



Article Robustness of Real-World Networks after Weight Thresholding with Strong Link Removal

Jisha Mariyam John ¹, Michele Bellingeri ^{2,3,*}, Divya Sindhu Lekha ¹, Davide Cassi ^{2,3} and Roberto Alfieri ^{2,3}

- ¹ Indian Institute of Information Technology, Kottayam 686635, India;
- jishamariyam.phd201010@iiitkottayam.ac.in (J.M.J.); divyaslekha@iiitkottayam.ac.in (D.S.L.)
- ² Dipartimento di Scienze Matematiche, Fisiche e Informatiche, Università di Parma, via G.P. Usberti, 7/a, 43124 Parma, Italy; davide.cassi@unipr.it (D.C.); roberto.alfieri@unipr.it (R.A.)
- ³ Istituto Nazionale di Fisica Nucleare (INFN) Gruppo Collegato di Parma, 43124 Parma, Italy
- * Correspondence: michele.bellingeri@unipr.it

Abstract: Weight thresholding (*WT*) is a method intended to decrease the number of links within weighted networks that may otherwise be excessively dense for network science applications. *WT* aims to remove links to simplify the network by holding most of the features of the original network. Here, we test the robustness and the efficacy of the node attack strategies on real-world networks subjected to *WT* that remove links of higher weight (strong links). We measure the network robustness along node removal with the largest connected component (*LCC*). We find that the real-world networks under study are generally robust when subjected to *WT*. Nonetheless, *WT* with strong link removal changes the efficacy of the attack strategies and the rank of node centralities. Also, *WT* with strong link removal may trigger a more significant change in the node centrality rank than *WT* by removing weak links. Network science research with the aim to find important/influential nodes in the network has to consider that simplifying the network with *WT* methodologies may change the node centrality.

Keywords: complex networks; network robustness; weight thresholding; link removal; link pruning

MSC: 05C82

1. Introduction

Weight thresholding is a simple technique that aims to reduce the number of edges in weighted networks that are otherwise too dense for applying standard graph-theoretical methods [1]. *WT* is a methodology in sparsification approaches to reduce link density in different real-world networks [2]. *WT* has many real-world applications, such as sparsifying ecological, financial, brain, and biological networks [3–5]. The principal aim of *WT* is to remove links to simplify the networks and make them easier to analyze. Therefore, the *WT* policy should guarantee that the significant traits of the original network are retained intact. In short, the objective of the *WT* procedure is to prune the highest number of links, avoiding drastic alterations in the critical structure of the original real-world network. Unfortunately, many conventional network properties quickly change under the *WT* procedure [1,6].

WT finds applications in research focusing on neural networks (*NNs*) or other machine learning models. In essence, *WT* involves applying a threshold to the weights of links in a *NN*. Links with weights below the threshold are considered less significant and can be eliminated or considered inactive. This process reduces the overall number of connections in the model, making it simpler and often more computationally efficient [7,8].

Network robustness is an essential field of research in network science [9]. Robustness is the property of a system to maintain functioning when perturbed or attacked [10]. The seminal research of Albert et al. [10] explores the error and attack resilience of complex networks of both real-world and model networks. The author investigates how networks



Citation: John, J.M.; Bellingeri, M.; Lekha, D.S.; Cassi, D.; Alfieri, R. Robustness of Real-World Networks after Weight Thresholding with Strong Link Removal. *Mathematics* **2024**, *12*, 1568. https://doi.org/10.3390/ math12101568

Academic Editor: Shugang Li

Received: 25 March 2024 Revised: 6 May 2024 Accepted: 17 May 2024 Published: 17 May 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https:// creativecommons.org/licenses/by/ 4.0/). react against random node removal (error and failures) and the targeted removal of the most connected nodes (attack). The findings provide valuable insights into understanding the robustness of real-world networks by opening a vast and important area of research.

A recent study investigated how weight thresholding procedures, which remove weak links (links of lower weight), affect the robustness of real-world networks to node attacks, and the rank of node centrality [2]. The study found that real-world networks subjected to the WT procedure have a robust connectivity structure to node attacks.

Here, we test whether *WT* with strong link removal changes the efficacy of the node attack strategies and how it affects the robustness of a set of real-world networks. To do this, we perform a sparsification procedure by removing a fixed fraction of higher-weight links. After sparsification, we execute a network attack by removing nodes using different node centrality indicators from the literature.

Performing the removal of strong links followed by a node attack can clarify the role that links of higher weights play in maintaining network connectivity. Previous studies have analyzed how the removal of strong links affects network connectivity. These studies removed links in decreasing (or increasing) order of weight and measured the resulting network connectivity using network structural indicators [11–16]. Here, we adopt a different and novel approach by removing strong links and then testing the resulting network structure with further node removals (attacks).

Generally, the real-world networks under study show robust connectivity against the *WT* procedure. Differently, the *WT* procedure removing strong links induces a more significant change in the ranking of nodes than the weak *WT* procedure.

2. Methods

2.1. Real-World Networks

We implemented five different node attack strategies on nine real-world weighted networks from different domains. Table 1 summarizes the statistics of these real-world networks, with node, link, and link weight meaning.

Table 1. Statistics of real-world networks. *N* number of nodes; *L* number of links; *<k>* average node degree; *<w>* average link weight; *<CC>* global clustering coefficient; *LCC* size of the largest connected component.

