WILEY | Hindawi

Research Article

New Betweenness Centrality Node Attack Strategies for Real-World Complex Weighted Networks

Quang Nguyen [1,2] **Ngoc-Kim-Khanh Nguyen,** [3] **Davide Cassi,** [4,5] **and Michele Bellingeri** [4,5,6]

[1]*Institute of Fundamental and Applied Sciences, Duy Tan University, Ho Chi Minh City 700000, Vietnam*
[2]*Faculty of Natural Sciences, Duy Tan University, Da Nang City 550000, Vietnam*
[3]*Van Lang University, Ho Chi Minh City, Vietnam*
[4]*Dip. Scienze Matematiche, Fisiche e Informatiche, Università di Parma, Parco Area delle Scienze, 7/A, Parma 43124, Italy*
[6]*Dipartimento di Fisica, Politecnico di Milano, Piazza Leonardo da Vinci 32, Milano 20133, Italy*
[5]*INFN, Gruppo Collegato di Parma, I-43124 Parma, Italy*

Correspondence should be addressed to Quang Nguyen; nguyenquang29@duytan.edu.vn

In this work, we introduce a new node attack strategy removing nodes with the highest conditional weighted betweenness centrality (CondWBet), which combines the weighted structure of the network and the node's conditional betweenness. We compare its efficacy with well-known attack strategies from literature over five real-world complex weighted networks. We use the network weighted efficiency (WEFF) like a measure encompassing the weighted structure of the network, in addition to the commonly used binary-topological measure, i.e., the largest connected cluster (LCC). We find that if the measure is WEFF, the CondWBet strategy is the best to decrease WEFF in 3 out of 5 cases. Further, CondWBet is the most effective strategy to reduce WEFF at the beginning of the removal process, whereas the Strength that removes nodes with the highest sum of the link weights first shows the highest efficacy in the final phase of the removal process when the network is broken into many small clusters. These last outcomes would suggest that a better attacking in weighted networks strategy could be a combination of the CondWBet and Strength strategies.

1. Introduction

The study of real-world complex networks has attracted much attention in recent decades because a large number of real complex systems can be abstracted as networks [1, 2]. One of the fundamental research topics is their robustness (resilience), i.e., capacity of the network to hold its functioning when a proportion of nodes are removed/blocked [3–12]. The robustness is usually evaluated by the size of the largest connected cluster (LCC) in the network and is a fundamental problem of theoretical and applied network science with huge efforts made in recent years [3, 7, 9, 12–17]. Previous studies showed that most real-world networks are resilient to random failure [4] but can disintegrate quickly when a small proportion of most connected nodes (hubs) are attacked [13]. Furthermore, one of the remarkable observations is when the proportion of removed nodes is high enough, a phase transition occurs, and the probability for the existence of a LCC in the network abruptly collapses (i.e., the network loses its global nodes connectivity). Monte Carlo simulation is usually used to run the attack by removing nodes according to a structural criterion and tracing the network damage using measures/indicators of its functioning. The most common criteria are node's centrality such as degree centrality [3, 4, 5, 9], closeness centrality [9, 18], and betweenness centrality [7, 14, 19]. Other node attack strategies are based on eigenvector [9, 20], the degree of the second neighbors [7], and entropy [21].

Overall findings showed that nodes attack strategies based on betweenness centrality are highly efficient to dismantle the LCC for most model and real-world networks [7, 9, 10, 12]. The attack based on the betweenness centrality (Bet) ranks and removes nodes according to their role in routing the shortest paths in the network; i.e., nodes with higher betweenness centrality are the ones that pass the major number of shortest paths among other nodes [7, 12]. This finding indicated that the most important nodes are not necessarily the most connected ones (i.e., hubs), but they can be nodes with a medium or low connectivity level and high betweenness centrality. However, it is found that the Bet strategy becomes ineffective at the end of the removal process when the network may be broken in many fully connected subnetworks (clusters) [10]. This is due to the inherent nature of betweenness's definition for a fully connected network (e.g., a complete graph); all node's betweenness is zero. At this stage, the Bet strategy was unable to break the fully connected LCC for a very long time, and finally, it fails to completely break the network earlier than other strategies, such as the degree-based strategy. To overcome this limitation of the betweenness nodes attack strategies, Nguyen et al. [22] introduced the "conditional attack strategy" (CondBet) that removes nodes according to their betweenness only if they belong to the LCC, demonstrating how the new conditional strategy outperforms the classic nodes attack strategies. However, most of the aforementioned researches investigate binary complex networks, where the interaction between two nodes is either 1 (have an edge) or 0 (no edge).

In reality, many real-world complex networks are naturally weighted with some interaction value associated to the links. For example, in a communication network, the link weights may represent the frequency of e-mail exchanges [23, 24]; in the Internet network link weights could be the connection bandwidth magnitude [24, 25]; in a flights transportation network, the link weights could be the total passengers flowing between airports ([26]. In some cases, real networks own very large link weight heterogeneity, with link weights spanning over five order of magnitude, such as the passengers flow between airports [8]. This implies that, to perform more real and precise network description, it is necessary to account the link weights heterogeneity.

Bellingeri and Cassi [27] analyzed the robustness of real-world networks to different nodes attack strategies using binary and weighted indicators of network functioning. The authors found that the inclusion of link weights in the analyses changes the network response to nodes attack, outlining the importance to investigate the performance of nodes attack strategies in weighted networks [27]. This leads us to the questions: (1) Which is the best nodes attack strategy to harm real weighted networks? And (2) How does the inclusion of link weight changes the global efficacy of the nodes attack betweenness-based strategies? (3) What metric should we use to measure the robustness of a weighted network?

In this paper, we analyzed the robustness of a high-quality dataset of real-world weighted networks from different fields of science. We implemented six attack strategies based on binary and weighted properties of the nodes. We adopted the

recalculated (adaptive) version of the nodes attack strategies, in which the nodes rank is updated after each nodes removal [12]. Here, we introduce for the first time the conditional nodes weighted betweenness centrality (CondWBet) attack strategy (the weighted version of the CondBet strategy) that remove nodes of highest weighted betweenness centrality inside the LCC. We test the efficacy of the different nodes attack strategies by computing their impact on network robustness using both binary and weighted measures of network functioning, i.e., the largest connected cluster (LCC) and the weighted efficiency (WEFF).

We found that the conditional betweenness attack strategy (CondBet) is the best to fragment the LCC in all the networks confirming previous evidences [22]. Then, when the goal is to decrease the network efficiency measure (WEFF), in 3 out of 5 networks, the new weighted conditional betweenness strategy (CondWBet) is the best strategy. On the other hand, in two other networks, the nodes strength attack strategies (Strength that removes nodes with highest sum of link weights) and the binary conditional betweenness strategy (CondBet) are the best strategies. We propose that an efficient breakdown of a weighted network, as measured by the WEFF, can be divided into two phases: the first one is when the network is still dense and the CondBet/CondWBet are more appropriate to dismantle the network as they can separate the network into smaller clusters; the second one is when the network is broken into many small clusters, and then the Strength strategy is more efficient as it breaks the nodes that reduce the most of the WEFF. The introduction of the new node attack strategy CondWBet and its efficacy comparison with other strategies from literature is the main contribution our work.

The paper is organized as follows. In Section 2, we describe the data and methods used in this work. Section 3 presents the empirical findings: the efficacy of six intentional attacking strategies over five real-world weighted networks. Finally, Section 4 summarizes the results and concludes the study.

2. Method and Data

2.1. The Nodes Attack Strategies. We simulate nodes attacks strategies belonging to two main groups, the binary and the weighted-based strategies. In each group, we use node degree, node betweenness centrality, and the conditional betweenness centrality for nodes ranking. In case of nodes ranking ties (e.g., nodes with equals rank), we randomly sort one of them.

In total, we adopt six nodes attack strategies:

(i) Deg: degree-based attack strategy removes nodes according to their degree; i.e., the degree of the node is the number of links to it [3]. This is the simplest and the oldest type of node attack widely used to test the networks robustness.

(ii) Bet: betweenness attack strategy removes the nodes with the highest betweenness centrality first. The betweenness centrality of the node is a macroscale network metric measuring the number of times a

node appears in the shortest path between all pairs of nodes in the network [7, 12, 28].

(iii) CondBet: conditional betweenness attack strategy is the improved version of the Bet [22]. The CondBet removes the node with highest betweenness if it is in the LCC; otherwise, it removes the node with highest betweenness inside the LCC. In other words, the CondBet removes nodes inside the LCC only, maximizing the efficacy of the Bet strategy into disrupting the LCC. It has been shown that the CondBet performs better than Bet, and it owns higher efficacy to fragment the LCC on a variety of real-world networks.

(iv) Strength: strength attack strategy removes nodes in decreasing rank of strength. The strength of the node is the sum of the link weights pointing to it [27]. The strength of the node is defined as the "weighted degree" and for this reason, the Strength can be viewed as the weighted counterpart of the Deg strategy.

(v) WBet: Weighted Betweenness attack strategy removes nodes according to their weighted betweenness centrality. Like the binary version of the nodes betweenness centrality, the weighted betweenness centrality for each node is the number of the shortest paths that pass through the nodes. The difference in this case is that the shortest paths among nodes are weighted shortest path (WSP). To compute WSP, it is necessary to distinguish whether the link weights indicate "flow" or "cost" among nodes [28]. Examples of "weight as a flows" are the strength of friendship among individuals, the number of common papers among authors, the number of flights among airports, and the number of synapses among neurons. Examples of "weight as a cost" are the distance among cities, the resistance in a circuit, and the time to information delivery between computers. If they are "flow," in WSP, each link contributes to the shortest path according to their inverse weight. This is a standard procedure making links of higher weight equivalent to "larger, faster, and shorter route" between nodes. Since our real-world networks have link weights as a flow, the WSP between two nodes is the path minimizing the sum of the inverse link weights between that pair of nodes. Differently, if they are "cost," the WSP between nodes is the path minimizing the sum of the link weights to travel between them.

(vi) CondWBet: conditional weighted betweenness attack strategy is introduced for the first time in this work. It is the weighted version of the CondBet. The CondWBet strategy removes nodes inside the LCC according to their recalculated weighted betweenness centrality.

