Physics, Faculty of Science, University of Colombo, Colombo 00700, Sri Lanka; pshanarat@gmail.com (P.R.); sugeeweragoda@gmail.com (S.W.); janaka.wansapura@phys.cmb.ac.lk (J.W.)
[2] Faculty of Computing, Sri Lanka Institute of Information Technology, B263, Malabe 10115, Sri Lanka; dharshana.k@sliit.lk
[3] Complex Systems Research Group, Faculty of Engineering, University of Sydney, Camperdown, NSW 2006, Australia
* Correspondence: mahendrarajah.piraveenan@sydney.edu.au

Abstract: The robustness of a complex network measures its ability to withstand random or targeted attacks. Most network robustness measures operate under the assumption that the nodes in a network are homogeneous and abstract. However, most real-world networks consist of nodes that are heterogeneous in nature. In this work, we propose a robustness measure called fitness-incorporated average network efficiency, that attempts to capture the heterogeneity of nodes using the 'fitness' of nodes in measuring the robustness of a network. Further, we adopt the same measure to compare the robustness of networks with heterogeneous nodes under varying topologies, such as the scale-free topology or the Erdős–Rényi random topology. We apply the proposed robustness measure using a wireless sensor network simulator to show that it can be effectively used to measure the robustness of a network using a topological approach. We also apply the proposed robustness measure to two real-world networks; namely the CO_2 exchange network and an air traffic network. We conclude that with the proposed measure, not only the topological structure, but also the fitness function and the fitness distribution among nodes, should be considered in evaluating the robustness of a complex network.

Keywords: complex networks; network robustness; network efficiency; node heterogeneity

Citation: Ratayake, P.; Weragoda, S.; Wansapura, J.; Kasthurirathna, D.; Piraveenan, M. Quantifying the Robustness of Complex Networks with Heterogeneous Nodes. *Mathematics* **2021**, *9*, 2769. https://doi.org/10.3390/math9212769

Academic Editor: Nikita Frolov

Received: 14 August 2021
Accepted: 27 October 2021
Published: 1 November 2021

1. Introduction

Most real-world networks exhibit self-organizing and emergent behavior. Further, they possess non-trivial complex topological features [1–3]. Such networks can be observed in domains ranging from social networks, neural networks and collaboration networks [3–5].

Network science attempts to model the structure and function of complex networks. Under network analysis, numerous structures properties such as the centrality, assortativity and robustness can be measured [3,6]. Further, the dynamic growth models such as the Barabasi-Alert model attempt to model the growth and evolution of complex networks [7,8].

One of the key functional and structural characteristics that is widely studied is the robustness of a complex network [3,9]. Robustness measures a network's ability to withstand random and targeted attacks or failures. Numerous robustness measures have been proposed in the literature that attempt to quantify a network's ability to withstand attacks or failures. Accurately quantifying the robustness of a network may be critical in understanding its structure. Further, depending on the application of a network, understanding its robustness may be useful in avoiding failures. For instance, in software networks [10], the robustness measure may help to compare multiple software systems in their ability to withstand failures. Transport networks like air traffic networks [11–17] may be another example where the quantification of the robustness could be applicable.

However, most of the existing measures make the assumption that the nodes in the network are homogeneous and abstract entities. While this ensures that the topological effect on the robustness is captured, most real-world networks may have nodes with heterogeneous characteristics. For instance, individuals in a social network may vary in their cognitive capacity, wealth or creativity, which may not be apparent from their social structure [18]. Therefore, it may be important to consider these heterogeneous and node-specific features, in order to quantify the robustness in a holistic and accurate manner.

In this paper, we present a robustness metric that could incorporate the heterogeneity of nodes. We extend the average network efficiency measure using a generic node fitness function to capture the heterogeneous characteristics of nodes. We then apply the proposed metric on a simulated wireless sensor network to test whether it adequately captures the heterogeneity of nodes in quantifying the network robustness. We apply the proposed measure on real-world networks to further test their applicability. The proposed measure may be used to accurately measure the robustness of a network, particularly if the nodes are heterogeneous in nature.

This paper is organized as follows. Section 2 provides the background on the complex networks and robustness of networks. Section 3 elaborates on the proposed measure and explains the experiments conducted to test the proposed measure both on synthetic and real-world networks. In Section 4, we present the results, followed by a discussion of the results including some directions for potential future work.

2. Background

Networks offer a generalized and a scalable method to model complex systems across multiple domains. Complex networks are graphs consisting of nodes and links with non-trivial topological features [3,7]. The study of the structure and function of complex networks has gained momentum in the recent past due to its wide applicability in numerous fields such as biology, social sciences and supply chains [3,19–21].

The topological structure of complex networks is one of the most important and widely studied characteristics of networks. The two most widely studies topologies are scale-free networks and Erdős–Rényi random networks [6,8]. These topologies are characterized by the degree distributions of a network. A scale-free network is a network whose degree distribution is a power law distribution, at least asymptotically. That is, the fraction $P(k)$ of nodes in the network having k connections to other nodes, in such a manner where $P(k) \sim k^{-\gamma}$. γ is the scale-free parameter and it has been observed that for most real-world networks $2 < \gamma < 3$ [3].

Preferential attachment and growth model have been conjectured as the mechanisms to explain the scale-free topology in real-world network. Preferential attachment suggests that more connected a node is, the more likely it is to receive new links [8]. Erdős–Rényi Random graphs on the other hand, could be represented by a binomial degree distribution, where the degree distribution may fit a Poisson distribution as the network grows in size [8].

However, recent studies suggest that fitness-based generative models that treat a node's ability to attract links as being dependent on node fitness, with fitness being a more abstract and general concept than degree, may be better able to describe the emergence of scale-free behaviour in real-world networks [22,23]. While fitness may be represented by topological attributes of a node, it may also represent topologically independent 'inherent' features of a node, or even a combination of the two [22,24].

The Lognormal Fitness Attachment (LNFA) model proposed in [25] is a fitness based generative model that considers the node's ability to attract links to be purely fitness based. In this model, the fitness Φ_i representing the ability of node i to attract links formed multiplicatively from a number of factors $\{\Phi_1, \Phi_2, \ldots, \Phi_L\}$ as follows [22]:

$$\Phi_i = \prod_{l=1}^{L} \Phi_l \tag{1}$$

Here, each factor Φ_l is a real non-negative value. The number of such factors is assumed to be reasonably large and statistically independent of each other. Under such conditions the fitness distribution will be a lognormal distribution of the form [22]

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma x} e^{\frac{(\log_e x - \mu)^2}{2\sigma^2}} \tag{2}$$

This lognormal distribution has parameters μ and σ where the associated normal distribution $f(y)$ with $y = \log_e x$ has a mean μ and a standard deviation σ. The range of the lognormal distribution is $x \in (0, \infty)$. Further, it can be assumed without loss of generality that $\mu = 0$. Then, by varying the σ values it is possible to generate power-law to winner-take-all [22].

Note that the robustness measure suggested here is equally applicable to planar networks and non-planar networks. Planar networks [26] are a widely studied sub-domain of networks with the key property that they have links that do not intersect. This feature is significant in certain domains, such as transport analysis, where the links have a physical interpretation (roads or railways) and must lie on a plane, and thus their layout is important. In networks from many other domains however (such as biological networks, online social networks and shareholding networks), the links or edges do not have a physical form and represent a conceptual relationship. In yet other domains (such as the Internet), the links do have a physical interpretation, but do not lay on a plane, and their layout thus is immaterial. The robustness measure suggested here is applicable to all networks, and focuses on topological characteristics that are primarily encapsulated by the relevant degree distributions, rather than spatial layout. Therefore, while the suggested measure is applicable to planar networks, it is not specifically designed to analyze the robustness of networks in which the spatial layout is important.

Robustness is a key characteristic of a network that is useful in determining its ability to withstand attacks, failures or perturbations [27]. The attacks on a network may be random or targeted, where a node or link may be removed based on a topological property, such as degree. Networks such as the world wide web, metabolic networks and most of the communication networks are potentially subject to unrealistically high failure rates. Most of these attacks are unpredictable random attacks. Robustness against random decay therefore is very relevant in monitoring the evolution and the existence of a given network. On the other hand, robustness may help to quantify the ability of a network with a negative utility, such as infectious disease networks, to withstand attacks aimed to dismantle them.