Networks	Key	Ref.	Туре	Node	Link	Weight	Ν	L	< <i>k</i> >	<i><w></w></i>	< <i>CC</i> >	LCC
C. Elegans	Eleg	[17,18]	Biological	Neurons	Neurons connection	Number of Connections	297	2344	15.8	3.761	0.181	297
Cargoship	Cargo	[19]	Transport	Ports	Route	Shipping journeys	834	4348	10.4	97.709	0.222	821
US airport	Air	[20]	Transport	Airports	Route	Passengers	500	2979	11.9	152320.2	0.351	500
E. Coli	Coli	[19,21]	Biological	Metabolites	Common reaction	Number of Common reactions	1100	3636	6.61	1.364	0.139	1100
Netscience	Net	[22]	Social	Authors	Coauthorship	Number of Common papers	1461	2741	3.75	0.434	0.693	379
Human 12a	Hum	[23,24]	Biological	Brain regions	Connection between regions	Connection density	501	6038	24.1	0.01	0.457	501
Caribbean	Carib	[25,26]	Ecological Food web	Species	Trophic relation	Amount of biomass	249	3503	28.13	0.067	0.172	249
CypDry	Сур	[16,27]	Ecological Food web	Species	Trophic relation	Amount of biomass	66	503	15.24	0.358	0.421	65
Budapest	Buda	[28]	Biological	Brain regions	Neural connection	Amount of track flow	480	1000	4.167	5.024	0.120	467

2.2. Attack Strategies

We simulated the following centrality-based node attacks in the networks: nodes with the highest centrality were removed first.

Random (*Ran*): Nodes are removed randomly. Selecting nodes at random is analogous to simulating errors or failures in the network [9,10].

• Degree (*Deg*): The degree of a node is the number of links connected to it [10,29–32]. The degree k_i of node *i* is given by the following:

$$k_i = \sum_{j=1}^N a_{ij},\tag{1}$$

where $a_{ij} = 1$ indicates the presence of a link between nodes *i* and *j* and is 0 otherwise. *N* is the number of nodes in the network.

• Strength (*Str*): A node's strength is the total weight of the links connected to that node [33], also called a weighted degree.

Mathematically, the strength s_i of node *i* is as follows:

$$s_i = \sum_{j=1}^{N} a_{ij} \cdot w_{ij},$$
 (2)

where $a_{ij} = 1$ indicates the presence of a link between nodes *i* and *j* and is 0 otherwise. w_{ij} is the weight of the link between *i* and *j*.

• Betweenness (*Bet*): Betweenness of a node is the number of shortest paths passing through it [29–31]. This binary metric defines the shortest path between two nodes as the minimum number of links needed to travel between them.

Mathematically, the betweenness b_i of node *i* is as follows:

$$b_i = \sum_{s,t=1}^N \frac{\sigma_{st}(i)}{\sigma_{st}},\tag{3}$$

where $\sigma_{st}(i)$ is the number of shortest paths between nodes *s* and *t* passing through the node *i*. σ_{st} is the total number of shortest paths between nodes *s* and *t*.

• Weighted Betweenness (*WBet*): Weighted betweenness of a node is defined as the number of weighted shortest paths passing through that node [34].

Weighted betweenness b_i^w of node *i* is as follows:

$$b_i^w = \sum_{s,t=1}^N \frac{\sigma_{st}^w(i)}{\sigma_{st}^w} , \qquad (4)$$

where $\sigma_{st}^{w}(i)$ is the number of weighted shortest paths between nodes *s* and *t* passing through the node *i*. σ_{st}^{w} is the total number of weighted shortest paths between nodes *s* and *t*.

While computing shortest paths, it is fundamental to consider whether the link weight corresponds to "flows" or "costs" [35]. If link weight means flow, then the shortest path is computed by summing the inverse of link weights. If link weights are costs, shortest paths are computed directly by summing the link weights.

These attacks are performed by removing nodes and the links incident on them by targeting the nodes according to the decreasing order of their centrality values (*Deg, Str, Bet, WBet*). First, the node with the highest centrality is targeted, and the attack is continued on lesser centrality nodes until the network collapses. Attacking the nodes based on their pre-calculated rank is known as an initial (not recalculated) or simultaneous attack strategy [29]. However, the network structure may change after each attack, and the nodes' importance may also change. In such a scenario, the pre-calculated ranking of nodes may no longer be valid. Here, we recalculated the node centrality values and updated the node's rank after each attack [29]. This attack strategy is known as a recalculated (also named adaptive) attack strategy. In the case of ties (i.e., nodes with equal centrality value), we randomly selected the node to remove. These node ties were randomized by averaging the outcomes over 100 simulations.

2.3. Weight Thresholding

We investigated the effect of strong link removal on the robustness of real-world networks under various node attack strategies. This analysis was performed using the weight thresholding (*WT*) technique. The *WT* is performed by removing a fraction of the strong links. Given a weighted network *G*, the first step is to apply the weight thresholding on *G*. In our study, we took nineteen discrete threshold values $WT = \{0.0, 0.05, 0.1, 0.15, 0.2, 0.25, 0.3, 0.35, 0.4, 0.45, 0.5, 0.55, 0.6, 0.65, 0.7, 0.75, 0.8, 0.85, 0.9\}$ (i.e., from 0% to 90% removal of strong links). In the case of ties (links having the same weight), the links were selected randomly. These ties were randomized by averaging the outcomes over 100 simulations. The thresholded network *G'* was the subgraph of *G* with the same number of nodes. Then, node attack strategies on *G'* were applied by identifying the nodes in the decreasing order of their centrality measures (*Deg, Bet, Str, and WBet*) computed from *G'*. This procedure was repeated for each *WT*.

2.4. Network Robustness Indicator

The largest connected component (*LCC*) is the simplest binary measure of the network's functioning along node removal. It is defined as the highest number of connected nodes in the network [9,10,32]. Here, normalized *LCC* against the fraction (q) of nodes removed was used to measure network damage. The normalization was performed in two ways.

- 1. One way was to normalize the *LCC* after node removal using the initial *LCC* value (before node attack) of the network after *WT*. In this case, we considered each thresholded network an independent network, and we did not account for the *LCC* decrease directly caused by the *WT* procedure.
- 2. A second way was to normalize the *LCC* after node removal using the initial *LCC* at WT = 0, i.e., we normalized using the *LCC* of the original network. In this second case, we also considered the *LCC* decrease triggered by the link removal of the WT procedure. This normalization was intended to analyze the joint effect of the weight thresholding and node attack to decrease the *LCC* (total *LCC* decrease).

For ease of comparison, the response of networks to each attack strategy was represented by a single number called robustness (R). It was defined as the area under the curve of network functioning measure (here, *LCC*) against the fraction (q) of nodes removed. From now on, we refer to 'robustness' as the R measure computed with the first *LCC* normalization (R) and 'total robustness' (R_{tot}) as the measure computed with the second *LCC* normalization. Table 2 lists the abbreviations used in this manuscript.