3. The Real-World Networks

We analyzed the efficacy of the nodes attack strategies on five real-world weighted networks (Table 1). The first two are financial networks constructed from stock prices in the SP500 index in the US market using a threshold method [29]. We keep the value of the correlation coefficients that are higher than the threshold as link weights. By adjusting the threshold, we can obtain two networks with a similar average degree and number of nodes but with different topological structures. The first network, (i) the SP500_1, is built from Feb. 1993 to Feb. 1996, which is relatively uniformly connected and contains 315 nodes and 8706 links (see Figure 1(a)). The second, (ii) SP500_2, is built from May 1999 to May 2002, which contains 371 nodes and 10636 links [22]. This network has several well-connected clusters connecting to the central bulk through intermediated stocks (see Figure 1(c)). The other three networks are as follows: (iii) the co-authorship undirected network of scientists working on network theory and experiments (NetScience) compiled by Newman [30]. Nodes represent authors, and link weights represent the number of common papers. The network includes 1589 nodes and 2742 links (see Figure 1(e)); (iv) the network of 6005 peoples who trade Bitcoin on a platform called Bitcoin OTC (Bitcoin) [31, 29]. The weight of the links represents the rate of members on other members, which is in a scale of −10 (total distrust) to +10 (total trust) with steps of 1. We simplified the network by removing self-node links and converting it to an undirected network. It results in a network of 6005 nodes and 21492 links (see Figure 1(g)); (v) the network of 500 busiest airports in U.S. [33], where nodes represent airports, and the weight of a link identifies the normalized passengers flowing between two airports/nodes (see Figure 1(i)) [8].

A detailed summary of all networks is represented in Table 1. Overall, our networks have a number of nodes from 315 to 6005. The two financial networks, the SP500_1 and SP500_2, are very dense with an average degree of more than 27, while the NetScience network is the sparsest one with an average degree of 1.72. Also, the NetScience network is composed of 396 subgraphs, and its LCC's size is only 23.9% of the total number of nodes N. Its clustering coefficient C and modularity Q are also highest among networks. Other networks are almost connected with the LCC's size of more than 99% of N. Finally, we compute the new metric $<r>$ that characterizes the influence of network link weights to the ranking of nodes' betweenness, where $<r>$ is the ratio of pair of nodes whose relative betweenness changes when weights are included. Concretely,

$$<r> = \frac{1}{N \cdot (N-1)} \sum_{i \neq j \in G} I\{((bet_i > bet_j) \, and \, (wbet_i < wbet_j)) \, or \, ((bet_i < bet_j) \, and \, (wbet_i > wbet_j))\}, \quad (1)$$

TABLE 1: Structural statistics of the real-world networks: network type, nodes type, weight definition, N nodes' number, L links' number, <k> average node degree, <s> average node strength, $\langle w \rangle$ average link weight, C global clustering coefficient, Q the modularity, LCC the size of initial largest clustered components, LCC% the relative size of the LCC, Nb_comp the number of components, the new metric <r> for the ratio of pair betwenneess ranking change, and the network Reference.

| Name	SP500_1	SP500_2	NetScience	Bitcoin	USAirport
Type	Financial	Financial	Co-author	Social	Traffic
Nodes	Stock	Stock	Scientist	Member	Airport
Weight	Correlation coefficient	Correlation coefficient	Number of common papers	Trusting rate	Passengers flow
N	315	317	1589	6005	499
L	8706	10636	2742	21492	2980
<k>	27.6	28.6	1.72	3.57	5.97
<s>	10.349	18.817	1.497	6.538	0.0675
$\langle w \rangle$	0.187	0.328	0.433	0.913	0.018
C	0.511	0.718	0.878	0.288	0.726
Q	0.253	0.373	0.954	0.498	0.283
LCC	315	369	379	5875	499
LCC%	100.0%	99.5%	23.9%	97.8%	100.0%
Nb	1	2	396	128	1
<r>	0.39300	0.43626	0.00752	0.14802	0.31965
Ref.	[22]	[22]	[30, 34]	[31, 32]	[33]
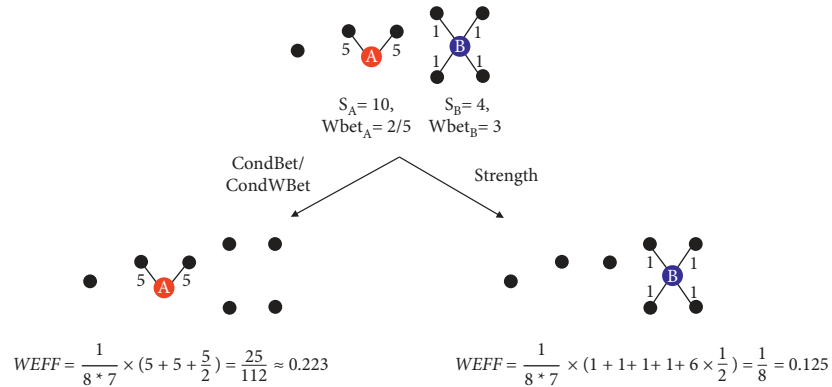
FIGURE 1: An example of weighted network composed of two small clusters. The node A has higher strength (10) and lower weighted betweenness (2/5) than node B; the node B presents lower strength (4), and higher weighted betweenness (3) than node (A) For this reason, the CondWBet will remove the node B and the Strength will remove node (A) The CondWBet is more efficient than Strength when the robustness measure is the LCC, but it is worse than the Strength when measured by the WEFF.

4. The Network Robustness Measures

We used two measures of network robustness along the nodes attack process, e.g., the size of the largest connected cluster (LCC) and the weighted efficiency (WEFF). The LCC is a simple indicator evaluating the binary-topological connectivity of the network nodes.

The network efficiency (WEFF) is introduced by [35] with the goal to account the network information delivery rate in the network. WEFF is the average of the sum of the inverse of the weighted shortest paths (WSP) among nodes:

where N is the total number of nodes of network G, bet_i (wbet$_i$) is the betweenness (weighted betwennness) of node i and I{} is the indicator function. In other words, higher <r> indicates that the inclusion of link weights changes the node betweenness centrality ranking.

$$\text{WEFF} = \frac{1}{N \cdot (N-1)} \sum_{i \neq j \in G} \frac{1}{d(i,j)}, \quad (2)$$

where N is the total number of nodes of network G and d(i, j) is the length of the WSP between node i and node j.

The WEFF is a measure that considers the difference in link weights in the evaluation of the weighted network functioning (integrity) and can be viewed like an indicator of how efficiently the network nodes exchange information [19, 35]. Recently, the network efficiency has been used to evaluate and compare the efficacy of different nodes attack strategy for weighted networks [27, 36, 37].

In addition, for each attack strategy, we compute a single value defined as the network robustness (R) as done in [37]. The value of R corresponds to the area below the curve of the system functioning indicators (LCC and WEFF) against the fraction of nodes removed. We denote
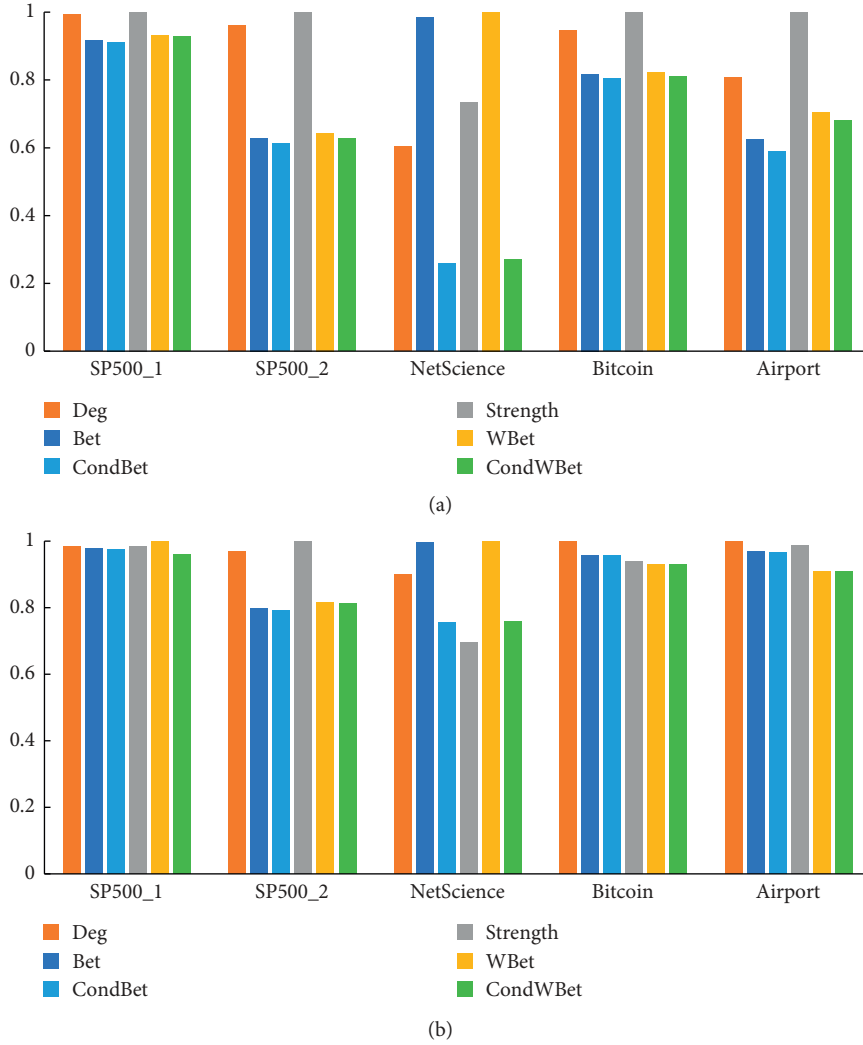
(a)

(b)

Figure 2: The single value network robustness (R) for LCC (a) and WEFF (b) for each network and strategy. To easily compare the efficacy of the strategies, we normalized R by the max robustness value produced by a strategy for that network.

R_{LCC} and R_{WEFF} as the network robustness when using LCC and WEFF functioning indicators, respectively.