Most of the network failures can be described as disconnecting of components from the network. This will disorganize the network and may sometimes cause the network to collapse into sub-networks leading to malfunctions and disruptions. However, it has been observed that, while most of the graphs collapse under random attacks or failures, scale-free networks tend to shrink and demonstrate higher level of robustness [27]. Further, such behavior can be observed in some naturally occurring networks such as metabolic systems. As metabolic systems tend to stabilize under regularly occurring drastic environmental interventions [27].

Where the analysis of network robustness is concerned, the current studies consider nodes to be abstract homogeneous entities [27–29]. Based on that premise, it has already been established that the network topology plays an important role in determining network resilience in multiple domains expanding from social networks to economic networks [27,30]. For instance, it has been shown that the scale-free networks are more resilient against random node failures, compared to random networks, but are more vulnerable to targeted attacks [27]. Vulnerability of a network is its inability to withstand random failures of the nodes or links or persistent targeted attacks that may eliminate the nodes or links. If a network has lower robustness, it may be more vulnerable to random failures and targeted attacks.

Numerous attempts have been made to analyze and quantify the robustness of networks. Albert et al. [27] evaluated the error and attack tolerance of complex networks by

removing the nodes from complex networks one by one until all nodes are extracted or a sustained attack, and studied the variation of topological properties in networks under these perturbations.

Topological properties such as the network diameter, size of the largest component and the average size of the rest of the components have been used in profiling the network robustness under sustained attacks, among others [9,11,17,31,32].

However, these measures have largely relied on profiles of quantities, rather than a single robustness measure. Following their work, a plethora of metrics have been proposed to measure the topological robustness of networks as a single quantity. However, these measures typically measure the averaged effects of single node removals, rather than effects of sequential removals, or are too simplistic. For instance, the network efficiency, which is a measure that we employ in this work, is the average of inverted shortest path lengths [33], and it has been used for quantifying the robustness of a network. However, network perturbations have not been explicitly considered for this measure.

Likewise, Dekker and Colbert [31] introduced two concepts of connectivity for a graph which can be used to model network robustness: the node connectivity and link connectivity, which are the smallest number of nodes and links, respectively, whose removal results in a disconnected or single-node graph. The robustness coefficient [32] gives a single measure by obtaining the ratio between the area under the curve (AUC) of the plot showing the variation of the size of the largest connected component under sustained attacks and the area under the ideal curve where the largest connected component degrades in a linear fashion. In this work, we adopt the network efficiency in a similar approach to obtain a single measure of robustness.

While the existing generalized robustness measures assume that the nodes in a network are homogeneous, heterogeneity of nodes have been addressed in network optimization [34]. It has been observed that when designing an optimal network with maximum efficiency, the network becomes heterogeneous. Thus, network efficiency may be vital in determining the robustness of a network against failures or attacks [34], particularly when it consists of heterogeneous nodes.

The existing literature discusses a particular example where failure in node heterogeneity with respect to power grids [35]. However, this model may only be applicable for power grids and may not be generalized easily. Further, there have been simulations on the dynamic nature of systems illustrating on the crashing of the network from a probabilistic point of view in [35]. Still, such an approach fails to incorporate the heterogeneity of node attributes. Musmeci et al. [36] proposed a fitness model that is used on network subsets of a synthetic network, without considering the significance of node heterogeneity.

Furthermore, our work has focused on wireless sensor networks where the model could be implemented to distinguish the most prominent nodes. Guidoni et al. [37] proposed a resilience-optimized wireless sensor network as a heterogeneous network, which inspired us to consider wireless sensor networks as a possible application to validate a robustness metric that incorporates the heterogeneity of nodes.

The notion of an abstract fitness function to evaluate the optimality of agents is used in heuristic-based and population-based optimization algorithms such as genetic algorithms [38]. In this work, we adopt a similar approach to model and quantify the heterogeneity of nodes.

Based on our survey, incorporating node heterogeneity in a generic robustness metric has not been addressed in the existing work. This paper extends the current standard methods in robustness analysis to include node heterogeneity. By utilizing the proposed metric, we further investigate the effect of node fitness, and thereby the node heterogeneity, on the robustness of complex networks with varying topological features.

3. Methodology

3.1. Fitness-Incorporated Average Network Efficiency

The average efficiency proposed in [39] is a measure of network resilience. In this model, a network is represented as a generic weighted graph G described by two matrices to represent whether two nodes are connected and if so, to quantify the distance between each pair of nodes. These two matrices are the adjacency matrix $\{a_{ij}\}$ and the matrix $\{l_{ij}\}$. Here, l_{ij} can be the space distance between the two vertices i and j or the strength of their possible interaction. The shortest path length d_{ij} between the two vertices i and j is the smallest sum of the distances between them throughout all possible paths in the graph. The efficiency ε_{ij} with respect to the vertices i and j is defined as being inversely proportional to the shortest distance between them: $\varepsilon_{ij} = \frac{1}{d_{ij}} \forall i, j$. When there is no path defined between i and j, $d_{ij} = +\infty$. Extending from this definition, the average efficiency of a network G can be defined as [39]

$$E(G) = \frac{1}{N(N-1)} \sum_{i \neq j \in G} \varepsilon_{ij} \tag{3}$$

where N denotes the number of nodes in the network.

We define the fitness-incorporated average efficiency for directed networks, as a metric of network robustness. Thus, the undirected networks could be transformed to bidirectional directed networks, prior to applying this measure. In the formulation of the fitness-incorporated average efficiency, the matrix $\{l_{ij}\}$ of average efficiency derivation is replaced by the matrix of edge weights $\{w_{ij}\}$. In a directed network each node has outgoing links and incoming links. The node heterogeneity is transformed into each outgoing link weight w_{ij} of node i with a neighbor j as follows:

$$w_{ij} = k\big(f(A_i), g(A_{ij})\big); i \in M, j \in K_i \tag{4}$$

Here, A_i is the node attribute vector node i and $A_{ij}$is the link attribute vector of link j. A_i holds the attribute values of a node such as node size, capacity or age. A_{ij} holds the attribute values of a link such as link length, capacity or strength. Here, M is the set of nodes of the network, while K_i is the set of outgoing neighbors of a particular node i. The function f and g are the node attribute function and the link attribute function, respectively. These functions map the corresponding attribute values into the respective fitness values. The function k is used to derive a link weight that captures the cumulative node fitness and link fitness but projecting the node fitness to the outgoing links and combining that with the original link weights. The weighted shortest path length d_{ij} between two generic nodes i and j is defined as

$$d_{ij} = \sum_{kl \in D_{ij}} w_{kl} \tag{5}$$

Here, kl is a link along the shortest path D_{ij} from i to j and w_{kl} is the weight of that link. Then, the fitness-incorporated efficiency can be defined as $\varepsilon_{ij} = \frac{1}{d_{ij}}$. We define the fitness-incorporated average efficiency E_F of a network G as

$$E_F(G) = \frac{1}{N(N-1)} \sum_{i \neq j \in G} \varepsilon_{ij} \tag{6}$$

Thus, E_F can be regarded as a measure of network robustness that encapsulates the node fitness and thereby the node heterogeneity. We define it for directed networks so that the link weights of all outgoing links along each directed shortest path are accumulated. Thus, the weight of a link connecting two nodes, that is obtained by combining both the

original link weight and the derived weight from node fitness, would vary depending on the link direction.

In order to obtain a measure the robustness of a network under sequential and continuous attacks, we plot the fitness-incorporated average efficiency over the number of nodes removed, under random or targeted failures. Then we compute the area under the curve (AUC) to obtain a single measure of the robustness. Figure 1 depicts a sample plot of the variation of the fitness-incorporated average efficiency.