Full Name
Weight thresholding
Size of largest connected component
Number of nodes
Number of links
Average link weight
Average node degree
Global clustering coefficient
Random node attack
Degree node attack

Table 2. List of the abbreviations used in this manuscript.

Abbreviation	Full Name								
Str	Strength node attack								
Bet	Betweenness node attack								
WBet	Weighted Betweenness node attack								
G	Weighted network								
G'	Thresholded network								
L'	Number of links in <i>G</i> ′								
9	Fraction of nodes removed								
R	Robustness								
R _{tot}	Total Robustness								
<\$>	Average node strength								
Initial_Weak WT	WT by weak link removal with initial node attack strategy								
Initial_Strong WT	WT by strong link removal with initial node attack strategy								
Recalculated_Weak WT	WT by weak link removal with recalculated node attack strategy								
Recalculated_Strong WT	WT by strong link removal with recalculated node attack strategy								

Table 2. Cont.

3. Results and Discussion

3.1. Robustness against WT

We investigate the role of strong links on the robustness of networks to node attack strategies. The WT removes a fixed fraction of strong links, and then, we perform the node attack strategies on each thresholded network. These strategies are performed using initial and recalculated node attack methods.

Figures 1 and 2 show the *LCC* and the robustness *R* as a function of *WT* for different real-world networks. First, we analyze the *LCC* decrease induced by the *WT* procedure. The bar plots in the first column of Figures 1 and 2 depict this *LCC* decrease. The networks *C. Elegans*, Caribbean, Human12a, and US airports show the slowest *LCC* decrease when subjected to the *WT* procedure. The *WT* procedure corresponds to the classic strong link removal [36]. Specifically, *C. Elegans* and the Caribbean keeps 80% of the *LCC* even up to 75% removal of strong links (*WT* = 0.75), and Human12a keeps 85% of the *LCC* for 80% removal of strong links (*WT* ≤ 0.85). The smallest network in our study, Cypdry (*N* = 66), and the air transportation network, US airports, also maintain a comparable *LCC* up to WT = 0.7.

The other networks, such as *E. Coli*, Budapest, Cargoship, and Netscience, present a lower robustness against the *WT* procedure, showing a faster *LCC* decrease than other networks. Budapest and Netscience networks show a faster *LCC* disruption under the *WT* procedure. Removing strong links accelerates the fragmentation of the science coauthorship networks (Netscience). In this network, dense local neighborhoods of scientists are primarily composed of weak links. In contrast, the strong links represent more significant and enduring connections among leading scholars, bridging distant research communities and thus playing a crucial role in overall network connectivity [14].

In summary, the real-world networks under study are robust to the strong *WT* procedure regarding the *LCC*. For this reason, the real-world networks under study unveil general robustness to strong link removal [36].



Figure 1. The LCC after each weight thresholding (*WT*) value (left column), the robustness (*R*) of the network under the initial (middle column), and the recalculated attack strategies (right column) as a function of the weight thresholding (*WT*) value for the networks *C. Elegans* (Eleg), Caribbean (Carib), Human12a (Hum), Cypdry (Cyp), and *E. Coli* (Coli).



Figure 2. The LCC after each weight thresholding (*WT*) value (left column), the robustness (*R*) of the network under the initial (middle column), and the recalculated attack (right column) strategies as a function of the weight thresholding (*WT*) value for the networks Budapest (Buda), Cargoship (Cargo), US airports (Air), and Netscience (Net).

3.2. Robustness to WT and Node Attack

We investigate the network robustness against the coupled effect of the *WT* and node attack strategies in two ways.

First, we normalize the *LCC* along node removal with the initial *LCC* of the network after the *WT* procedure. This normalization does not consider the *LCC* decrease triggered by the *WT* link removal. This normalization evaluates the network after *WT* as an independent system and accounts only for the *LCC* decrease caused by the node attack. The trends of the robustness *R* with this normalization procedure for the node attack strategies, *Ran*, *Deg*, *Str*, *Bet*, *and WBet* are represented in Figures 1 and 2.

Another way is to compute the relative robustness normalizing the *LCC* over the original *LCC* size, i.e., before *WT* and a node attack. In this manner, we can understand the decrease in network functioning by the joint effect of *WT* and a node attack (i.e., total robustness R_{tot}). The R_{tot} for the node attack strategies, *Ran*, *Deg*, *Str*, *Bet*, and *WBet*, is represented in Figures 1 and 2.

We find a gradual change in *R* along the *WT* in both the initial and recalculated strategies for most of the networks. The C. *Elegans* network almost maintains steady robustness in all the attack strategies up to WT = 0.75. After removing 75% of the strong links, we can see a drop in the robustness of the network. The *C. Elegans* network, with the remaining 25% weak links, is highly vulnerable to all the attack strategies. The networks Caribbean, Human12a, *E. Coli*, Cargoship, and US airports show gradual changes in robustness after each thresholding even up to WT = 0.90. Instead of a smooth change in *R*, the network Cypdry shows some spikes in *R*, especially towards the *Bet* (red) and *Str* (purple) attack strategies.

The total robustness R_{tot} (solid lines) follows a similar pattern of robustness decrease for all the attack strategies except *Ran* (see green dotted and solid lines). In networks such as *C. Elegans*, Human12a, and *E. Coli*, the joint effect of thresholding and node attacks (R_{tot}) returns roughly the same robustness computed with the first normalization procedure (*R*). In all other networks, we can observe only a small difference in the values of these two types of robustness when focusing on targeted attacks. Differently, the robustness of the networks against random removal is always lower when considering the joint effect of *WT* and random node attacks. The solid green lines describing the R_{tot} decrease with increasing *WT* in Figures 1 and 2 are significantly lower than the dotted green lines (*R*).

The principal aim of *WT* is to remove links to simplify the networks, making them easier to analyze and reducing the simulation time. Previous analyses showed that many standard network features quickly change under the *WT* procedure [1,6]. Here, we test whether *WT* with strong link removal changes the robustness of real-world networks when subjected to a node attack. Combining these results leads to the point that the real-world networks analyzed here hold comparable robust connectivity using both the two normalization procedures of the *LCC*.