5. Results

5.1. Efficacy of the Attack Strategies with LCC. We found that the CondBet is the best strategy to minimize R_{LCC} for all the five networks (Figure 2 and Table 2). This outcome confirms the highest efficacy of this strategy into fragment the LCC [22]. CondWBet is the second best strategy to fragment the LCC outperforming the WBet counterpart in all the network (Figure 2). We can see that, at the end of the removal process, the nodes betweenness centrality strategies such as Bet and WBet become inefficient, especially in the SP500_2 and NetScience networks (Figure 3). The Bet and WBet nodes attack strategies get stuck in a fully connected LCC and are not able to break it for a long fraction of removals. In fact, both binary (Bet) and weighted (WBet) nodes betweenness centrality inside a fully connected LCC are zero for all the nodes (because all the paths inside this subnetwork are

shortest paths); thus, the Bet and WBet strategies are not able to select nodes and fragment the LCC. In Figure 3, we compare the LCC size after removing a fraction q of nodes by the attack strategies outlining the conditional strategies CondBet and CondWBet as the best to attack the LCC; this is because they are able to select the most important nodes during the entire removal process.

Moreover, we find that including link weights as done in the Strength, WBet, and CondWBet worsens the efficacy of the strategies with respect to their corresponding nonweighted strategy counterpart (Deg, Bet, and CondBet, respectively) (Figures 2 and 3). For example, in the SP500_2 network, the removal of $q = 4\%$ nodes by CondBet strategy triggers the faster network fragmentation with respect the same nodes removal fraction performed by the CondWBet (Figure 4). In Figure 4, we can see that the CondBet strategy is able to fragment the network isolating a large cluster, thus producing a sharper LCC decrease. This would suggest that, for the networks analyzed here, adding information about link weights may degenerate the efficacy of the attack strategies to select important nodes

TABLE 2: The most efficient nodes attack strategy for the five real-world weighted networks.

Network	Best strategy to reduce R_{LCC}	Best strategy to reduce R_{WEFF}
SP500_1	CondBet	CondWBet
SP500_2	CondBet	CondBet
NetScience	CondBet	Strength
Bitcoin	CondBet	CondWBet
USAirport	CondBet	CondWBet

(a)

(b)

(c)

(d)

(e)

(f)

(g)

(h)
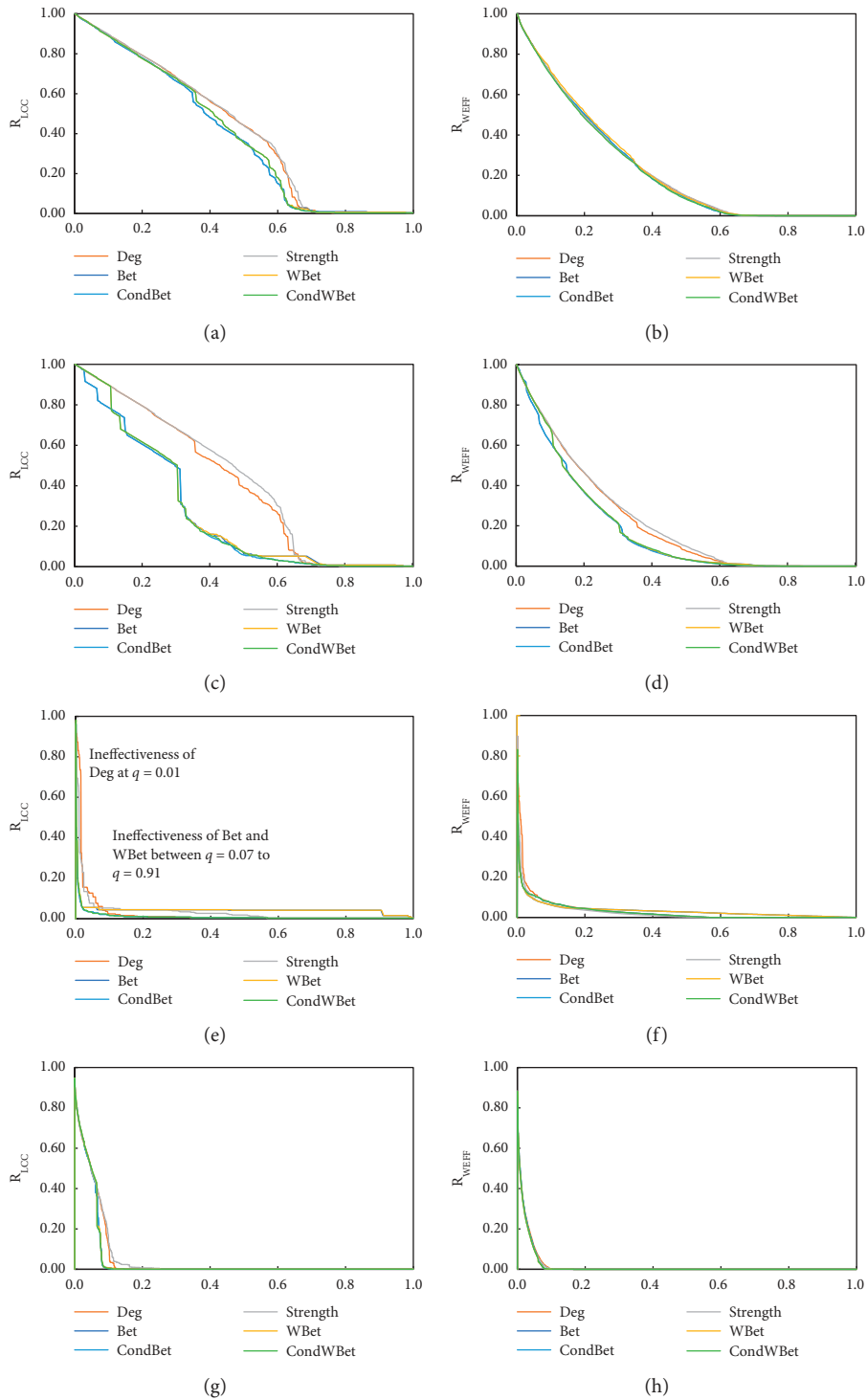
FIGURE 3: Continued.

(i)

(j)
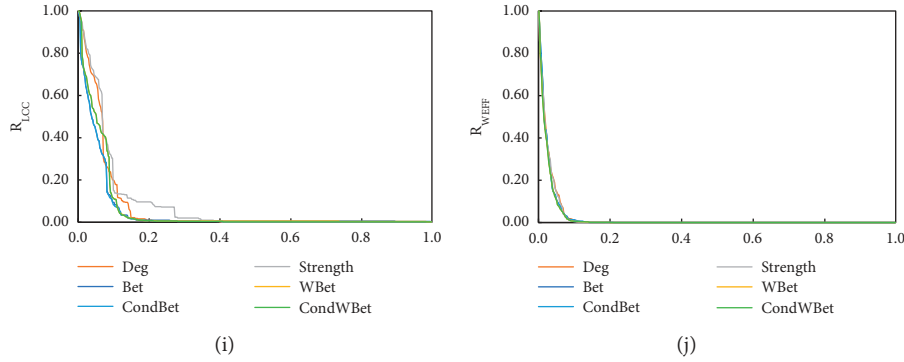
FIGURE 3: Normalized largest connected cluster (LCC) and weighted efficiency (WEFF) as a function of the fraction of removed nodes (q) under different attack strategies: (a, b) the SP500 threshold network from Feb. 1993 to Feb. 1996 (SP500_1), (c, d) the SP500 threshold network from May 1999 to May 2002 (SP500_2), (e, f) the co-authors network (NetScience), (g, h) the Bitcoin trust network, and (i, j) the USAirport network.
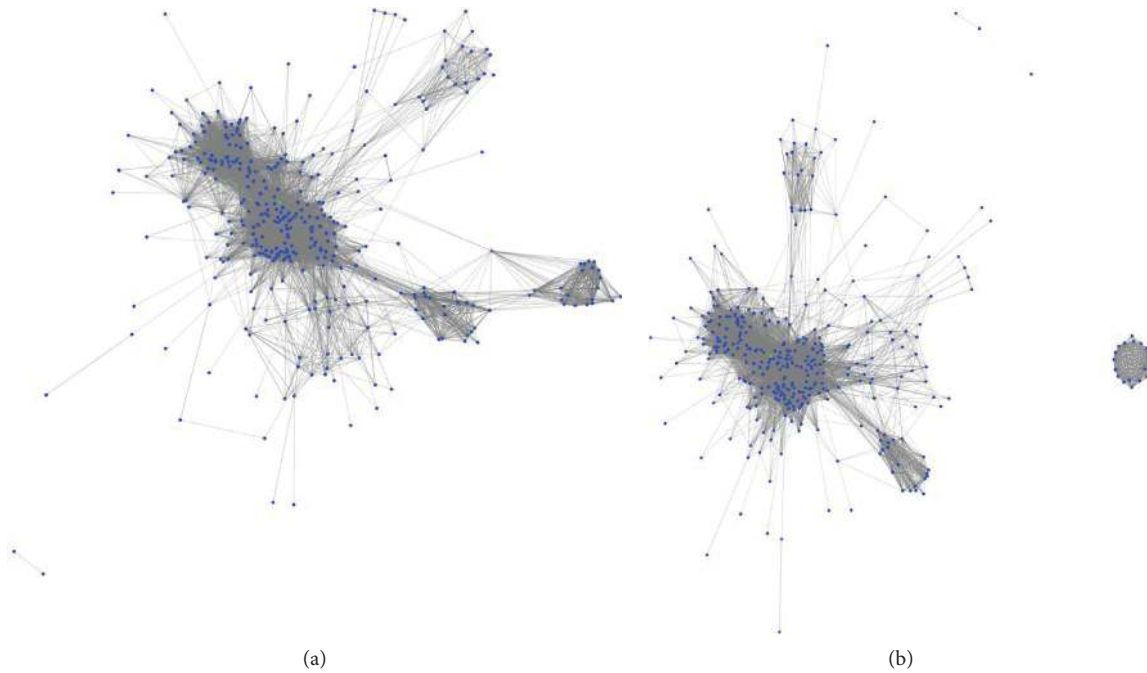

(a)   (b)

FIGURE 4: Topological structure of the SP500_2 network at $q = 4\%$ attacked by (a) the CondWBet strategy and (b) the CondBet strategy. The CondBet strategy is more effective to fragment the network isolating a large cluster, thus producing a sharper LCC decrease than the weighted counterpart CondWBet.

supporting the simple binary-topological connectivity (measured by the LCC). In fact, the link weights structure may induce the weighted attack strategies to remove nodes; hence, it is important for network functioning, not playing a major role in shaping the topological network connectivity. As we may expect, this effect is more important for networks with high $<r>$ value (the SP500_1, SP500_2, and the USAirport).