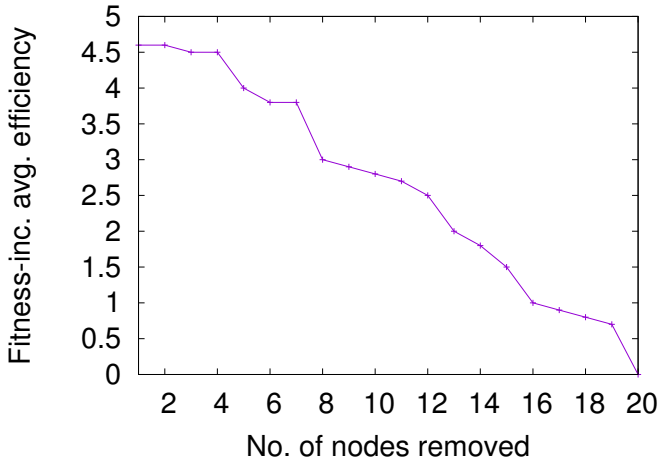


Figure 1. Sample plot showing the variation of the fitness-incorporated average efficiency.

Similarly, in order to quantify the impact on the robustness of the removal of a particular node or a link, we measure the percentage of reduction of fitness-incorporated average efficiency due to the removal of that node or link. We can rank a node or a link in terms of its relative importance using this method.

In the following sub-sections, we use the proposed metric to observe the effect of topology and fitness on the robustness of scale-free and random networks.

3.2. Effect of Network Topology on Fitness-Incorporated Average Efficiency

In this section, we examined the effects of topology on the fitness-incorporated robustness by comparing the scale-free topology and Erdős–Rényi random network topology. Each network consisted of 100 nodes with 130 links and 300 links, respectively, and the node fitness values were assigned based on an equivalent lognormal distribution. The analysis is performed on fifty networks of each topology and the average values noted.

In this experiment, we assumed that the network has uniform link weights and the node fitness values are divided equally to each outgoing link from each node. The cumulative link weights to derive the fitness-incorporated average efficiency is obtained by adding the original link weight and the link weight derived by propagating node fitness to the outgoing links of each node.

In other words, based on the notation given in the Equation (4), $f(A_i)$ is the node fitness value assigned using a lognormal distribution, $g(A_{ij}) = 1$ as links have uniform weights. The node fitness propagated to each of the outgoing link weights is $f(A_i)/d_i$, where d_i is the degree of node i and $j \in K_i$ where K_i is the neighborhood of node i. Further, $k\big(f(A_i), g\big(A_{ij}\big)\big)$ was defined as $f(A_i)/d_i + g(A_{ij})$.

First, we investigated the effect of node heterogeneity on node robustness. We ranked the five most prominent nodes of a scale-free network in terms of node efficiency under both heterogeneous and homogeneous node configurations. In order to rank the nodes in terms of their effect on the robustness of the network, we used the percentage reduction of the robustness upon their removal.

Once the effect of heterogeneity is thus established, we performed robustness analysis under random and targeted attacks on the selected networks. Under random failures,

we removed nodes randomly from the networks and with respect to targeted attacks, we ranked the nodes on order of prominence based on topological measures such as degree or node fitness or a combination of a topological measure and fitness. The network is subjected to the sustained removal of nodes and the E_F calculated after each node removal, until the network is completely dismantled. The plot of E_F against the number of nodes removed is plotted and the area under the resulting curve (AUC) is taken as the quantitative singular measure of robustness. It has to be noted that this approach of computing the AUC of a robustness profile has been used previously to obtain a singular value for the robustness of a network [40]. We then compare the results obtained using the fitness-incorporated efficiency measure of both topological structures against the network efficiency without considering the node fitness, under both random and targeted attacks.

3.3. Applying the Fitness-Incorporated Network Efficiency to Simulated Wireless Sensor Networks

We applied the proposed robustness measure on a simulated wireless sensor network, in order to validate its effectiveness. For this purpose, we used two simulated wireless sensor networks with a random and a scale-free topology. The two networks consist of 100 nodes and 1630 links and 46 nodes and 41 links, respectively. The scale-free network size had to be limited due to the limitations in the simulation environment. Fitness of the nodes in both networks were assigned from a lognormal distribution. All simulations were done using the NS3 wireless network simulator [41]. The NS3 simulator is a tool that's widely used to simulate the network communication of wireless sensor networks that use wireless communication models to emulate the signal transmission and routing in a wireless environment. As there's no direct topological arrangement in a wireless sensor network, the position and a threshold radius were assign around each node to infer a topological arrangement.

The simulated wireless sensor networks were designed to have a predefined amount of power. Thereafter, we assigned variable amounts of packets to each node to simulate varying level of fitness, where node fitness was assumed to be the number of data packets stored, to be transmitted, in a given node.

Applying the notation given in the Equation (4), $f(A_i)$ is the node fitness value assigned, which is the number of packets stored in each node, and $g(A_{ij}) = 1$ as links were assumed to have uniform weights, originally. The node fitness of each node propagated as link weights into outgoing links would be $f(A_i j)/d$. Here, d_i is the degree of node i and $j \in K_i$ where K_i is the neighborhood of node i. Further, $k\big(f(A_i), g\big(A_{ij}\big)\big)$ was defined as $f(A_i)/d_i + g(A_{ij})$.

First, we assigned the data packets to all of the nodes randomly and then made the nodes send their respective data packets from each node to the entire network. Thereafter, we measured the number of packets received by the network from this node. This process is applied for all of the nodes and the nodes are ranked based on the number of packets transmitted by them and receive by the other nodes in the network. As the number of packets that each node can transmit is distributed randomly, we consider this scenario to have the node fitness being uncorrelated to the node degree.

In the second arrangement, we assign the number of data packets stored in each node or fitness of each node such that the node fitness would be correlated to its degree. Then, we used the same ranking technique mentioned above to rank the nodes with respect to their prominence.

To compare the simulator generated ranking with the rankings obtained using the average efficiency and fitness-incorporated average efficiency, we rank the nodes based on the percentage reduction of the robustness calculated using these measures, upon the removal of each node.

By ranking the nodes with respect to the packets they generated that were successfully received, we can order nodes based on the effect they have on the network, using the data transmission model used in the simulator. Then, we can compare those rankings against the node rankings obtained using the proposed fitness-incorporated average efficiency,

in order to check how much the rankings correlate with each other. If the rankings correlate better when fitness values are considered, that is, when using the proposed metric against the existing average efficiency metric, that may justify the usage of the proposed fitness-incorporated average efficiency as a means of quantifying robustness in a network with heterogeneous nodes.

3.4. Applying the Fitness-Incorporated Network Efficiency to Real-World Networks

In order to further investigate the utility of the proposed robustness measure, we applied it on two real-world networks and analyzed the results. Unlike simulated networks, for real-world networks, we derive the fitness from the actual node attributes. The CO_2 exchange network [42] quantifies the transfer of CO_2 across countries in terms of goods and services. It is essentially a supply chain network where the commodity CO_2 is embedded with the products and services exchanged among countries. The countries are modeled as nodes with connecting links provided that there are CO_2 transfers between them. A link is weighted by a measure of how much of ones country's production emission corresponds to the connecting country's consumption emission. The size of the network, N, is 112. We define the node fitness as the CO_2 produced and consumed within the same country.

Despite being designed systems, air traffic can be modeled as complex self-organizing networks [43]. The air traffic network that we consider is an undirected weighted network as obtained by considering the 500 US airports with the largest amount of traffic from publicly available data. Nodes represent US airports and edges represent air travel connections among them. The dataset contains an anonymized list of connected pairs of nodes and the weight associated to the edge, expressed in terms of number of available seats on the given connection on a yearly basis. Therefore, fitness of one such airport was taken as the cumulative number of seats corresponding to a given airport.

In each of the above scenarios, the node fitness propagated to outgoing link weight were calculated by $f(A_i)/d_i$ where $f(A_i)$ is the node fitness in the respective scenarios and d_i is the degree of node i. Similar to the previous experiments, $k\big(f(A_i), g\big(A_{ij}\big)\big)$ was defined as the additive function $f(A_i)/d_i + g(A_{ij})$.

4. Results and Discussion

4.1. Effect of Network Topology and Node Fitness on Robustness Analysis

In Table 1, we present the results of the ranking of nodes in the scale-free network considered, under both homogeneous and heterogeneous node configurations, with more prominent nodes considered as being the ones that are more influential to the network robustness. We compare the ranking of nodes when the average efficiency, where node fitness is not considered, with the fitness-incorporated average efficiency based rankings.