There are some exceptions in Budapest, Netscience, and CypDry networks when we consider the normalization with the initial *LCC* of the network after the *WT* procedure. The Budapest network shows a higher robustness structure towards the end of thresholding (*WT*>0.7) (see Figure 2). Figure 3 shows the *LCC* as a function of the fraction of nodes removed *q* for *Ran*, *Deg*, *Str*, *Bet*, and *WBet* attacks in the Budapest network for *WT* values 0.75, 0.8, 0.85, and 0.9. It clearly shows that a higher *WT* value returns a slower *LCC* decrease.



Figure 3. The *LCC* as a function of the fraction of nodes that removed *q* for *Ran*, *Deg*, *Str*, *Bet*, and *WBet* (both initial and recalculated) attacks in the Budapest network for *WT* values 0.75, 0.8, 0.85, and 0.9.

In Figure 2, the Netscience also shows a higher robustness structure for some thresholding (WT > 0.2). The effect is also visible in Figure 4. The Cyp network also exhibits an increase in robustness when nodes are removed randomly (*Ran*). This interesting and counterintuitive result reveals that the network structures after *WT* may show a more robust *LCC* connectivity structure to node removal. In other words, the strong link removal performed by applying *WT* can strengthen networks against node attacks.



Figure 4. The *LCC* as a function of the fraction of nodes removed *q* for *Ran*, *Deg*, *Str*, *Bet*, and *WBet* (both initial and recalculated) attacks in the Netscience network for *WT* values 0.25, 0.45, 0.55, and 0.65.

Scientific collaboration networks present links of higher weight connecting different communities of nodes [14]. Removing the strong links could fragment the scientific social network (Net) into smaller communities. Figure 5 shows that the Net network's node clustering coefficient ($\langle CC \rangle$) increases as a function of WT; that is, $\langle CC \rangle$ decreases when strong links are removed. Figure 5 shows an analogous $\langle CC \rangle$ increase at the end of the WT procedure for the Buda (WT > 0.8) and Cyp networks (WT > 0.75). The $\langle CC \rangle$ rise can explain why an increase in robustness is also observed for different node attacks (Ran, Str, WBet) in the Buda network (Figure 2) and for the Cyp network for random node removal (Ran) (Figure 2) at the end of the WT procedure. The Buda network is a complex brain network where nodes are brain regions and links indicate electrical activity between them [28]. The Cyp network is a food web ecological network in which nodes are species and links depict trophic interactions among them [16]. Global node clustering $\langle CC \rangle$ is a simple measure evaluating the presence of communities of nodes in networks [28], and it is a measure that counts node triplets in the network. A triplet is three nodes connected by either two (open triplet) or three (closed triplet) links. <*CC*> is the ratio between the number of closed triplets and the total number of triplets (both open and closed) in the network [28]. The higher the <CC>, the higher the node's tendency to cluster in communities.

Taking together the results would suggest that the removal of strong links can lead, in some cases, to the fragmentation of the network into communities (clusters of nodes) that are more resistant to node removal than the original network. This last pattern may explain the counterintuitive finding of increased network robustness in these real-world networks after applying strong *WT*. At the same time, this result would indicate that in highly clustered networks, removing bridge links (here, the strong links) connecting different communities of nodes may lead to a sparser network that is more resistant to node removal than the original one. Nonetheless, for the Air network, that is, the network of US airports [20], we observe a *<CC>* increase with *WT* but not a corresponding increase in robustness to node removal. For this reason, further mechanisms must be elucidated to understand why, in some real-world networks, the removal of strong links is associated with an increased robustness of the remaining network.



Figure 5. Real-world network features as a function of WT for each network. $\langle k \rangle$: average node degree; $\langle s \rangle$: average node strength; $\langle w \rangle$: average link weight; $\langle CC \rangle$: global clustering coefficient. For the ease of analysis, the network features are normalized by their maximum value.

3.3. The Efficacy of the Node Attack Strategies

Figures 6 and 7 list the best attack strategy, returning the lowest *R* value for each real-world network and each *WT* value. We find that with increasing *WT*, the efficacy of the attack strategy changes as well, and this is for both the normalization procedures of the *LCC*. For example, for initial node attack strategies, the best attack strategy for the *C. Elegans* network is the degree-based strategy (*Deg*) for $WT \le 0.1$, whereas for WT > 0.1, the betweenness attack strategy (*Bet*) becomes the best method to dismantle the *LCC* (Figure 7). For the Cargo network, the best attack strategy is *Str* for 0.25 $\le WT \le 0.4$; in the remaining *WT* parameter space, the best attack strategy is *Deg*.

These findings show that the strong link removal performed using the WT procedure changes the efficacy of the node attack strategies. This last result has two important consequences. (I) Finding the best node attack strategies in real-world networks is a fundamental problem in network science with many real applications [31,35,37,38]. The WT procedure aiming to simplify the network by reducing the number of links induces structural changes that affect the efficacy of the node attack strategies. For this, network science research focusing on node attack strategies must consider that applying WT may significantly change the node attack efficacy. (II) Finding the best attack strategies is a heuristic way to select important nodes in the network [35]. Here, we show that WT performed with strong link removal changes the efficacy of the attack strategies. Therefore, strong WT is likely affecting the node rank in the network [2]. To test how WT affects the rank of the different node centralities, we use Kendall's tau coefficient (τ) to evaluate the change in node rank after weight thresholding [39]. The τ coefficient is a measure of the magnitude of correspondence between two ranked pieces of data, i.e., the higher the Kendall's τ coefficient, the more similar the two ranking sequences. The range of Kendall's τ coefficient is from -1 to 1. We depict the results of this analysis in Figure 8. The τ coefficient decreases by increasing WT, indicating changes in the node rank after the WT procedure. Comparing the τ coefficient for strong WT (Figure 8, solid lines) with the τ coefficient discovered in a previous work by applying weak WT [2] (Figure 8, dashed lines), we find that strong WT produces a faster decrease in the τ coefficient. John et al. [2] found that applying the WT weak link removal decreases the τ coefficient to around 0.3 for most networks. By applying strong WT, we can lower the τ coefficient to 0 or even

negative values (Figure 8, solid lines). This indicates that sparsification procedures based on strong link removal may trigger a greater change in the node centrality rank concerning the sparsification procedures removing weak links. Network science research focusing on developing algorithms to find important influential nodes [40] has to consider that simplifying the network with *WT* methodologies may also change the node importance evaluated by different node centrality indicators in the network.