6. Efficacy of the Attack Strategies with WEFF

When the goal is to minimize R_{WEFF}, the CondWBet strategy is the best strategy in 3 out of 5 networks (Figure 3, Table 2). The CondWBet is highly effective to reduce the network

efficiency (WEFF), because it is able to remove the nodes passing the most of the weighted shortest paths in the network (equation (2)). For this reason, when removing the nodes with highest weighted betweenness, we trigger the disruption of many weighted shortest paths in the network with a quick WEFF decrease. Furthermore, CondWBet is more efficient than the classic betweenness strategy WBet for all networks. This can be explained by the ability of CondWBet to select nodes only inside the LCC at the end of the removal process when the LCC may be fully connected, thus providing further efficacy with respect classic WBet.

For the SP500_2, the efficacy difference between CondBet and CondWBet to reduce WEFF is very low, with
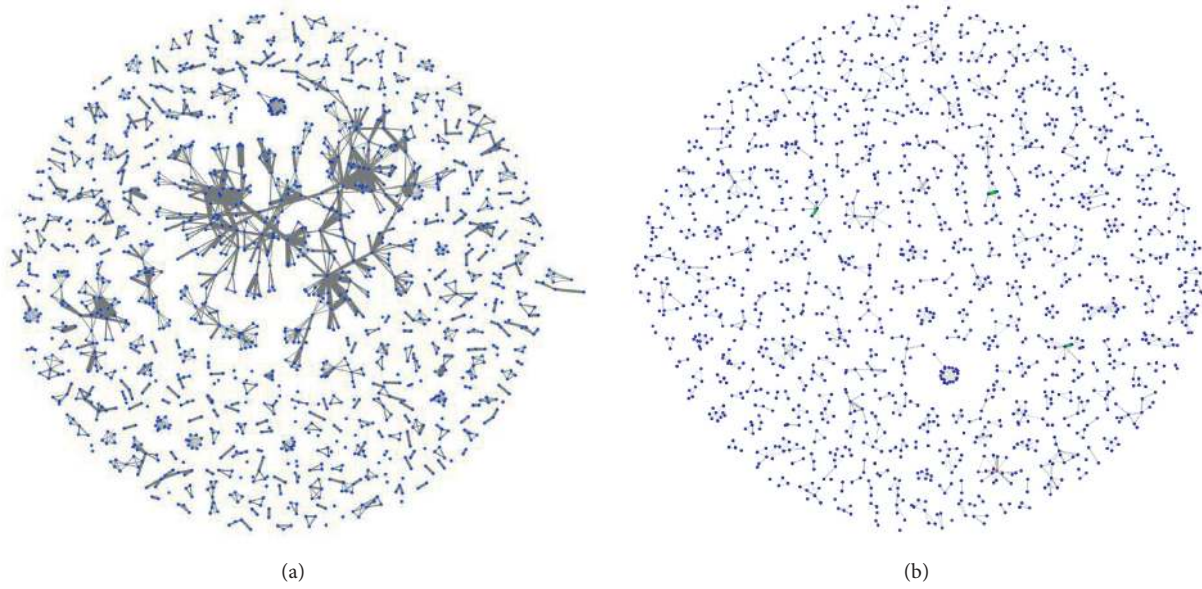
(a)

(b)

FIGURE 5: Topological structure of the NetScience network at (a) the beginning and (b) following the $q = 2\%$ of node removal by Strength strategy. The thicker links in the left panel indicate the links of highest weight (strong links representing higher number of common papers) occurring between most prolific scholars. The removal of the 2% of the higher strength nodes produces both the network fragmentation and the deletion of the strong links thus inducing an abrupt WEFF decrease.
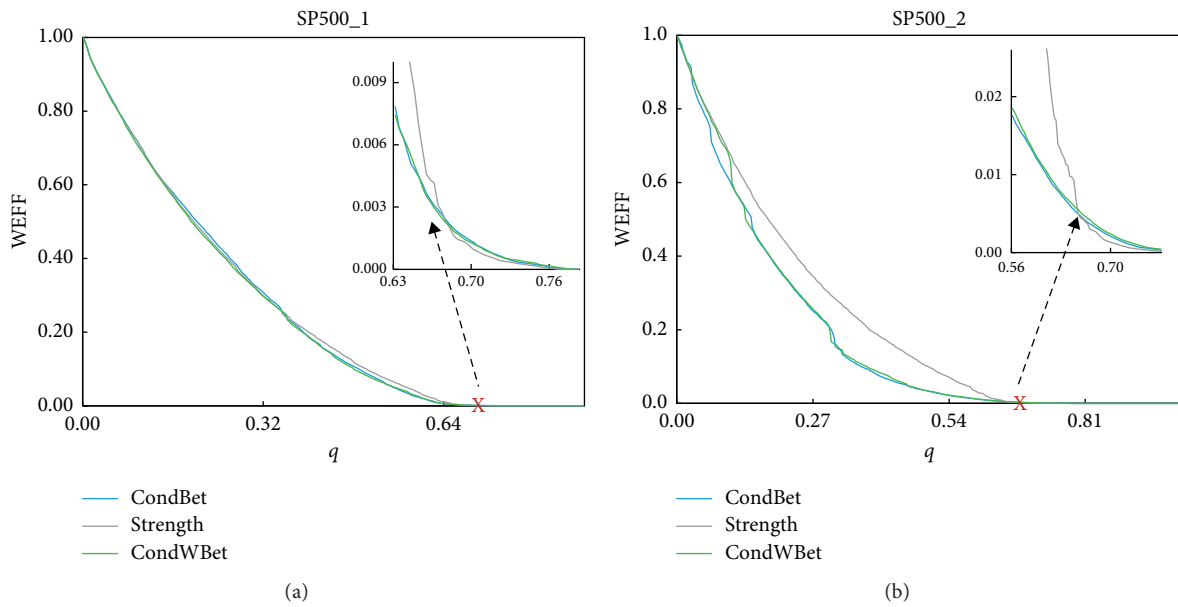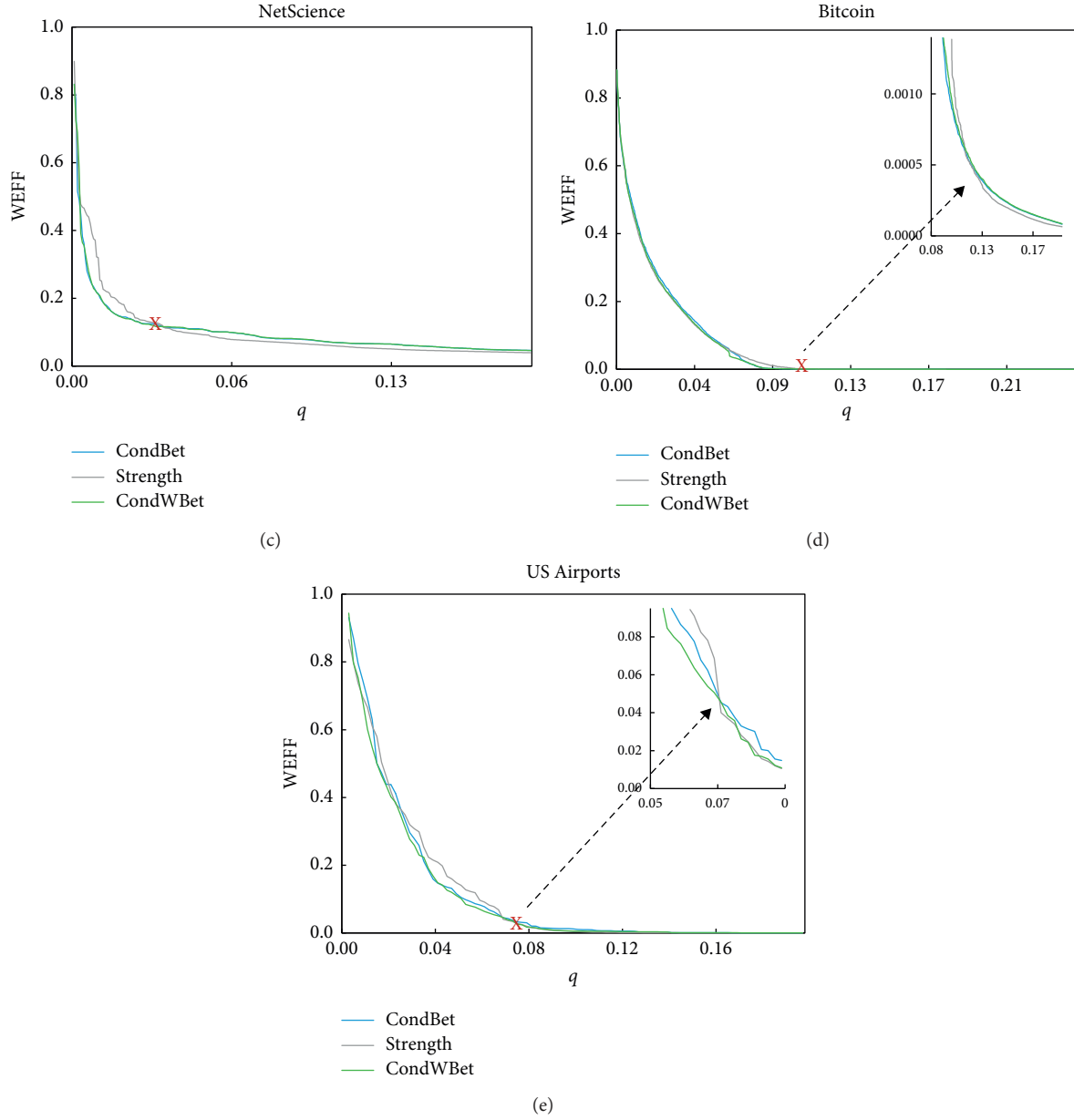

(a)

(b)

FIGURE 6: Continued.

(c)

(d)

(e)

FIGURE 6: The weighted efficiency (WEFF) as a function of the fraction of removed nodes (q) under three most efficacy attack strategies: Strength, CondBet, and CondWBet. (a) SP500_1, (b) SP500_2, (c) NetScience, (d) Bitcoin, and (e) US Airports. The transition between phases is marked by the red crosses. The inserts show a magnification of the transition (except for NetScience where this transition is clear).

TABLE 3: Fraction of removed nodes q range corresponding to the transition when the Strength strategy becomes more efficient than other strategies.