Table 1. Rankings of nodes based on each robustness method in the scale-free network with 100 nodes and 130 links.

Rank	Node ID	
	Avg. Efficiency	**Fitness-Incorporated Avg. Efficiency**
1	29	85
2	74	54
3	96	89
4	97	36
5	41	20

These results indicate that the prominence and the robustness of nodes differ under heterogeneous and homogeneous node configurations.

In order to compare the robustness using the fitness-incorporated average efficiency and average efficiency, we subject the synthetic Erdős–Rényi random network and scale-free network to multiple attacks and failures. In each attack, nodes are removed sequentially based on a selection criteria. The selection criteria include random failures, degree-based attacks, betweenness-based attacks, fitness-based attacks, fitness and degree-based attacks, and fitness and betweenness-based attacks. In each timestep, the node with the highest centrality value is selected to be removed until the entire network is dismantled. The variation of the robustness is plotted against the timestep and the area under the curve is calculated to obtain an aggregated robustness value based on the robustness measure in concern. The two robustness columns represent the area under the curve of the variation of the efficiency metric, under sustained attacks. The fitness-incorporate robustness uses the fitness-incorporated average efficiency and the topology-based robustness is measured using the average efficiency, without considering node fitness.

Tables 2 and 3 present the results of the resilience analysis on the simulated random and scale-free networks, respectively.

Table 2. Robustness analysis of the Erdős–Rényi random network topology with 100 nodes and 300 links, under different attacks.

Node Removal Criteria	Fitness-Incorporated Robustness (Units)	Topology-Based Robustness (Units)
Random failures	875.1	29.2
Degree Centrality	8641.4	11.5
Betweeness Centrality	7381.1	11.3
Fitness	29,127.7	27.8
Fitness + Degree centrality	8805.5	14.2
Fitness + Between. centrality	9707.6	13.8

Table 3. Robustness analysis of the scale-free network topology with 100 nodes and 130 links, under different attacks.

Node Removal Criteria	Fitness-Incorporated Robustness (Units)	Topology-Based Robustness (Units)
Random failures	363.5	14.9
Degree Centrality	1550.4	0.9
Betweenness Centrality	3321	1.2
Fitness	14,859.7	2.4
Fitness + Degree centrality	2001.3	2.2
Fitness + Between. centrality	3067.9	2.5

The results depict that both random and scale-free topologies depict less robustness against random failures when measuring robustness using the proposed fitness-incorporated average efficiency measure. However, the average efficiency measure gives higher robustness against random failure, in both topologies.

In random topology, random failures seem to affect the robustness most, when using the fitness-incorporated average efficiency, while betweenness based attacks are most effective when the fitness is not considered. On the other hand scale-free networks seem to be least robust against random failures when the node fitness is considered, while it appears to be more vulnerable against degree-based attacks, when node fitness is not considered.

In general, these results depict that incorporating the fitness of the nodes may be critical in analyzing the robustness of a network against a given type of failure or an attack.

4.2. Analysis of Robustness Based Node Rankings

We present the rankings obtained for for the first few nodes in both random and scale-free networks that were simulated using NS3 simulator in Tables 4–7. We compare each network with the robustness-based node rankings made when the node fitness is distributed in a degree-correlated and a random manner. The nodes with highest rankings have the most impact on robustness upon their removal. The results indicate that the fitness-incorporated average efficiency based node rankings may depend on the node fitness distribution in the network.

Table 4. Node ranking comparison in the random topology (presented with Node IDs), based on fitness-incorporated average efficiency, when the node fitness is distributed in a degree-correlated and a random manner.

Node Rank	Degree-Correlated		Randomly Assigned	
	NS3	Fitness-Inc. Avg. Efficiency Based Ranking	NS3	Fitness-Inc. Avg. Efficiency Based Ranking
1	45	45	56	56
2	46	47	57	48
3	47	46	45	58
4	48	48	44	99
5	49	49	43	26

Table 5. Node raking comparison in random topology, based on average efficiency, when the node fitness is distributed in a degree-correlated and a random manner.

Node Rank	Degree-Correlated		Randomly Assigned	
	NS3	Average Efficiency Based Ranking	NS3	Average Efficiency Based Ranking
1	45	45	56	45
2	46	44	57	44
3	47	43	45	43
4	48	42	44	42
5	49	4	43	4

Table 6. Node ranking comparison of scale-free topology, based on fitness-incorporated average efficiency, when the node fitness is distributed in a degree-correlated and a random manner.

Node Rank	Degree-Correlated		Randomly Assigned	
	NS3	Fitness-Inc. Avg. Efficiency Based Ranking	NS3	Fitness-Inc. Avg. Efficiency Based Ranking
1	37	17	17	17
2	34	22	29	29
3	15	29	22	22
4	28	6	36	36
5	44	19	27	27

Table 7. Node ranking comparison in the scale-free topology, based on average efficiency, when the node fitness is distributed in a degree-correlated and a random manner.

Node Rank	Degree-Correlated		Randomly Assigned	
	NS3	Average Efficiency Based Ranking	NS3	Average Efficiency Based Ranking
1	37	17	17	17
2	34	36	29	36
3	15	27	22	27
4	28	29	36	29
5	44	22	27	22

Tables 8 and 9 denote the Pearson correlations of the rankings obtained using different robustness rankings. The NS3 simulator based rankings are used as the ground truth as the robustness rankings based on the NS3 simulator is based on the number of packets received, which in turn is calculated using the wireless communication based signal transmission, error propagation and network quality related models. Each node in the simulator may generate packets that could be received by the other nodes, and the percentage of packets received by the rest of the nodes are used to quantify the robustness of the wireless network in a more concrete manner. The node rankings that were used to obtain these correlation coefficients are shown in the Appendix A.

By comparing the correlation of the rankings obtained using the simulator with the average efficiency based rankings and fitness-incorporated average efficiency based rankings, we can deduce whether incorporating node fitness in network efficiency can provide a better understanding of a node's relative importance in a network, with respect to its impact on the network's robustness. Thus, if the fitness-incorporated average efficiency based rankings correlate more with the simulator based rankings, that may be used as evidence for the effectiveness of using node fitness in measuring network robustness. We do this comparison for both the scenarios where the node fitness is distributed in a manner that correlates with the node degree and where the node fitness is distributed in a manner that does not correlate with the node degree.

Table 8. Ranking correlations for the Random topology, when the node fitness is distributed in a degree-correlated and a random manner.

Correlation Coefficient Type	Degree Correlated	Randomly Assigned
NS3 ranking and fitness inc. avg. efficiency based ranking	0.5587	−0.1933
NS3 ranking and average efficiency based ranking	0.5430	−0.2587
Fitness and degree	0.995	0.2045
Fitness and betweenness	−0.2587	−0.328

Table 9. Ranking correlations for the scale-free topology, when the node fitness is distributed in a degree-correlated and a random manner.

Correlation Coefficient Type	Degree-Correlated	Randomly Assigned
NS3 ranking and fitness inc. avg. efficiency based ranking	0.7636	0.1491
NS3 ranking and average efficiency based ranking	0.5430	−0.03156
Fitness and Degree	0.8882	0.1103
Fitness and Betweenness	0.5770	−0.0583

The results show that both for random and scale-free topologies, the NS3-based ranking correlates more with the fitness-incorporated average efficiency based ranking, in comparison with the average efficiency-based ranking, that does not take the node fitness into account. This is particularly evident with the scale-free topology when the node fitness is distributed in a manner that correlates with the degree. This may be used as some evidence of the effectiveness of the proposed fitness-incorporated average efficiency measure of robustness.

Note that the node rankings obtained using the fitness-incorporated average efficiency depends on the fitness function used to derive node fitness. Nevertheless, incorporating fitness to order nodes based on their relative importance, in terms of their effect on the robustness of a network, may be useful in determining the significance of a node.