Figure 6. Best attack strategy returning the lowest *R* value for each real-world network and each *WT* value. In each cell, we indicate the best attack strategy and its R_{tot} value. The R_{tot} value is computed by normalizing the *LCC* with the initial *LCC* for *WT* = 0. Colors indicate the different attack strategies.

								WT by	strong link	removal (R)								
	Initial attack								Recalculated attack									
wT	Eleg	Carib	Hum	Сур	Coli	Buda	Cargo	Air	Net	Eleg	Carib	Hum	Сур	Coli	Buda	Cargo	Air	Net
0	Deg	Bet	Bet 0.441	Bet	Deg	Deg	Deg	Deg	Bet	Bet 0.212	Bet	Bet 0.219	Bet 0.254	Bet	Bet	Bet	Bet 0.051	Deg
0.05	Deg	Bet	Bet	Bet	Deg	Deg	Deg	Bet	Bet	Bet	Bet	Bet	Bet	Bet	Bet	Bet	Bet	Bet
	0.320	0.265	0.445	0.287	0.105	0.051	0.172	0.079	0.042	0.209	0.202	0.222	0.253	0.061	0.027	0.095	0.051	0.031
0.1	0 309	8et	Bet 0.438	Bet 0.288	Deg 0.102	Deg 0.050	Deg 0.172	Bet 0.077	Bet 0.039	Bet 0.208	Bet 0.192	Bet 0.223	Bet 0.253	Bet 0.060	Bet 0.026	Bet 0.092	Bet 0.051	Str 0.032
	Bet	Bet	Bet	Bet	Deg	Deg	Deg	Bet	Bet	Bet	Bet	Bet	Bet	WBet	Bet	Bet	Bet	Str
0.15	0.325	0.243	0.435	0.294	0.099	0.048	0.164	0.075	0.036	0.210	0.184	0.224	0.252	0.058	0.027	0.090	0.051	0.031
0.2	Bet	Bet	Bet	Bet	Str	Deg	Deg	Bet	Bet	WBet	Deg	Bet	Bet	WBet	Bet	Bet	Bet	Str
	Ret	0.255 Ret	Ret	0.200 Ret	Deg	0.047 Deg	0.167 Str	Bet	Bet	Bet/MBet	Bet	0.225 Ret	0.234 Ret	WRet	0.026 Ret	Ret	Bet	Str
0.25	0.298	0.226	0.423	0.295	0.086	0.046	0.164	0.074	0.036	0.199	0.169	0.222	0.246	0.051	0.025	0.090	0.050	0.035
0.2	Bet	Bet	Bet	Bet	Deg	Deg	Str	Deg	Bet	WBet	Bet	Bet	Bet	WBet	Bet	Bet	Bet	Str
0.5	0.292	0.222	0.416	0.284	0.079	0.042	0.161	0.076	0.049	0.194	0.162	0.222	0.246	0.047	0.023	0.089	0.050	0.042
0.35	Bet	Bet	Bet	Bet	Deg	Deg	Str 0.152	Deg	Bet	Wbet	Deg 0.157	Bet	Bet	Bet	WBet	Bet	Bet	Str
	Bet	Bet	Bet	Bet	Str	Deg	Str	Deg	Bet	WBet	Bet	Str						
0.4	0.286	0.193	0.393	0.275	0.064	0.033	0.149	0.074	0.068	0.191	0.147	0.207	0.248	0.041	0.020	0.083	0.050	0.059
0.45	Bet	Bet	Bet	Bet	Deg	Deg	Str	Deg	Bet	Bet	Bet	Bet	Bet	WBet	Bet	Bet	Bet	Str
0.45	0.280	0.189	0.386	0.279	0.058	0.031	0.146	0.072	0.066	0.184	0.142	0.202	0.243	0.038	0.019	0.079	0.050	0.056
0.5	Bet	Bet	Bet	Bet	Deg	Deg	Deg	Deg	Bet	Bet	Bet	Bet	Bet	WBet	WBet	Bet	Bet	Deg
	0.261 Rot	0.1//	0.369 Rot	0.264	0.052	0.028	0.135 Dog	0.073	0.074 Rot	0.1//	0.132 Rot	0.191 Rot	0.221 Rot	0.034	0.018	0.075 Pot	0.050	0.057
0.55	0.261	0.165	0.344	0.263	0.046	0.025	0.122	0.071	0.085	0.170	0.122	0.183	0.215	0.031	0.017	0.072	0.050	0.061
0.6	Bet	Deg	Deg	Bet	Deg	Deg	Deg	Deg	Bet	WBet	Bet	Bet	Bet	WBet	WBet	Bet	Bet	Deg
0.0	0.246	0.139	0.324	0.240	0.041	0.023	0.121	0.073	0.086	0.164	0.111	0.173	0.180	0.029	0.017	0.068	0.051	0.067
0.65	Bet/WBet	Deg	Bet	Bet	Str	Deg	Deg	Deg	Bet	WBet	Bet	Bet	Bet	WBet	Deg	Bet	Bet	Deg
	Bet/WBet	Deg	0.310 Ret	0.224 Ret	0.035 Deg	0.021 Deg	Deg	0.068 Deg	0.098 Ret	U.166 WRet	Bet	0.166 Ret	0.155 Ret	0.026 Ret	0.016 Deg	0.060 Ret	0.047 Bet	0.071 Deg
0.7	0.239	0.121	0.266	0.192	0.030	0.020	0.100	0.061	0.083	0.166	0.097	0.146	0.144	0.022	0.014	0.058	0.048	0.053
0.75	Bet/WBet	Deg	Str	Bet	Str	Deg	Deg	Deg	Deg	WBet	Bet	WBet	Deg	WBet	Deg	Bet	Bet	Deg
0.75	0.239	0.105	0.240	0.180	0.025	0.022	0.085	0.056	0.057	0.166	0.085	0.127	0.155	0.019	0.014	0.050	0.042	0.034
0.8	Deg	Bet	Bet	Bet	Str	Str	Deg	Deg	Bet	Bet	Bet	WBet	Deg	WBet	Deg	Bet	Bet	Deg
	0.133 Dog	0.093	0.189	0.151	0.019	0.031	0.069 Dog	0.053	0.046	U.U95	0.077	0.100 Rot	0.124 Dog	0.016	0.020	0.041 Rot	0.039	0.032 Dog
0.85	0.093	0.076	0.123	0.124	0.014	0.048	0.052	0.048	0.024	0.067	0.066	0.069	0.102	0.011	0.028	0.033	0.032	0.017
	Deg	Deg	Deg	Bet	Str	Str	Deg	Deg	Bet	WBet	Deg	WBet	WBet	Str	Str	Bet	WBet	Deg
0.9	0.051	0.062	0.077	0.101	0.011	0.064	0.032	0.029	0.014	0.036	0.052	0.047	0.083	0.008	0.034	0.023	0.022	0.012