	SP500_1	SP500_2	NetScience	Bitcoin	USAirport
Approximate range of the fraction of removed nodes q when the transition happens	0.67	0.66	0.04	0.12	0.07
Overall highest efficacy strategy for reducing R_{WEFF}	CondWBet	CondBet	Strength	CondWBet	CondWBet

The overall highest efficacy strategy with minimum R_{WEFF} for each real-world networks is shown in the last row. See Figures 4–5 for more details.

almost negligible advantage of the CondBet strategy (Figures 2 and 3). For this reason, we can consider CondBet and CondWBet roughly equally performing, with negligible difference due to some less important effect.

Differently, in the NetScience network, the Strength is clearly the best strategy to decrease WEFF. The highest efficacy of the Strength strategy for the NetScience network can be explained by the peculiar embedding of the highest
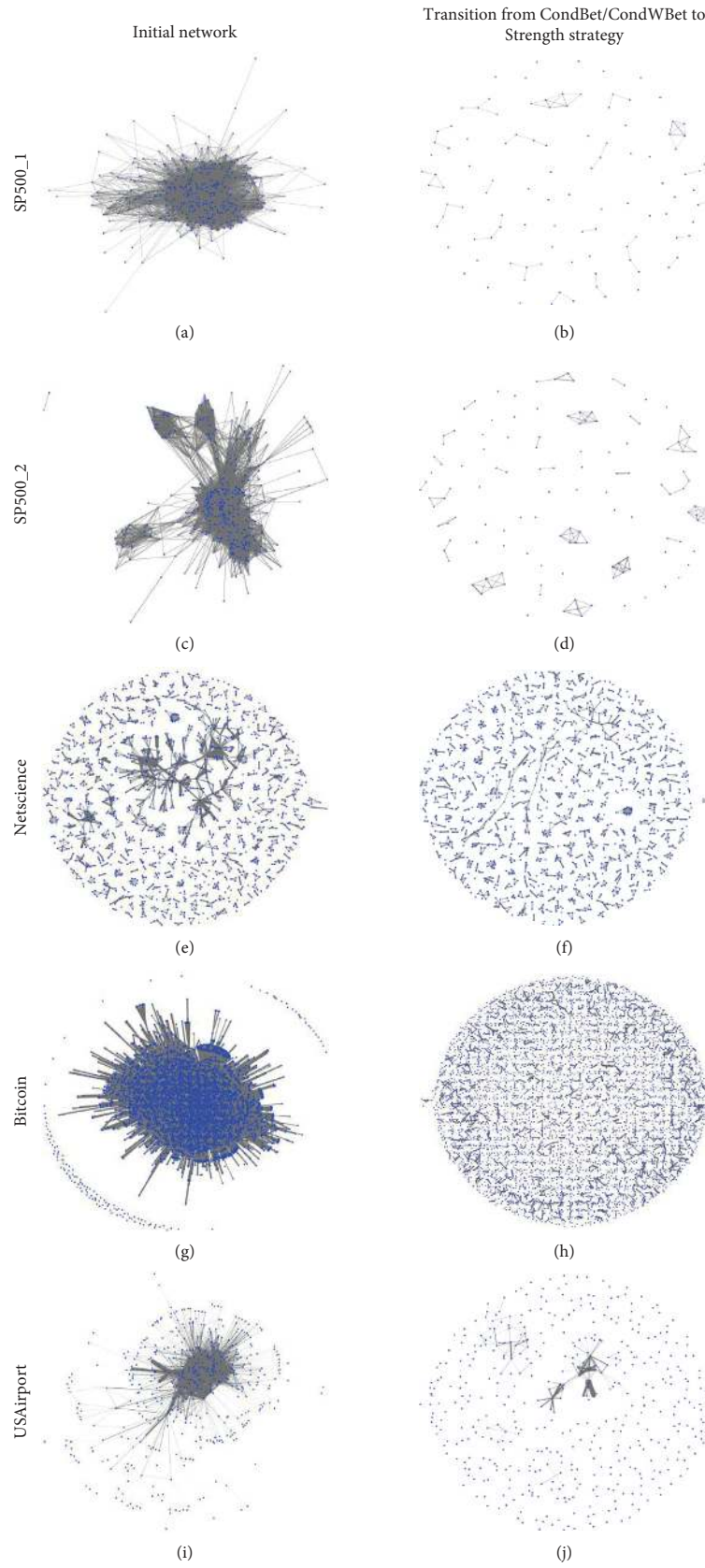
FIGURE 7: The topological image of the each networks in the beginning of the removal process where the conditional betweenness strategies CondBet/CondWBet are the most efficacy strategies (left column); at the transition point when the network are broken in many small clusters and the Strength strategy becomes the most efficacy for the remaining nodes removal process (right column).
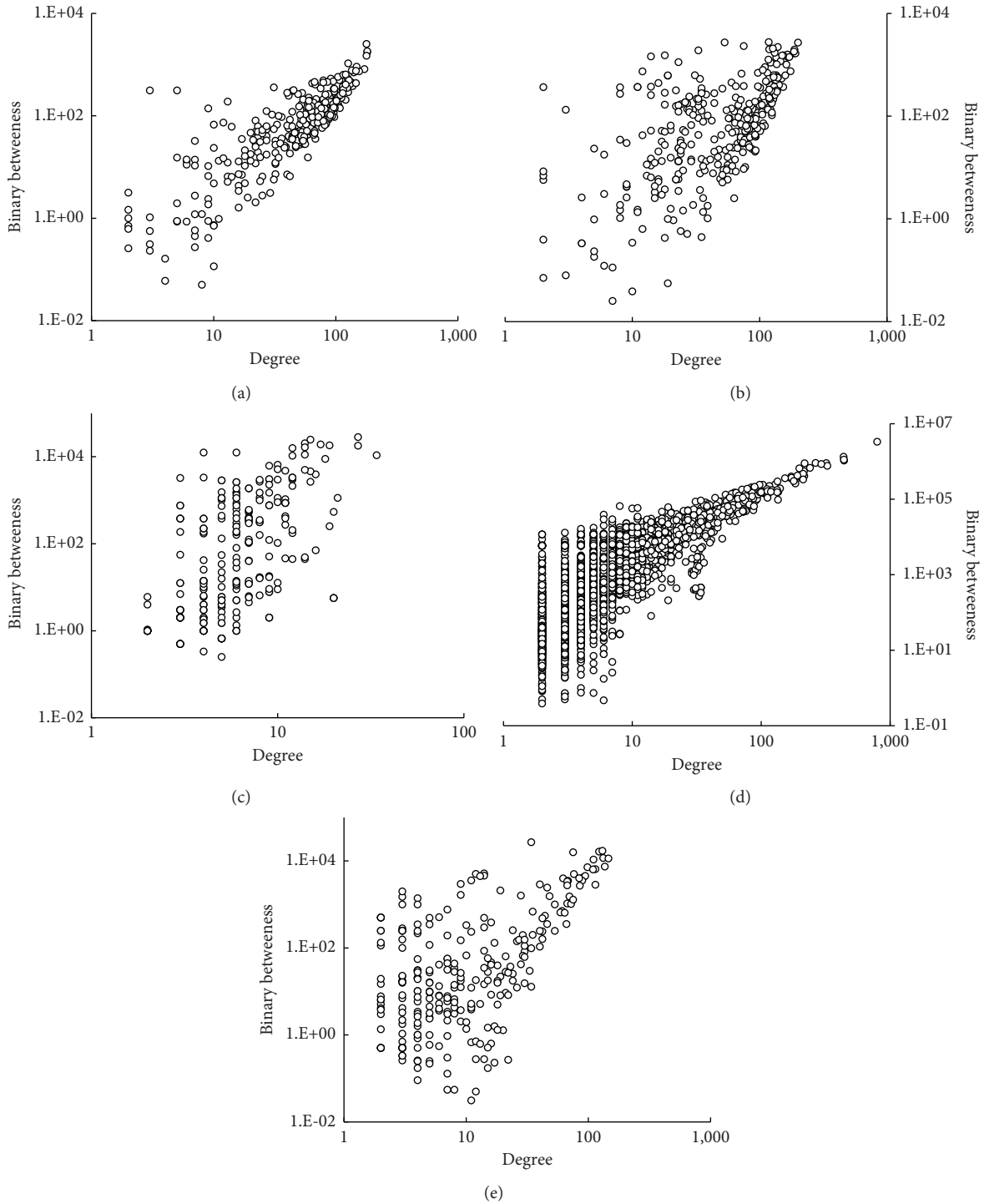
Figure 8: Scattered plot of binary betweenness and degree for (a) SP500_1 network, (b) SP500_2 network, (c) NetScience network, (d) Bitcoin network, and (e) Airport network.

weight links (strong links). The NetScience is the social network of co-authorship, where nodes are scholars, and links weights indicate the number of common papers. In this network, the strong links occur between senior and most prolific scholars leading different research groups; i.e., the strong links act as bridges between different research communities [36, 38]. The nodes of higher strength are these senior scholars publishing many papers and holding the collaborations with different research groups. For this reason, when removing nodes of higher strength in the NetScience network, we remove strong links (playing a major role into shaping WEFF) and at the same time, we delete the bridge links between communities, producing abrupt collapses of WEFF (Figure 5).
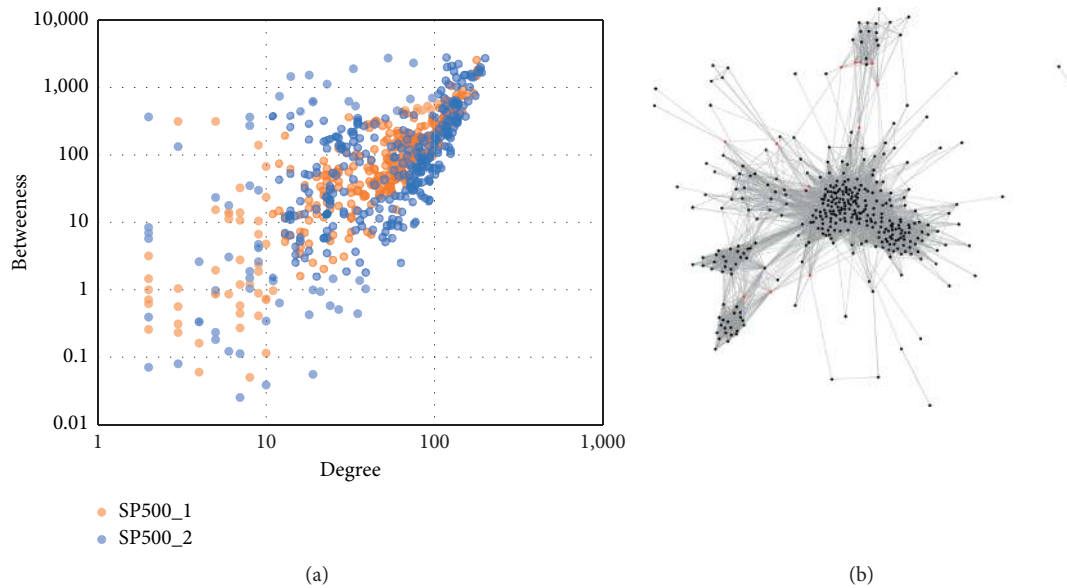
(a)   (b)