4.3. Robustness Analysis of Real-World Networks

Here, we present the results of the robustness analysis of two real-world networks; the CO_2 exchange network and the air traffic network. Table 10 presents the results of the robustness analysis of the CO_2 exchange network. The robustness analysis on the CO_2 exchange network shows that it is more susceptible to random failures when the robustness is measured using the fitness-incorporated average efficiency. Under the average efficiency measure that does not take node fitness into account, degree based attacks prove to be most effective. While these results may depend on the node fitness measure and the fitness function used, it is evident that incorporating node fitness can give a significantly different picture of network robustness in a real-world network.

Table 10. Robustness analysis of the CO_2 exchange network, N = 112.

Attack Type	Robustness Based on Fitness Inc. Avg. Efficiency	Robustness Based on Avg. Efficiency
Random failures	2626.3	79.5
Degree Centrality	5233.5	42.3
Betweenness Centrality	4666.4	54.0
Fitness	8233.4	51.2
Fitness + Degree	6006.8	46.1
Fitness + Betweenness	5211.1	51.3

Table 11 presents the results of the robustness analysis of the air-traffic network. In this network, the degree-based attack appears to be most effective in terms of the effect on the network, when the fitness-incorporated average efficiency is used as the robustness measure. On the other hand, betweenness centrality-based attacks are most damaging when the average efficiency measure is used.

Table 11. Robustness analysis of the air traffic network, N = 500.

Attack Type	Robustness Based on Fitness Inc. Avg. Efficiency	Robustness Based on Avg. Efficiency
Random failures	19,550.1	79.5
Degree Centrality	10,916.2	7.7
Betweenness Centrality	11,234.1	7.4
Fitness	59,883.0	29.8
Fitness + Degree centrality	13,018.1	9.0
Fitness + Betweenness centrality	12,209.1	9.6

The air traffic network is seen to be the most susceptible to attacks targeting the nodes with higher fitness. It is interesting to note that attacks based on a topological measure such as betweenness does not outperform the attacks made by combination of node fitness with a topological measure such as betweenness, suggesting that considering node fitness in combination with a topological centrality may not always be favorable in attacking a network.

Unlike the CO_2 exchange network, the air traffic network appears to be more susceptible to attacks targeting the nodes with the highest fitness value combined with the topological centrality measures, in comparison to attacks based on fitness only. In the air traffic network, the correlation between node degree and node fitness is relatively high, compared to the CO_2 exchange network. This suggests that the fitness of individual nodes may be more relevant to determine network robustness, when the correlation of the node fitness with the topological property is high. In such a scenario, the effect of node fitness on the robustness may decrease.

Another important observation is that the power-law exponent of the degree distribution of the CO_2 exchange network is comparatively high, shifting it towards the realm of random networks (in a typical scale-free network the exponent value ranges between 2 and 3 [43], whereas that of the air traffic network is closer to a typical(scale-free) value.

In observing the resilience of real-world networks with heterogeneous nodes, it is important to note that the robustness that we may observe is subjective to how we define the node fitness. Thus, even the same network may show varying robustness features

under different fitness functions. However, in general, the fitness distribution and the topology both play key roles in the definition of the network robustness. Further, these results suggest that not just the fitness distribution and topology, but also the correlation between the fitness distribution and centrality distribution affects the robustness of the overall network.

While we consider the topological centrality values and node fitness as two separate classes of attributes in the experiments conducted, it is important to note that even the centrality measures can be considered as a special case of node fitness. However, in this work, we consider node fitness as an explicitly non-topologically derived value that is based on the inherent heterogeneity of the nodes.

5. Conclusions and Future Work

In this work, we proposed a node fitness-based approach to capture the node heterogeneity in proposing a novel robustness measure. The results indicate that the fitness function, network topology, fitness distribution and the correlation between the node topological features and node fitness are all relevant in determining the robustness of a network. Further, we use the proposed measure as a means of ranking nodes in a network in terms of their impact on the robustness. The rankings derived from the proposed measure against the rankings obtained from a wireless sensor network based simulation, we could empirically validate the proposed robustness measure. The proposed measure may be useful in analyzing the real-world networks for their robustness, as most real-world networks constitute of heterogeneous nodes.

The synthetic networks and the simulated wireless sensor networks may be limited in their resemblance to real-world networks, which may be considered a limitation. Possible future work may look into the applicability of the proposed measure in real-world networks such as power-grid networks, social networks and financial networks where heterogeneous nodes may be present and the heterogeneity may be defined in a context sensitive manner. The effectiveness of the proposed measure may be further validated by comparing it with the real-world robustness data, where available.

In this work, we predominantly focused on empirical analysis of the proposed metric, rather than focusing on a more analytical approach in explaining it. The main reason for this approach was the fact that robustness is an emergent property where not only network topology, but also the node fitness function and fitness distribution may play a synergistic role in determining it, based on the proposed measure. As possible future work, further analytical analysis on the proposed measure could be conducted by quantifying its properties and limiting conditions such as its expected value, variance and effective range, under different network conditions. Furthermore, the robustness values obtained from the proposed measure can be compared with existing measures and also more failure/robustness data of real-world networks to further validate its validity and applicability.

Another possible extension of this work may be to study how the node heterogeneity incorporated network robustness varies when the degree distribution exponent of a scale-free network is changed. Identifying the scale-free exponent value for which the robustness is maximal, may be helpful in determining the optimal topological structure that would be most robust in the presence of attacks. The knowledge of such an optimal topological structure may be helpful in design considerations of complex engineered systems and in quantifying the robustness of real-world networks.

Author Contributions: Conceptualization, D.K. and M.P.; Investigation, P.R. and S.W.; Methodology, J.W. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

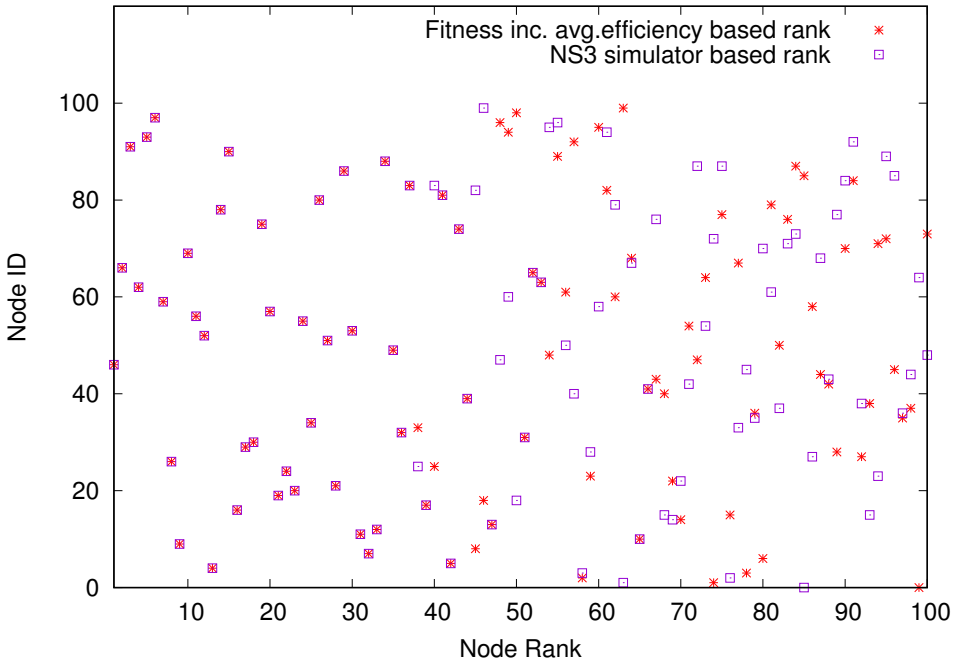
Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