Figure 7. Best attack strategy returning the lowest *R* value for each real-world network and each *WT* value. In each cell, we indicate the best attack strategy and its *R* value. The *R* value is computed by normalizing the *LCC* with the initial *LCC* at each *WT* value. Colors indicate the different attack strategies.



Figure 8. Kendall's tau coefficient (τ) for centrality measures *Deg*, *Str*, *Bet*, and *WBet*. Correlation is measured between the initial network's node rank and the network's node rank after *WT*. We compute τ using the top 30% of nodes of the network. Solid lines indicate τ for *WT* with strong link removal; dashed lines indicate τ for *WT* with weak link removal as in [2].

3.4. Comparing Strong and Weak WT Procedures

John et al. [2] investigated the effect of weight thresholding (WT) on the robustness of real-world complex networks against various node attack strategies by removing a fixed fraction of weak links. In this study, we investigate the opposite perspective and perform WT by removing strong links. Figure 9 compares the R_{tot} against the initial node attack when weak and strong WT procedures simplify networks. We do not find a clear trend; in some cases, weak WT triggers a faster robustness decrease, and in others, it is to the contrary. For example, the weak WT induces a higher decrease in robustness concerning the strong WT for the Eleg, Cyp (except under WBet initial attack), Air, and Cargo networks for both the initial (Figure 7, red lines) and recalculated attacks (Figure 9, green lines). These results agree with the study by Onnela et al. [13] on mobile communication networks [13]. Onnela et al. [13] show the counterintuitive consequence that real-world social networks are robust to removing the strong links but fall apart quickly if the weak links are removed. Onnela et al. [13] analyzed the network's robustness to removing links only. Our study, on the other hand, analyzes the combined effect of removing links and then attacking the network by removing nodes. Despite the differences between Onnela et al. [13] and our approaches, similar systems' responses are observed: for certain types of real-world networks, removing weak links can induce greater fragility. Our results show that this may happen not only in social networks [13] but also in transportation and biological networks.



--- Initial_WeakWT --- Initial_StrongWT --- Recalculated_WeakWT --- Recalculated_StrongWT

Figure 9. Comparison between the total robustness (R_{tot}) against weak and strong *WT* procedures. Network robustness under the initial attack (dotted lines) and recalculated attack (solid lines) strategies as a function of the weight thresholding (*WT*) value for the networks *C. Elegans* (Eleg), Caribbean (Carib), Human12a (Hum), Cypdry (Cyp), *E. Coli* (Coli), Budapest (Buda), Cargoship (Cargo), US airports (Air), and Netscience (Net).

However, the strong *WT* returns lower robustness in the Carib and Hum networks, especially for initial node attacks. The Car network is a food web ecological network depicting who eats whom in the ecosystem. Hum is the human brain network modeling the electrical communication activities sharing information among brain regions [23]. Hence, from very different domains of science, these real-world networks show a higher fragility when combining strong *WT* and node attacks (Figure 9). From these results, it is possible to infer possible dynamics of these real-world networks. In food webs, removing

14 of 16

(or disrupting) the higher magnitude trophic connections between species may trigger the remaining ecological network to be more sensitive to species removal. The removal of species in food webs models the case of species extinction [16]. Our results would suggest that removing strong trophic interactions would make the ecosystem more prone to biodiversity loss. The higher vulnerability to strong *WT* in the brain network would indicate that once the connections with the highest electrical activity between different brain regions are removed, the remaining brain network is more prone to brain region malfunctioning (node removal) and becoming more easily disconnected. This may help in understanding the mechanisms by which brain networks and which brain regions play the main routing information.

Our latest results show the difficulty in predicting how different sparsification procedures may affect the robustness of node attacks on real-world networks. Different real-world networks may exhibit opposite behaviors regarding sparsification through removing the heaviest-weight links (strong *WT*) compared to removing the links of lower weights (weak *WT*).

4. Conclusions

We analyzed the impact of weight thresholding on the robustness of real-world networks to different node attack strategies. Here, weight thresholding is performed by removing a fixed fraction of strong links. Generally, the real-world networks under study show robust connectivity against the *WT* procedure. In other words, real-world networks maintain a robust structure regarding the *LCC* to strong link removal. These results suggest that strong link removal can be used as a method for the sparsification of networks for applications in which the robustness to node attacks is important.

Then, we find that applying *WT* may significantly change the node attack efficacy and the rank of different node centrality measurements. The strong *WT* procedure induces a greater change in the ranking of nodes than the weak *WT* procedure. For this reason, network research focusing on finding the efficacy of node attack strategies or finding important nodes in the network has to consider the network structural changes caused by the weight thresholding (sparsification) procedures.