FIGURE 9: (a) Binary betweenness vs. degree of the two financial networks, the SP500_1 and SP500_2 where in the SP500_2 we see a group of high betweenness-medium degree nodes of the SP500_1 (highlighted by the rectangle). (b) Those nodes are colored red in the corresponding topological graph and appear to be bridge points of the network.
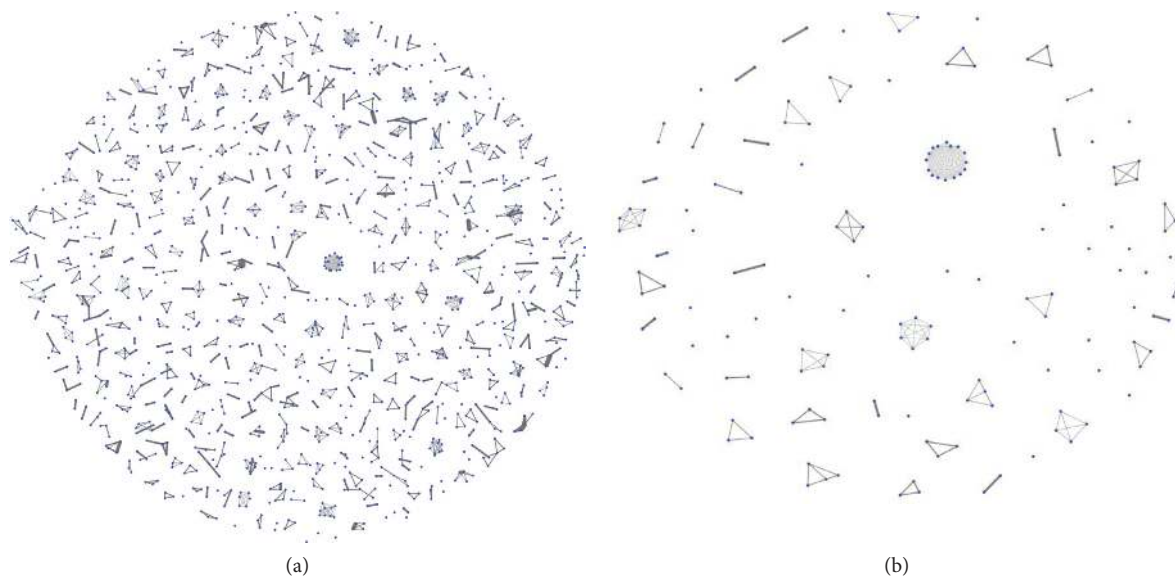

(a)   (b)

FIGURE 10: The NetScience network attacked by the Bet strategy at (a) $q = 7\%$ when the LCC is a large complete graph of 16 nodes and the Bet strategy will ignore it until at (b) $q = 91\%$ when the remaining network only contain individual nodes or complete graph.

Last, we find an interesting "efficacy transition" between CondBet/CondWBet and Strength strategies for all the networks; i.e., at the beginning of the removal process, CondBet/CondWBet are the best strategies, whereas, at the end of the removal process, Strength strategy shows the highest efficacy into reducing WEFF (Figure 6). This may be explained by the fact that when the network is broken apart in many small clusters, the nodes betweenness-based strategies would become ineffective to produce a significant further network fragmentation. For this reason, at this stage by selecting the highest strength nodes, we can intercept the remaining strong links that play the major contribution in shaping the network efficiency (WEFF). For example, USAirport network is mostly broken and contains a large number of small clusters as soon as $q = 7\%$ (Figure 6). At this stage, the betweenness-based strategies CondBet/CondWBet lose their global ability to intercept nodes bridge, and the Strength strategy that removes nodes with highest link weights becomes the most efficacy strategy for reducing WEFF. Thus, before the transition, the CondBet/CondWBet are the most efficacy strategies and after this
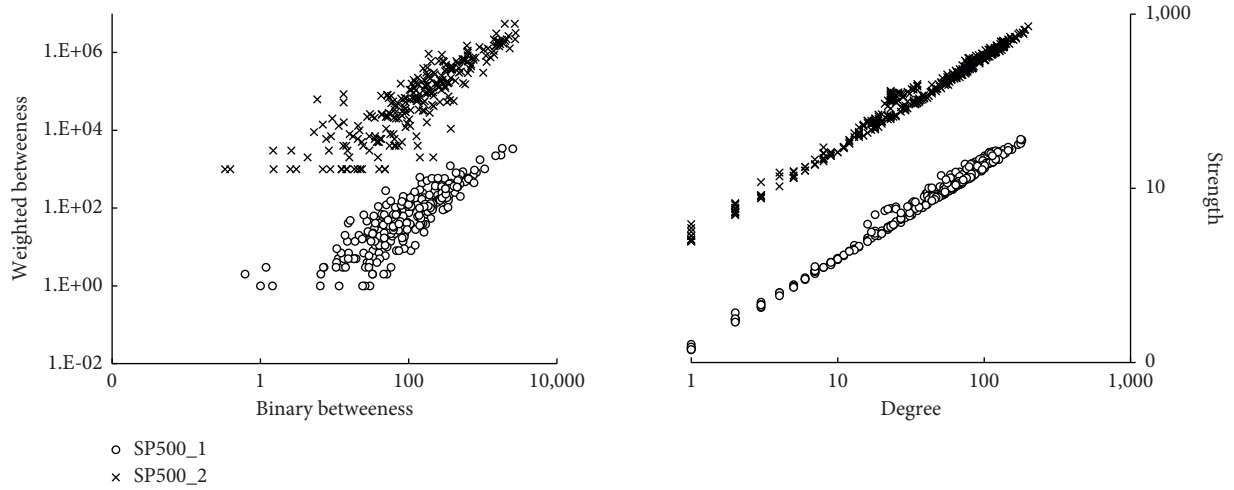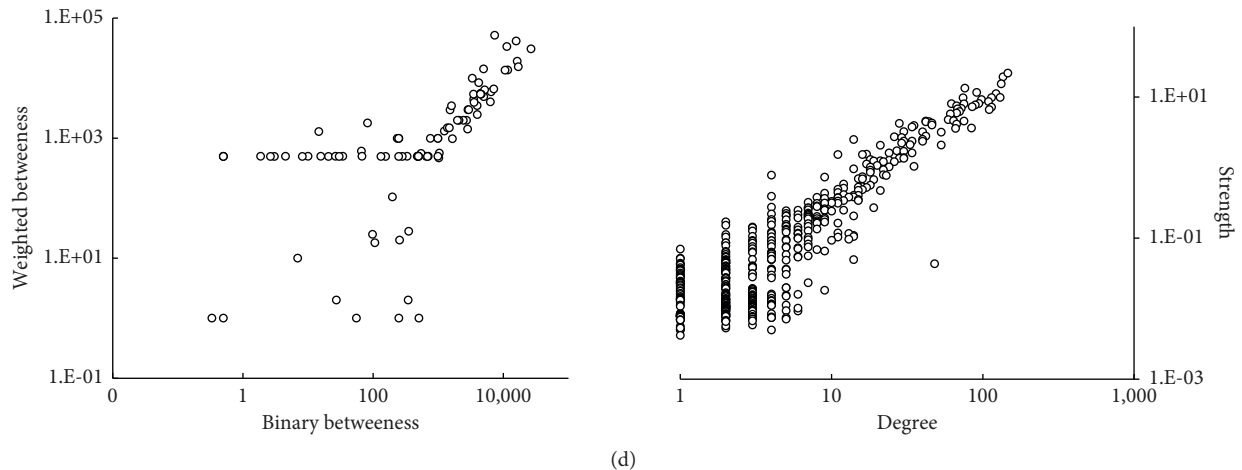
(a)
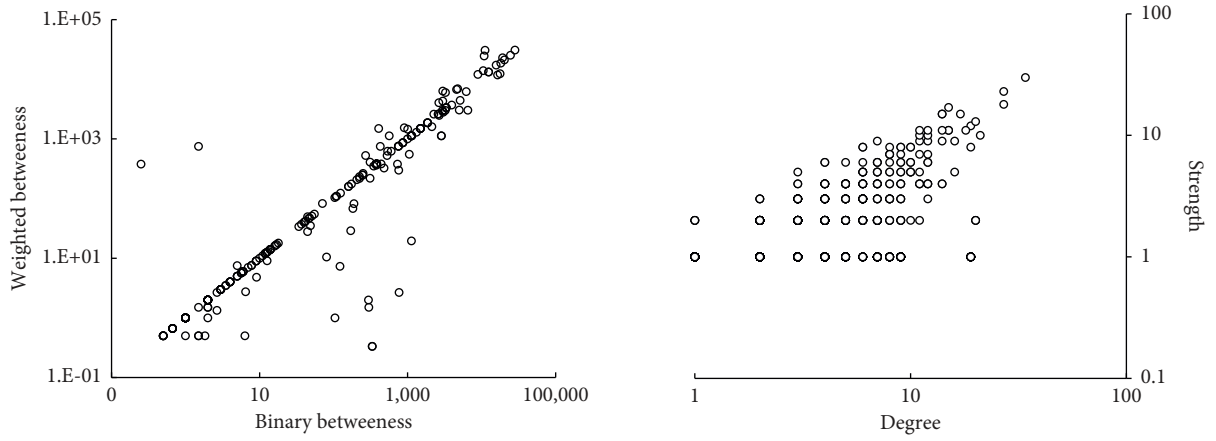
(b)
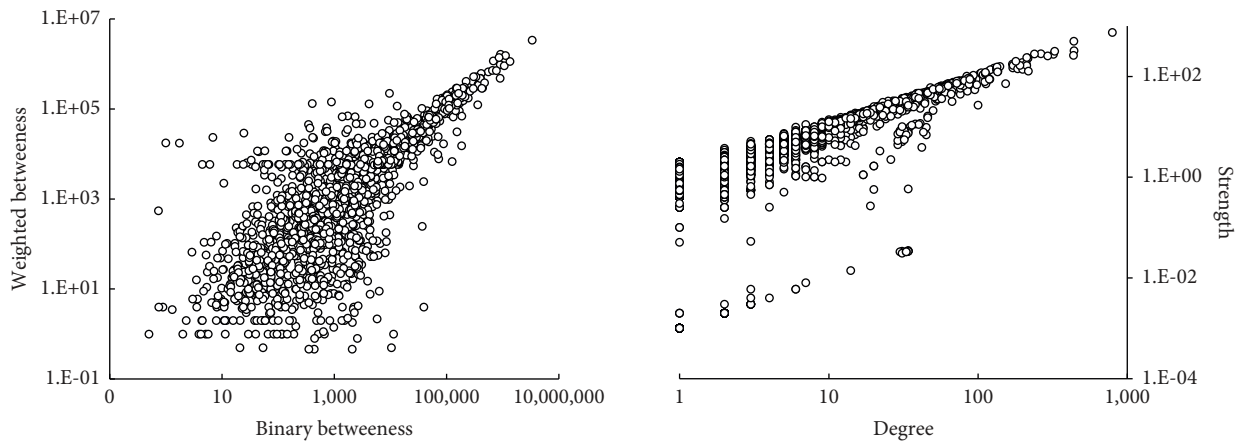
(c)

Figure 11: Continued.