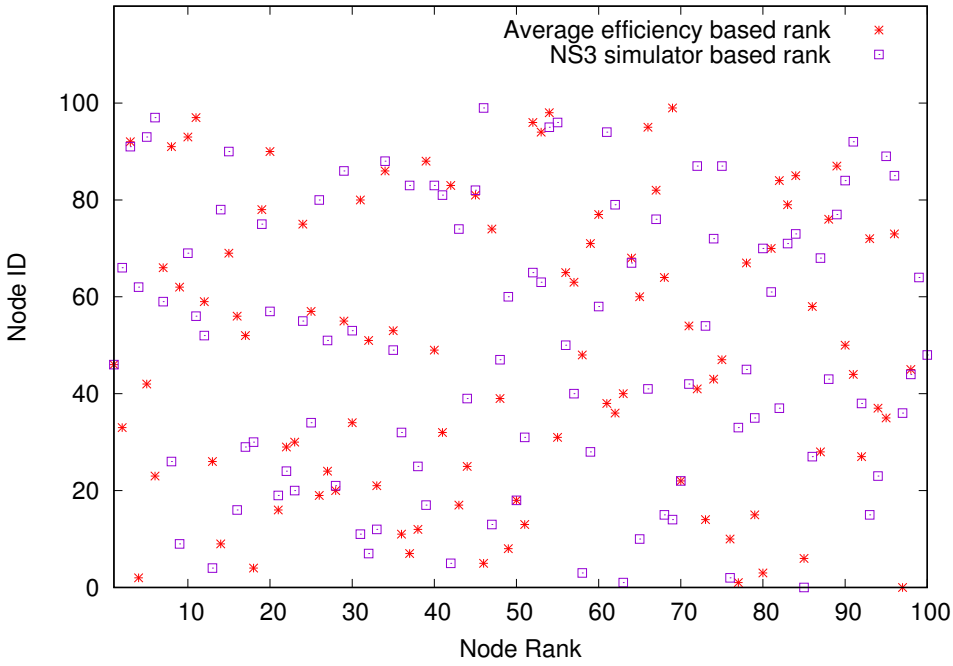
Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

Figures A1 and A2 show the node ranking comparisons based on the random topology. The Pearson correlation coefficients denoted in Table 8 are derived from these rankings.

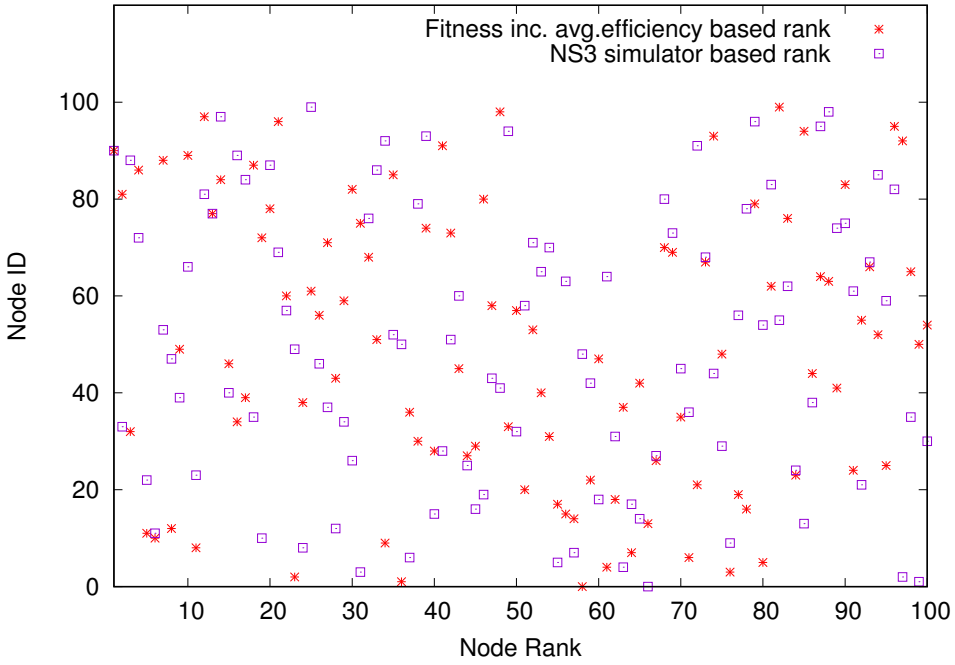


(**a**)

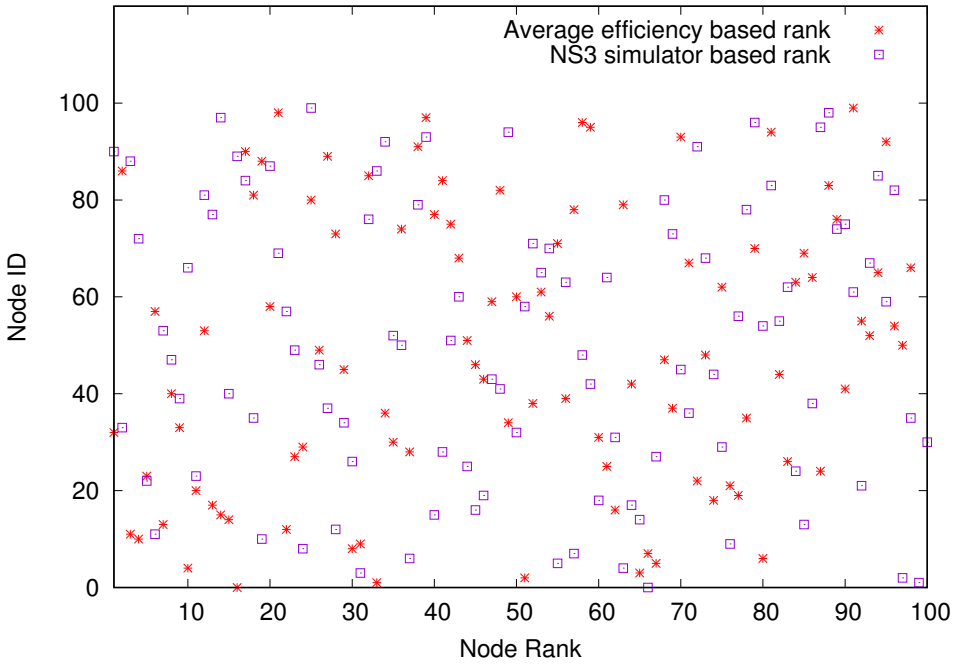


(**b**)

Figure A1. Node ranking comparison for random topology, when the fitness distribution correlates with the degree. (**a**) Fitness incorporated average efficiency-based ranking vs. NS3 simulator ranking. (**b**) Average efficiency-based ranking vs. NS3 simulator-based ranking.



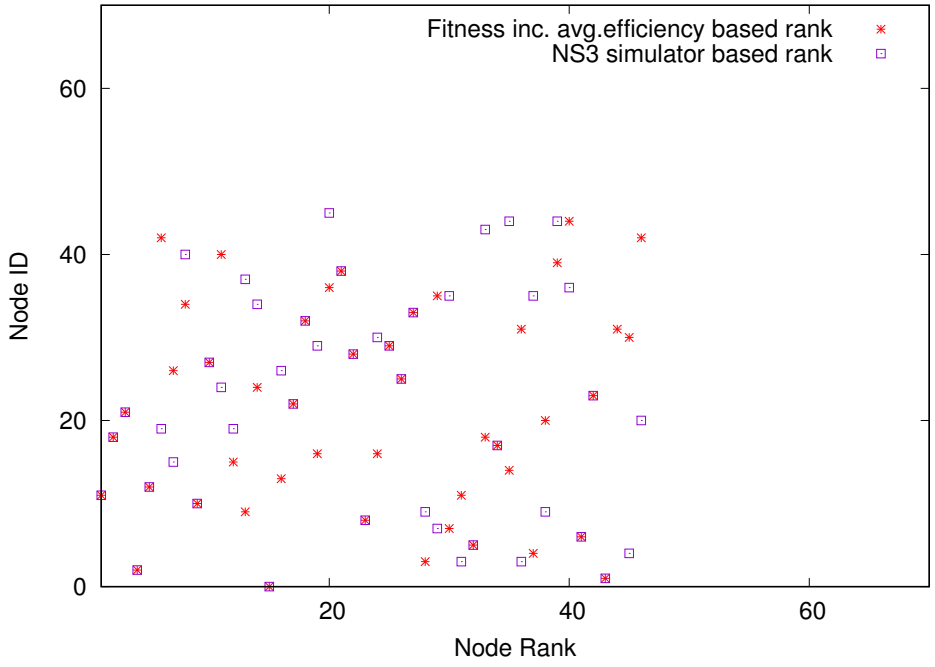
(**a**)