Studying the robustness against node attacks after strong WT may have different real-world applications. Removing links with higher weights and then performing node attacks could help identify the parts of the network that are more robust (or less affected) when removing key connections. In the real world, this can be useful for designing network protection or reinforcement strategies in critical infrastructure networks, such as those for energy, transportation, or communications. These vital systems can benefit from identifying the robustness of network components resulting from attacks on strong links, planning their protection, and developing risk mitigation strategies.

Our research also has significant implications for understanding ecological networks. By identifying keystone species in food web ecological networks, we can gain insights into the mechanisms of biodiversity loss in ecosystems. Food webs are networks of species and their trophic interactions [16,41]. Strong *WT* can simulate the deletion of strong trophic interactions that occur with the extinction or decreasing abundance of the most general species/resources in ecosystems. The subsequent node removal can then model the occurrence of species extinction in the remaining parts of the food web ecological network, providing a deeper understanding of biodiversity loss mechanisms.

Moreover, the emergence of the role of strong and weak links is associated with the local structure of the social networks [42], and understanding the specific embedding of strong links is important to comprehend complex social systems.

For example, scientific collaboration networks present links of higher weight connecting different communities of nodes [14]. Removing the strong links could fragment the scientific social network into smaller communities. Subsequently, removing nodes from these communities can help us better understand the robustness and relationships within specific groups of scientists. Last, the results presented in this study can be useful in network science research that needs to simplify complex networked systems and in machine learning and neural network research that needs to reduce model complexity or eliminate less important network connections.

Author Contributions: Conceptualization, M.B, D.C. and R.A.; Methodology, J.M.J. and M.B.; Software, J.M.J.; Formal analysis, J.M.J.; Investigation, J.M.J., M.B., D.S.L. and R.A.; Writing—original draft, J.M.J., M.B., D.S.L. and D.C.; Writing—review & editing, J.M.J., M.B. and D.S.L. All authors have read and agreed to the published version of the manuscript.

Funding: This study is funded by the IIT Palakkad Technology IHub Foundation Technology Development Grant No IPTIF/HRD/DF/019/SEP33 and IIT Palakkad Technology IHub Foundation Doctoral Fellowship IPTIF/HRD/DF/019. This study is funded by the Ecosister project, under the National Recovery and Resilience Plan (NRRP), by Mission 4 Component 2 Investment 1.5—Call for tender No. 3277 of 30 December 2021 of the Italian Ministry of University, and by the European Union—NextGenerationEU. Award Number: Project Code ECS00000033, Concession Decree No. 1052 of 23 June 2022, adopted by the Italian Ministry. We acknowledge the CINECA award under the ISCRA initiative for the availability of high-performance computing resources and support.

Data Availability Statement: Data are contained within the article.

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

References

- 1. Yan, X.; Jeub, L.G.S.; Flammini, A.; Radicchi, F.; Fortunato, S. Weight thresholding on complex networks. *Phys. Rev. E* 2018, 98, 042304. [CrossRef]
- 2. John, J.M.; Bellingeri, M.; Lekha, D.S.; Cassi, D.; Alfieri, R. Effect of Weight Thresholding on the Robustness of Real-World Complex Networks to Central Node Attacks. *Mathematics* **2023**, *11*, 3482. [CrossRef]
- 3. Namaki, A.; Shirazi, A.H.; Raei, R.; Jafari, G. Network analysis of a financial market based on genuine correlation and threshold method. *Phys. A Stat. Mech. Its Appl.* **2011**, *390*, 3835–3841. [CrossRef]
- Lynall, M.-E.; Bassett, D.S.; Kerwin, R.; McKenna, P.J.; Kitzbichler, M.; Muller, U.; Bullmore, E. Functional connectivity and brain networks in schizophrenia. J. Neurosci. 2010, 30, 9477–9487. [CrossRef] [PubMed]
- 5. Allesina, S.; Bodini, A.; Bondavalli, C. Secondary extinctions in ecological networks: Bottlenecks unveiled. *Ecol. Model.* 2006, 194, 150–161. [CrossRef]
- 6. Garrison, K.A.; Scheinost, D.; Finn, E.S.; Shen, X.; Constable, R.T. The (in) stability of functional brain network measures across thresholds. *Neuroimage* **2015**, *118*, 651–661. [CrossRef]
- Kavzoglu, T.; Mather, P.M. Assessing artificial neural network pruning algorithms. In Proceedings of the 24th Annual Conference and Exhibition of the Remote Sensing Society, Greenwich, UK, 9–11 September 1998.
- Hongwu, P.; Deniz, G.; Shaoyi, H.; Tong, G.; Weiwen, J.; Orner, K.; Caiwen, D. Towards Sparsification of Graph Neural Networks. In Proceedings of the IEEE 40th International Conference on Computer Design (ICCD), Olympic Valley, CA, USA, 23–26 October 2022.
- 9. Albert, R.; Barabasi, A.-L. Statistical mechanics of complex networks. Rev. Mod. Phys. 2002, 74, 47–97. [CrossRef]
- Albert, R.; Jeong, H.; Barabasi, A.-L. Error and attack tolerance of complex networks. *Nat. Vol.* 2000, 406, 378–382. [CrossRef] [PubMed]
- 11. Garas, A.; Argyrakis, P.; Havlin, S. The structural role of weak and strong links in a financial market network. *Eur. Phys. J. B* 2008, 63, 265–271. [CrossRef]
- 12. Onnela, J.-P.; Saramäki, J.; Hyvönen, J.; Szabó, G.; De Menezes, M.A.; Kaski, K.; Barabási, A.-L.; Kertész, J. Analysis of a large-scale weighted network of one-to-one human communication. *New J. Phys.* **2007**, *9*, 179. [CrossRef]
- 13. Onnela, J.-P.; Saramäki, J.; Hyvönen, J.; Szabó, G.; Lazer, D.; Kaski, K.; Kertész, J.; Barabási, A.-L. Structure and tie strengths in mobile communication networks. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 7332–7336. [CrossRef] [PubMed]
- 14. Pan, R.K.; Saramaki, J. The strength of strong ties in scientific collaboration networks. Europhys. Lett. 2012, 97, 18007. [CrossRef]
- 15. Pajevic, S.; Plenz, D. The organization of strong links in complex networks. *Nat. Phys.* **2012**, *8*, 429–436. [CrossRef] [PubMed]
- 16. Bellingeri, M.; Vincenzi, S. Robustness of empirical food webs with varying consumer's sensitivities to loss of resources. *J. Theor. Biol.* **2013**, 333, 18–26. [CrossRef] [PubMed]
- 17. Watts, D.J.; Strogatz, S.H. Collective dynamics of 'small-world' networks. Nature 1998, 393, 440–442. [CrossRef]
- 18. Latora, V.; Nicosia, V.; Russo, G. Complex Networks: Principles, Methods and Applications; Cambridge University Press: Cambridge, UK, 2017.
- Allard, A.; Serrano, M.A.; Garcia-Perez, G.; Boguna, M. The geometric nature of weights in real complex networks. *Nat. Commun.* 2017, *8*, 14103. [CrossRef]