(d)

FIGURE 11: Scattered plot of weighted betweenness and binary betweenness, Strength and degree for (a) two financial networks (with shifted scale for the SP500_2, (b) NetScience network, (c) Bitcoin network, and (d) USAirport network.

transition, when the network is broken into many small clusters, the Strength strategy is the best one for the remaining removal process. This situation can be better illustrated by a sample network as shown in Figure 1. We have a small network of 9 nodes, with node A being less connected than node B, but its links have much higher weight than those of node B. If attacked by the CondBet/CondWBet, node B will be removed; however, the remaining WEFF is still higher than that if node A is removed, because of higher strength of the node A. In such case, the most efficient strategy must be the Strength strategy that removes node A of higher strength in the network.

As a consequence, we can infer a general result for which, when the network is broken enough, the Strength strategy may be one of the best strategies to decrease the network efficiency (WEFF). The transition point values are summarized in Table 3 and the topological images of each network at this point are shown in Figure 7. Further observations that are specific to each real-world network are presented in the Appendix.

7. Conclusion

In this work, we studied the efficacy of different nodes attack strategies on real-world complex networks adopting different measures of the network functioning, both accounting the binary-topological connectedness and the link weights structure of the network. We used both classic topological-binary attacks and introduced new attacks based on weighted properties of the nodes and found that the recently introduced conditional attack strategy (CondBet) still outperforms the other strategies to break the LCC in all the 5 networks, confirming the highest effectiveness of this strategy [22]. Furthermore, the inclusion of link weighs into nodes attack strategy results in a lower efficacy when the goal is to reduce the LCC. This would suggest that adding information about link weights may degenerate the efficacy of the attack strategies to select important nodes supporting the simple binary-topological connectivity (LCC).

Secondly, when measuring the network functioning with the network efficiency WEFF, we find that, in 3 out of 5 networks, the newly introduced conditional weighted betweenness strategy CondWBet outperforms all other strategies, showing that, with the aim to select the most important nodes in real networks, it is necessary to consider the link weights. Further analysis shows that if the target is to decrease WEFF, the node attack process can be divided into two phases: in the first phase, when the network is well connected, the CondBet/CondWBet are the most efficient. If the correlation between the links' weight and the node betweenness is low (i.e., high $<r>$), and the network modularity is high, the inclusion of weight may avoid the removal of important bridge node by CondWBet, decreasing the efficacy of this strategy; in the second phase, the Strength shows the highest efficacy because when the network is broken into many small clusters, the Strength strategy may break the most important cluster, thus being the best one to decrease the network efficiency (WEFF). The best strategy would be therefore a result of the balance between these two node attack strategies.

These last outcome would suggest that a better attacking strategy for real-world weighted networks could be a combination of CondBet/CondWbet and Strength. However, an analytic model is necessary to determine the transition time and construct the combined strategy. Our work may help best design the attack strategy, or inversely design a more robust weighted network structure in practice.

Appendix

In this Appendix, we discuss additional observations that are specific to the real-world networks, as a complement to the Result section.

A. Financial Networks (SP500_1 and SP500_2)

All strategies seem relatively inefficacy to break the LCC network SP500_1 in relative to other networks. The size of

the LCC only decreases linearly with the fraction of nodes removed q; i.e., the strategies are not able to break apart the networks and the LCCs linearly decrease as the effect of the node removal itself. Only until $q = 35\%$ then the conditional betweenness strategies (CondBet and CondWBet) can have some efficacy producing the disconnection of network nodes from the LCC (Figure 5(b)), and now they are able to lay below the linear decrease of the LCC (Figure 3(a)). Only until $q = 63\%$ then the conditional strategies can reduce the LCC to 10% of its initial size (Figure 1(a)). Thus, in the financial network SP500_1, the CondBet and CondWBet are clearly the most efficient than all other strategies with significantly better performance to break the LCC than the related Bet and WBet strategies. The difficulty of the nodes attack strategies into breaking apart the LCC can be due to the fact that the network is highly and relatively uniformly connected around the main bulk (Figure 7(a)).

In SP500_2, all strategies perform better than those in SP500_1, except for the degree-based strategies (Figure 7(b)). The reason can be seen from the topological image of this network as it shows higher communities effect (in Figure 7(c)). This structure has twofold consequences: the betweenness-based strategies are efficient because they remove the pivotal nodes connecting different communities first, thus breaking the network apart and reducing the LCC quickly. Differently, the degree-based strategies (Deg and Strength) are inefficient because high degree nodes are usually inside a dense community, and their removal does not trigger the fragmentation of the network in different clusters. This can be seen from the scattered plot of the initial betweenness vs degree of this network in Figure 8(b). We found that, in the SP500_2, there are groups of nodes with medium degree but very high betweenness. Those nodes playing the role of connecting the network are quickly removed by the betweenness-based strategy, but they are not selected in principle by the degree-based strategies. Consequently, these "higher betweenness-medium degree nodes" playing a major role in connecting the network are secondarily removed during the degree-based attack process producing a slower LCC decrease. We show the higher betweenness-medium degree nodes owing a pivotal role in connecting the network in Figures 9(a) and 9(b). We found a small difference in the efficacy of the nodes attack strategies into decreasing WEFF in the SP500_1 financial network (Figure 3(b)). Nonetheless, the CondWBet is the most efficient to decrease WEFF until $q = 67\%$, and then the Strength is more efficient (Figure 6(a)). A similar observation was found with SP500_2 (at $q = 67\%$ as shown in Table 3).

B. NetScience Network

We found all the attack strategies very efficient in reducing the LCC. This higher efficacy of the nodes attack strategies can be explained by the fact that the NetScience is the least connected network among the four studying networks. It has been shown that the highest linkage density level positively affects the robustness of the networks connectivity; i.e., networks with a higher number of link per node experienced a slower decrease of the LCC when subjected to nodes removal [3]. The NetScience network owns an average degree

$<k>=1.72$ equal to one-fifteen of the average degree of the two financial networks (Table 1), and the removal of a small fraction of nodes is able to trigger the network fragmentation with faster LCC decrease.

Again, we found that the conditional strategies outperform all the other strategies for reducing the LCC. Especially, when the Bet and WBet become inactive for a very long time, as soon as $q = 7\%$ until $q = 91\%$, the Bet and WBet strategies are able to break the largest LCC. This is an interesting situation because at $q = 7\%$, a complete (fully connected) graph whose size is 16 becomes the LCC (Figure 10(a)). As the remaining network still has 1477 nodes, the Bet and WBet will completely ignore the LCC until all the clusters that are not fully connected are removed (Figure 10(b)). The situation is different for the Deg and Strength strategies: at $q = 1\%$, the same 16-node cluster already contains nodes with the highest degree and is broken by the Deg and Strength, although it is not the LCC. This is why the Deg and Strength strategies are ineffective at the early stage as shown in Figure 3(e).

For the weighted efficiency measure WEFF, we found that, for this Netscience network, the Strength strategy is the most efficient as soon as $q = 4\%$ (Figure 6(c)). In consequence, the Strength strategy is the most efficient and significantly better than the second best CondBet one overall.

C. Bitcoin Network

Similar to the NetScience network, all strategies perform well into decreasing the LCC (Figure 3(g)). It is also because the Bitcoin network is sparse (low number of links per node) with an average degree $<k>=3.58$. We also found the inefficacy of the Bet and WBet strategies when the network is broken enough (to a lesser extent than the NetScience). At this later stage of the removal process, the Bet and WBet are not able to intercept nodes into the LCC in the network presenting a fully connected LCC and many clusters, and as a consequence, the conditional strategy CondBet that always removes nodes inside the LCC is the most efficient one to make it vanish. For the weighted efficiency measure WEFF, similar to the Netscience, the second phase starts as soon as $q = 12\%$ (Figure 6(d)). However, because of the large gain in efficacy in the first phase by the CondWBet strategy, the overall best strategy is still the CondWBet strategy.

D. USAirport Network

All strategies perform well into decreasing the LCC; in particular, the betweenness-based strategy Bet and CondBet strategies outperform the others significantly (Figures 2(b) and 3(i)).

Again, the CondBet strategy can break the complete graph at the end of the removal process, resulting in a higher efficacy than the Bet strategy. We found that, for this network, the inclusion of weights into strategies worsens the efficacy in reducing the LCC for all the three strategies: Strength, WBet, and CondWBet (Figures 2(a) and 3(i)). This can be seen from the scattered plot of the binary betweenness vs weighted betweenness of this network in

Figure 11(d) where one finds that their correlation is the worst among networks. This may come from the fact that high binary betweenness node/airport may represent a geographically important transit point. Meanwhile, when counting the passenger flow for weighted betweenness, other airports, although less well located, can still be higher ranked in terms of passenger transit. For the weighted efficiency measure WEFF, both degree-based strategies Deg and Strength improve significantly. This suggests that high weight nodes may have important contribution in the efficiency measure of this network. However, CondWBet is still slightly better than the Strength for the whole removal process for reducing WEFF. As shown in Figure 7(i), the USAirport network has a high degree of communities and therefore, CondWBet works efficiently. In other words, the US 500 airports network would be most affected if the top 7% (35 airports) of the most transit point is closed. Then, from this point, the whole network is almost broken (Figure 7(j)), and the Strength strategy based on the total number of passenger flow will become the most efficient again.