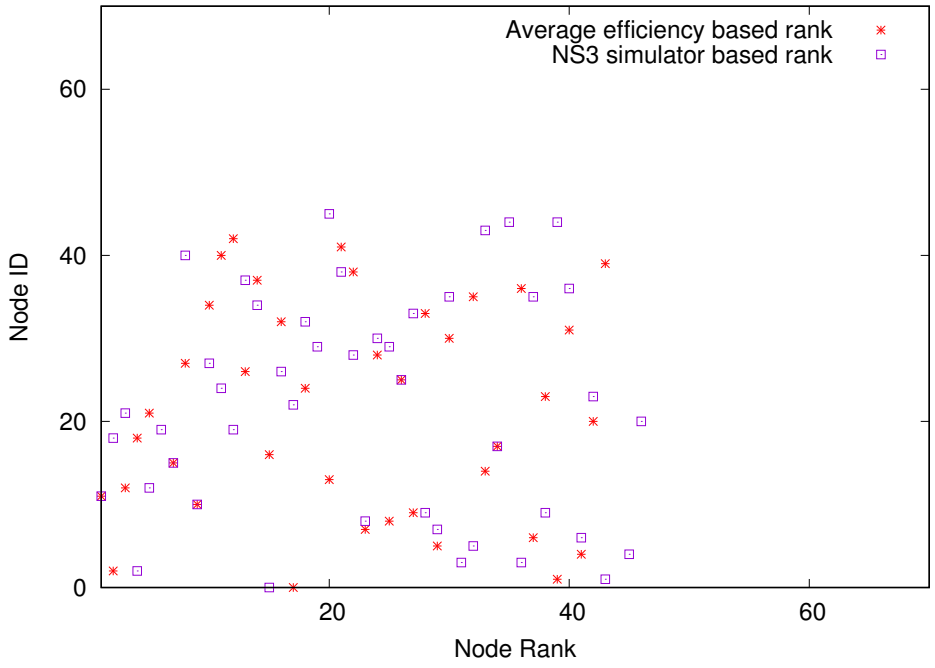
(**b**)

Figure A2. Node ranking comparison for random topology, when the fitness is distributed in a manner that does not correlate with the degree. (**a**) Fitness incorporated average efficiency-based ranking vs. NS3 simulator ranking. (**b**) Average efficiency-based ranking vs. NS3 simulator-based ranking.

Figures A3 and A4 show the node ranking comparisons based on the scale-free topology. The Pearson correlation coefficients denoted in Table 9 are derived from these rankings.

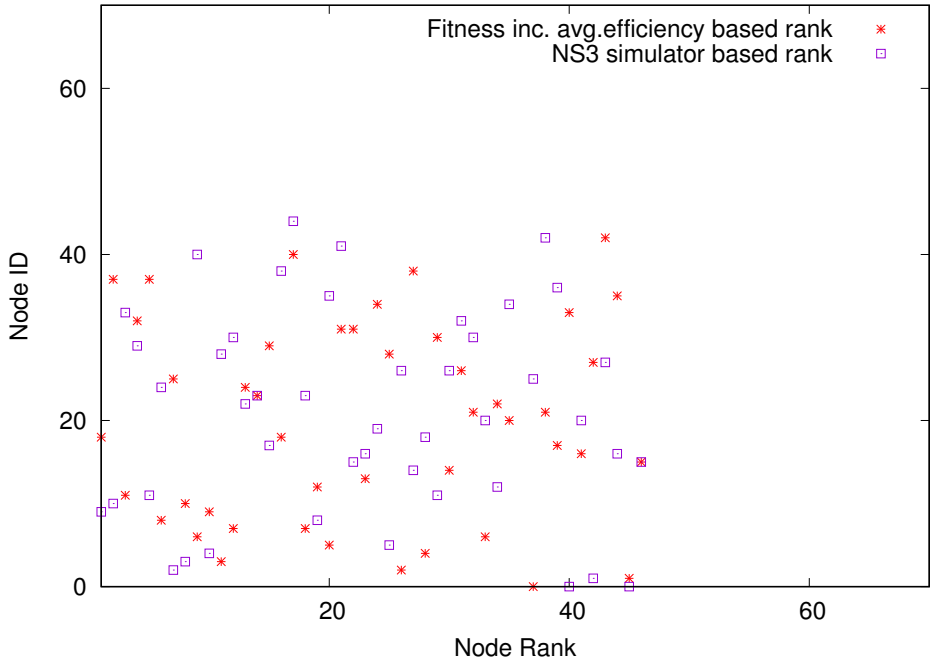


(**a**)

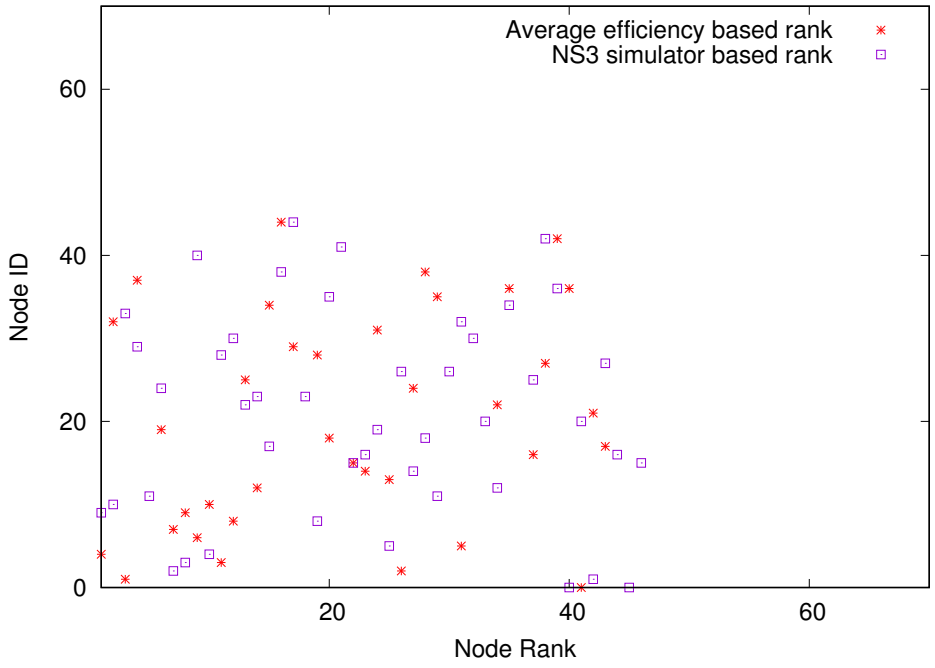


(**b**)

Figure A3. Node ranking comparison for scale-free topology, when the fitness is distributed in a manner that correlates with the degree. (**a**) Fitness incorporated average efficiency-based ranking vs. NS3 simulator ranking. (**b**) Average efficiency-based ranking vs. NS3 simulator-based ranking.



(**a**)



(**b**)

Figure A4. Node ranking comparison for scale-free topology, when the fitness is distributed in a manner that does not correlate with the degree. (**a**) Fitness incorporated average efficiency-based ranking vs. NS3 simulator ranking. (**b**) Average efficiency-based ranking vs. NS3 simulator-based ranking.