- 20. Colizza, V.; Pastor-Satorras, R.; Vespignani, A. Reaction-diffusion processes and metapopulation models in heterogeneous. *Nat. Phys.* **2007**, *3*, 276–282. [CrossRef]
- 21. Serrano, A.M.; Boguna, M.; Sagues, F. Uncovering the hidden geometry behind metabolic networks. *Mol. BioSystems* **2012**, *8*, 843–850. [CrossRef] [PubMed]
- 22. Newman, M.E.J. Finding community structure in networks using the eigenvectors of matrices. *Phys. Rev.* 2006, 74, 036104. [CrossRef]
- Avena-Koenigsberger, A.; Goni, J.; Betzel, R.F.; Heuvel, M.P.V.D.; Griffa, A.; Hagmann, P.; Thiran, J.-P.; Sporns, O. Using Pareto optimality to explore the topology and dynamics of the human connectome. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 2014, 369, 20130530. [CrossRef]
- 24. Hagmann, P.; Cammoun, L.; Gigandet, X.; Meuli, R.; Honey, C.J.; Wedeen, V.J.; Sporns, O. Mapping the structural core of human cerebral cortex. *PLoS Biol.* **2008**, *6*, 1479–1493. [CrossRef] [PubMed]
- 25. Bellingeri, M.; Bodini, A. Food web's backbones and energy delivery in ecosystems. Oikos 2016, 125, 586–594. [CrossRef]
- 26. Opitz, S. Trophic Interactions in Caribbean Coral Reefs; ICLARM: Penang, Malaysia, 1996.
- 27. Heymans, J.J.; Ulanowicz, R.E.; Bondavalli, C. Network analysis of the South Florida Everglades graminoid marshes and comparison with nearby cypress ecosystems. *Ecol. Model.* **2002**, *149*, 5–23. [CrossRef]
- Szalkai, B.; Kerepesi, C.; Varga, B.; Grolmusz, V. The Budapest Reference Connectome Server v2.0. Neurosci. Lett. 2015, 595, 60–62. [CrossRef] [PubMed]
- 29. Holme, P.; Kim, B.J.; Yoon, C.N.; Han, S.K. Attack vulnerability of complex networks. *Phys. Rev. E* 2002, 65, 056109. [CrossRef] [PubMed]
- 30. Bellingeri, M.; Cassi, D.; Vincenzi, S. Efficiency of attack strategies on complex model and real-world networks. *Phys. A Stat. Mech. Its Appl.* **2014**, 414, 174–180. [CrossRef]
- Iyer, S.; Killingback, T.; Sundaram, B.; Wang, Z. Attack robustness and centrality of complex networks. *PLoS ONE* 2013, *8*, e59613. [CrossRef]
- 32. Cohen, R.; Erez, K.; Ben-Avraham, D.; Havlin, S. Breakdown of the internet under intentional attack. *Phys. Rev. Lett.* 2001, *86*, 3682–3685. [CrossRef] [PubMed]
- 33. Bellingeri, M.; Cassi, D. Robustness of weighted networks. Phys. A Stat. Mech. Its Appl. 2018, 489, 47–55. [CrossRef]
- 34. Nguyen, Q.; Nguyen, N.K.K.; Cassi, D.; Bellingeri, M. New betweenness centrality node attack strategies for real-world complex weighted networks. *Complexity* **2021**, 2021, 1677445. [CrossRef]
- 35. Bellingeri, M.; Bevacqua, D.; Sartori, F.; Turchetto, M.; Scotognella, F.; Alfieri, R.; Nguyen, N.K.K.; Le, T.T.; Nguyen, Q.; Cassi, D. Considering weights in real social networks: A review. *Front. Phys.* **2023**, *11*, 1152243. [CrossRef]
- 36. Bellingeri, M.; Bevacqua, D.; Scotognella, F.A.; Alfieri, R.; Cassi, D. A comparative analysis of link removal strategies in real complex weighted networks. *Sci. Rep.* **2020**, *10*, 3911. [CrossRef] [PubMed]
- 37. Jordán, F. Keystone Species and Food Webs. Philos. Trans. R. Soc. Lond. Ser. B Biol. Sci. 2009, 364, 1733–1741. [CrossRef] [PubMed]
- Nie, T.; Guo, Z.; Zhao, K.; Lu, Z.-M. New attack strategies for complex networks. *Phys. A Stat. Mech. Its Appl.* 2015, 424, 248–253. [CrossRef]
- 39. Kendall, M.G. The treatment of ties in ranking problems. Biometrika 1945, 33, 239–251. [CrossRef] [PubMed]
- 40. Lü, L.; Chen, D.; Ren, X.-L.; Zhang, Q.-M.; Zhang, Y.-C.; Zhou, T. Vital nodes identification in complex networks. *Phys. Rep.* **2016**, 650, 1–63. [CrossRef]
- 41. Dunne, J.A.; Williams, R.J.M.D. Network structure and biodiversity loss in food webs: Robustness increases with connectance. *Ecol. Lett.* **2002**, *5*, 558–567. [CrossRef]
- 42. Shang, K.-K.; Small, M.; Yin, D.; Li, T.-C.; Yan, W. The key to the weak-ties phenomenon. *Europhys. Lett.* 2019, 127, 48002. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.