Abbreviations

CondWBet: The conditional weighted betweenness node attack strategy
CondBet: The conditional betweenness node attack strategy
WEFF: The network weighted efficiency
LCC: The network largest connected cluster
Deg: The degree-based node attack strategy
Bet: The betweenness node attack strategy
Strength: The strength node attack strategy
WBet: The weighted betweenness node attack strategy
SP500_1: The US daily stock price network (built from data between Feb. 1993 and Feb. 1996)
SP500_2: The US daily stock price network (built from May 1999 to May 2002)
NetScience: The co-authorship undirected network of scientists working on network theory and experiments
Bitcoin: The network of 6005 nodes-peoples who trade using Bitcoin on a platform called Bitcoin OTC
USAirport: The network of 500 busiest airports in USA
N: The network nodes' number
L: The network links' number
$<k>$: The network's average node degree
$<s>$: The network's average node strength
$\langle w \rangle$: The network's average link weight
R: The network robustness
R_{LCC}: The network robustness measured by the LCC
R_{WEFF}: The network robustness measured by the WEFF
WSP: Weighted shortest path.

Data Availability

Historical daily close price of the SP500 index was downloaded from the website finance.yahoo.com. The cross-correlation matrix and then the threshold networks SP500_1 and SP500_2 are prepared as described in [22]. The final network datasets can be provided upon reasonable request. The NetScience network dataset is downloaded from Newman's website http://www-personal.umich.edu/~mejn/netdata/. The Bitcoin network dataset is downloaded from https://snap.stanford.edu/data/ [31, 32]. The 500 US busiest airports network dataset is downloaded from Tore Opsahl's website https://toreopsahl.com/datasets/ [33].

Conflicts of Interest

The authors declare no conflicts of interest.

Authors' Contributions

QN and MB conceived the analyses. QN and NKKN performed the simulation. QN, MB, and DC wrote the manuscript.

References

[1] L. d. F. Costa, O. N. Oliveira, G. Travieso et al., "Analyzing and modeling real-world phenomena with complex networks: a survey of applications," *Advances in Physics*, vol. 60, no. 3, pp. 329–412, 2011.

[2] A. S. d. Mata, "Complex networks: a mini-review," *Brazilian Journal of Physics*, vol. 50, no. 5, pp. 658–672, 2020.

[3] R. Albert and A.-L. Barabási, "Statistical mechanics of complex networks," *Reviews of Modern Physics*, vol. 74, no. 1, pp. 47–97, 2002.

[4] R. Cohen, K. Erez, D. Ben-Avraham, and S. Havlin, "Resilience of the Internet to random breakdowns," *Physical Review Letters*, vol. 85, no. 21, pp. 4626–4628, 2000.

[5] D. S. Callaway, M. E. J. Newman, S. H. Strogatz, and D. J. Watts, "Network robustness and fragility: percolation on random graphs," *Physical Review Letters*, vol. 85, no. 25, 2000.

[6] M. Bellingeri, E. Agliari, and D. Cassi, "Optimization strategies with resource scarcity: from immunization of networks to the traveling salesman problem," *Modern Physics Letters B*, vol. 29, no. 29, Article ID 1550180, 2015.

[7] M. Bellingeri, D. Cassi, and S. Vincenzi, "Efficiency of attack strategies on complex model and real-world networks," *Physica A: Statistical Mechanics and its Applications*, vol. 414, pp. 174–180, 2014.

[8] L. Dall'Asta, A. Barrat, M. Barthélemy, and A. Vespignani, "Vulnerability of weighted networks," *Journal of Statistical*

Mechanics: Theory and Experiment P04006, vol. 2006, no. 4, p. P04006, 2006.

[9] S. Iyer, T. Killingback, B. Sundaram, and Z. Wang, "Attack robustness and centrality of complex networks," *PLoS One*, vol. 8, no. 4, Article ID e59613, 2013.

[10] N. K. Nguyen and Q. Nguyen, "Resilience of stock cross-correlation network to random breakdown and intentional attack," in *Econometrics for Financial Applications in Studies in Computational Intelligence*, pp. 553–561, Springer, Berlin, Germany, 2018.

[11] S. Sun, X. Liu, L. Wang, and C. Xia, "New link attack strategies of complex networks based on k-core decomposition," *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 67, no. 12, pp. 3157–3161, 2020.

[12] S. Wandelt, X. Sun, D. Feng, M. Zanin, and S. Havlin, "A comparative analysis of approaches to network-dismantling," *Scientific Reports*, vol. 8, no. 1, Article ID 13513, 2018.

[13] R. Albert, H. Jeong, and A.-L. Barabási, "Error and attack tolerance of complex networks," *Nature*, vol. 406, no. 6794, pp. 378–382, 2000.

[14] P. Holme, B. J. Kim, and C. N. Yoon, S. K. Han, Attack vulnerability of complex networks," *Physical review. E, Statistical, nonlinear, and soft matter physics*, vol. 65, no. 5 Pt 2, Article ID 056109, 2002.

[15] C. Li, L. Wang, S. Sun, and C. Xia, "Identification of influential spreaders based on classified neighbors in real-world complex networks," *Applied Mathematics and Computation*, vol. 320, no. C, pp. 512–523, 2018.

[16] F. Morone and H. A. Makse, "Influence maximization in complex networks through optimal percolation," *Nature*, vol. 524, no. 7563, pp. 65–68, 2015.

[17] J. Wang, C. Li, and C. Xia, "Improved centrality indicators to characterize the nodal spreading capability in complex networks," *Applied Mathematics and Computation*, vol. 334, pp. 388–400, 2018.

[18] A. E. Motter, "Cascade control and defense in complex networks," *Physical Review Letters Review*, vol. 93, Article ID 098701, 2004.

[19] P. Crucitti, V. Latora, M. Marchiori, and A. Rapisarda, "Efficiency of scale-free networks: error and attack tolerance," *Physica A: Statistical Mechanics and its Applications*, vol. 320, pp. 622–642, 2003.

[20] S. Allesina and M. Pascual, "Googling food webs: can an eigenvector measure species' importance for coextinctions?" *PLoS Computational Biology*, vol. 5, no. 9, Article ID e1000494, 2009.

[21] B. Wang, H. Tang, C. Guo, and Z. Xiu, "Entropy optimization of scale-free networks' robustness to random failures," *Physica A: Statistical Mechanics and Its Applications*, vol. 363, no. 2, pp. 591–596, 2006.

[22] Q. Nguyen, H. D. Pham, D. Cassi, and M. Bellingeri, "Conditional attack strategy for real-world complex networks," *Physica A: Statistical Mechanics and its Applications*, vol. 530, Article ID 121561, 2019.

[23] H. Yin, A. R. Benson, J. Leskovec, and D. F. Gleich, "Local higher-order graph clustering," in *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining KDD '17*, pp. 555–564, Halifax, Canada, August 2017.

[24] J. Leskovec, J. Kleinberg, and C. Faloutsos, "Graph evolution," *ACM Transactions on Knowledge Discovery from Data*, vol. 1, no. 1, 2007.

[25] M. Ripeanu, I. Foster, and A. Iamnitchi, "Mapping the gnutella network: properties of large-scale peer-to-peer systems and implications for system design," *IEEE Internet Computing Journal*, vol. 6, pp. 50–57, 2002.

[26] M. Barthélemy, A. Barrat, R. Pastor-Satorras, and A. Vespignani, "Characterization and modeling of weighted networks," *Physica A: Statistical Mechanics and Its Applications*, vol. 346, no. 1-2, pp. 34–43, 2005.

[27] M. Bellingeri and D. Cassi, "Robustness of weighted networks," *Physica A: Statistical Mechanics and its Applications*, vol. 489, pp. 47–55, 2018.

[28] U. Brandes, "A faster algorithm for betweenness centrality∗," *The Journal of Mathematical Sociology*, vol. 25, no. 2, pp. 163–177, 2001.

[29] V. Boginski, S. Butenko, and P. M. Pardalos, "Statistical analysis of financial networks," *Computational Statistics & Data Analysis*, vol. 48, no. 2, pp. 431–443, 2005.

[30] M. E. J. Newman, "The structure and function of complex networks," *SIAM Review*, vol. 45, pp. 167–256, 2003.

[31] S. Kumar, H. Bryan, D. Makhija, M. Kumar, C. Faloutsos, and V. S. Subrahmanian, "REV2: fraudulent user prediction in rating platforms srijan," in *Proceedings of the Eleventh ACM International Conference on Web Search and Data Mining–WSDM '18*, Los Angeles, CA, USA, 2018.

[32] S. Kumar, F. Spezzano, V. S. Subrahmanian, and C. Faloutsos, "Edge weight prediction in weighted signed networks," in *Proceedings of the IEEE International Conference on Data Mining, ICDM*, Barcelona, Spain, December 2017.

[33] V. Colizza, R. Pastor-Satorras, and A. Vespignani, "Reaction-diffusion processes and metapopulation models in heterogeneous networks," *Nature Physics*, vol. 3, no. 4, pp. 276–282, 2007.

[34] S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, and D. Hwang, "Complex networks, structure and dynamics," *Physics Reports*, vol. 424, pp. 175–308, 2006.

[35] V. Latora and M. Marchiori, "Efficient behavior of small-world networks," *Physical Review Letters*, vol. 87, no. 19, Article ID 198701, 2001.

[36] M. Bellingeri, D. Bevacqua, F. Scotognella, R. Alfieri, and D. cassi, "A comparative analysis of link removal strategies in real complex weighted networks," *Scientific Reports*, vol. 10, no. 1, p. 3911, 2020.

[37] M. Bellingeri, D. Bevacqua, F. Scotognella, and D. Cassi, "The heterogeneity in link weights may decrease the robustness of real-world complex weighted networks," *Scientific Reports*, vol. 9, no. 1, Article ID 10692, 2019.

[38] R. K. Pan and J. Saramäki, "The strength of strong ties in scientific collaboration networks," *Europhysics Letters*, vol. 97, p. 18007, 2012.