References

1. Waissi, G.R. Network Flows: Theory, Algorithms, and Applications. *Interfaces* **1994**, *24*, 133–155.
2. Morris, M. Epidemiology and social networks: Modeling structured diffusion. *Sociol. Methods Res.* **1993**, *22*, 99–126. [CrossRef]
3. Albert, R.; Barabási, A.L. Statistical mechanics of complex networks. *Rev. Mod. Phys.* **2002**, *74*, 47. [CrossRef]
4. Gupta, S.; Anderson, R.M.; May, R.M. Networks of sexual contacts: Implications for the pattern of spread of HIV. *AIDS (Lond. Engl.)* **1989**, *3*, 807–817. [CrossRef]

5. Davis, G.F.; Greve, H.R. Corporate elite networks and governance changes in the 1980s. *Am. J. Sociol.* **1997**, *103*, 1–37. [CrossRef]
6. Newman, M.E. The structure and function of complex networks. *SIAM Rev.* **2003**, *45*, 167–256. [CrossRef]
7. Banks, D.L.; Carley, K.M. Models for network evolution. *J. Math. Sociol.* **1996**, *21*, 173–196. [CrossRef]
8. Barabási, A.L.; Albert, R. Emergence of scaling in random networks. *Science* **1999**, *286*, 509–512. [CrossRef] [PubMed]
9. Callaway, D.S.; Newman, M.E.; Strogatz, S.H.; Watts, D.J. Network robustness and fragility: Percolation on random graphs. *Phys. Rev. Lett.* **2000**, *85*, 5468. [CrossRef] [PubMed]
10. Wen, L.; Kirk, D.; Dromey, R.G. Software systems as complex networks. In Proceedings of the 6th IEEE International Conference on Cognitive Informatics, Lake Tahoe, CA, USA, 6–8 August 2007; pp. 106–115.
11. Perera, S.; Bell, M.G.; Bliemer, M.C. Network science approach to modelling the topology and robustness of supply chain networks: A review and perspective. *Appl. Netw. Sci.* **2017**, *2*, 33. [CrossRef] [PubMed]
12. Wang, J.; Mo, H.; Wang, F.; Jin, F. Exploring the network structure and nodal centrality of China's air transport network: A complex network approach. *J. Transp. Geogr.* **2011**, *19*, 712–721. [CrossRef]
13. Maćkiewicz, A.; Ratajczak, W. Towards a new definition of topological accessibility. *Transp. Res. Part B Methodol.* **1996**, *30*, 47–79. [CrossRef]
14. Haggett, P.; Chorley, R. Network Analysis in Geography. In *Network Analysis in Geography*; Edward Arnold: London, UK, 1969.
15. Garrison, W.L.; Marble, D.F. Factor-analytic study of the connectivity of a transportation network. In *Papers of the Regional Science Association*; Springer: Cham, Switzerland, 1964; Volume 12, pp. 231–238.
16. Verma, T.; Araújo, N.A.; Herrmann, H.J. Revealing the structure of the world airline network. *Sci. Rep.* **2014**, *4*, 1–6. [CrossRef] [PubMed]
17. Kansky, K.J. Structure of Transportation Networks: Relationships between Network Geometry and Regional Characteristics. Ph.D. Thesis, The University of Chicago, Chicago, IL, USA, 1963.
18. Uchino, B.N.; Holt-Lunstad, J.; Smith, T.W.; Bloor, L. Heterogeneity in social networks: A comparison of different models linking relationships to psychological outcomes. *J. Soc. Clin. Psychol.* **2004**, *23*, 123–139. [CrossRef]
19. Gallos, L.K.; Song, C.; Havlin, S.; Makse, H.A. Scaling theory of transport in complex biological networks. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 7746–7751. [CrossRef]
20. Perera, S.S.; Bell, M.G.; Piraveenan, M.; Kasthurirathna, D.; Parhi, M. Topological structure of manufacturing industry supply chain networks. *Complexity* **2018**, *2018*, 3924361. [CrossRef]
21. Piraveenan, M.; Jing, H.; Matous, P.; Todo, Y. Topology of international supply chain networks: A case study using factset revere datasets. *IEEE Access* **2020**, *8*, 154540–154559. [CrossRef]
22. Nguyen, K.; Tran, D.A., Fitness-Based Generative Models for Power-Law Networks. In *Handbook of Optimization in Complex Networks: Theory and Applications*; Thai, M.T., Pardalos, P.M., Eds.; Springer: Boston, MA, USA, 2012; pp. 39–53. [CrossRef]
23. Bell, M.; Perera, S.; Piraveenan, M.; Bliemer, M.; Latty, T.; Reid, C. Network growth models: A behavioural basis for attachment proportional to fitness. *Sci. Rep.* **2017**, *7*, 1–11.
24. Bianconi, G.; Barabási, A.L. Competition and multiscaling in evolving networks. *Europhys. Lett. (EPL)* **2001**, *54*, 436–442. [CrossRef]
25. Chakravartula, S.; Killingback, T.; Sundaram, B.; Tran, D. A statistical construction of power-law networks. *Int. J. Parallel, Emergent Distrib. Syst.* **2010**, *25*, 223–235. [CrossRef]
26. Viana, M.P.; Strano, E.; Bordin, P.; Barthelemy, M. The simplicity of planar networks. *Sci. Rep.* **2013**, *3*, 1–6. [CrossRef]
27. Albert, R.; Jeong, H.; Barabási, A.L. Error and attack tolerance of complex networks. *Nature* **2000**, *406*, 378–382. [CrossRef] [PubMed]
28. Giannoccaro, I.; Albino, V.; Nair, A. Advances on the Resilience of Complex Networks. *Complexity* **2018**, *2018*, 1–3. [CrossRef]
29. Gao, J.; Barzel, B.; Barabasi, A.L. Universal resilience patterns in complex networks. *Nature* **2016**, *530*, 307–312. [CrossRef]
30. Caschili, S.; Reggiani, A.; Medda, F. Resilience and vulnerability of spatial economic networks. *Netw. Spat. Econ.* **2015**, *15*, 205–210. [CrossRef]
31. Dekker, A.H.; Colbert, B.D. Network robustness and graph topology. In Proceedings of the 27th Australasian Conference on Computer Science, Dunedin, New Zealand, 1 January 2004; Volume 26, pp. 359–368.
32. Piraveenan, M.; Thedchanamoorthy, G.; Uddin, S.; Chung, K.S.K. Quantifying topological robustness of networks under sustained targeted attacks. *Soc. Netw. Anal. Min.* **2013**, *3*, 939–952. [CrossRef]
33. Crucitti, P.; Latora, V.; Marchiori, M.; Rapisarda, A. Error and attack tolerance of complex networks. *Phys. A Stat. Mech. Appl.* **2004**, *340*, 388–394. [CrossRef]
34. Wang, B.; Tang, H.; Guo, C.; Xiu, Z.; Zhou, T. Optimization of network structure to random failures. *Phys. A Stat. Mech. Appl.* **2006**, *368*, 607–614. [CrossRef]
35. Li, J.; Wang, J.; Sun, S.; Xia, C. Cascading crashes induced by the individual heterogeneity in complex networks. *Appl. Math. Comput.* **2018**, *323*, 182–92. [CrossRef]
36. Musmeci, N.; Battiston, S.; Caldarelli, G.; Puliga, M.; Gabrielli, A. Bootstrapping Topological Properties and Systemic Risk of Complex Networks Using the Fitness Model. *J. Stat. Phys.* **2013**, *151*. [CrossRef]
37. Guidoni, D.L.; Mini, R.A.; Loureiro, A.A. On the design of resilient heterogeneous wireless sensor networks based on small world concepts. *Comput. Netw.* **2010**, *54*, 1266–1281. [CrossRef]

38. Bielli, M.; Caramia, M.; Carotenuto, P. Genetic algorithms in bus network optimization. *Transp. Res. Part C Emerg. Technol.* **2002**, *10*, 19–34. [CrossRef]
39. Latora, V.; Marchiori, M. Efficient Behavior of Small-World Networks. *Phys. Rev. Lett.* **2001**, *87*, 198701. [CrossRef]
40. Kasthurirathna, D.; Dong, A.; Piraveenan, M.; Tumer, I. The Failure Tolerance of Mechatronic Software Systems to Random and Targeted Attacks. In *International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*; American Society of Mechanical Engineers: New York, NY, USA, 2013; Volume 55928. [CrossRef]
41. Carneiro, G. NS-3: Network simulator 3. In Proceedings of the UTM Lab Meeting, Porto, Portugal, 30 April 2010; Volume 20, pp. 4–5.
42. Davis, S.J.; Peters, G.P.; Caldeira, K. The supply chain of CO_2 emissions. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 18554–18559. [CrossRef] [PubMed]
43. Colizza, V.; Pastor-Satorras, R.; Vespignani, A. Reaction-diffusion processes and metapopulation models in heterogeneous network. *Nat. Phys* **2007**, *3*, 276–282. [CrossRef]