Reachability-based Robustness of Controllability in Sparse Communication Networks

Sun, Peng; Kooij, Robert E.; Van Mieghem, Piet

DOI
10.1109/TNSM.2021.3082283

Publication date
2021

Document Version
Final published version

Published in
IEEE Transactions on Network and Service Management

Reachability-Based Robustness of Controllability in Sparse Communication Networks

Peng Sun, Robert E. Kooij, and Piet Van Mieghem, *Senior Member, IEEE*

***Abstract*—In this paper, we propose closed-form analytic approximations for the number of controllable nodes in sparse communication networks from the aspect of network controllability, considering link-based random attack, targeted attack, as well as random attack under the protection of critical links. We compare our approximations with simulation results on communication networks. Results show that our approximations perform well for all three attack strategies as long as the fraction of removed links is small. Only when the fraction of removed links is large, our approximation for targeted attacks does not fit well with simulation results. Finally, we validate our approximations using 200 communication networks and some synthetic networks. Results show that our approximations perform well in most cases.**

***Index Terms*—Network controllability, network robustness, reachability, communication networks.**

I. Introduction

IN RECENT years, the analysis of network controllability from a graph theoretic point of view has become an active area of research. Through the control of external inputs [1], a controllable system can be driven from any arbitrary state to any desired state in finite time. For example, a communication network can be controlled externally through input signals such as commands from control units connected to some of the work stations [2].

Most work regarding the robustness of controllability has focused on the number of controls required to maintain network controllability after link or node failures. Lou *et al.* [3] proposed a complex network model called q-snapback network which has the strongest robustness of controllability due to its advantageous inherent structure with many chain and loop motifs, when compared with the multiplex congruence network and the generic scale-free network. Pu *et al.* [4] found that the degree-based node attack is more efficient than a random failure for degrading the controllability in random and scale-free networks. Nie *et al.* [5] found that the controllability of Erdős-Rényi random graphs with a moderate average degree is not very robust, whereas a scale-free network with moderate power-law exponent shows a stronger ability to maintain its controllability, when these networks are under intentional link attack. Thomas *et al.* [6] identified that the potency of a degree-based attack is directly related to the betweenness centrality of the edges being removed. Chen *et al.* [7] evaluated the effect of the number of control inputs on the controllability for random networks and scale-free networks in the process of cascading failure. Lou *et al.* [8] proposed a framework of hierarchical attack by means of link- or node-removal attacks and suggest to protect the critical links and nodes to maintain network controllability. Yan-Dong *et al.* [9] proposed a method that modifies any given network with strict structural perturbation to make the network homogenous and effectively enhance its robustness against malicious attacks. Zhang *et al.* [10] optimized the robustness of interdependent network controllability by redundant design including node backup and edge backup. Sun *et al.* [11] proposed closed-form analytic approximations for the number of controls that are needed to maintain network controllability, where links are removed according to both random and targeted attacks.

The above work regarding the robustness of controllability assumes that the network operator has the capability to add additional controls at any location in the network in order to maintain the current network controllable after attacks or failures. In other words, the basic assumption of previous work mentioned above is that network operators have sufficient budget and quantity of resources that can be deployed in response to an attack or failure. However, a more realistic assumption is that network operators have a fixed budget and a limited quantity of resources. Moreover, the increase in additional controls is only a proxy for the most relevant information - how much of the network is still controllable (reachable) after an attack or failure. Parekh *et al.* [12] proposed the number of controllable nodes as a new metric to quantify the robustness of controllability under network perturbations. Thomas *et al.* [6] analyzed the changes in the controllability of synthetic networks from the perspective of reachability and found that scale-free networks evidence higher robustness to random failures than Erdős-Rényi networks. In this paper, we analyze and measure the robustness of network controllability in terms of reachability. In particular, we determine the maximum number of nodes that are still controllable when the number of driver nodes remains the same during the failure or attack process. Here, the driver nodes are the nodes into which the external control signals are directly injected.

Manuscript received October 31, 2020; revised February 15, 2021 and May 5, 2021; accepted May 15, 2021. Date of publication May 20, 2021; date of current version September 9, 2021. This research was supported by the China Scholarship Council (No. 201706220113). The associate editor coordinating the review of this article and approving it for publication was F. D. Turck. *(Corresponding author: Peng Sun.)*

The authors are with the Faculty of Electrical Engineering, Mathematics and Computer Science, Delft University of Technology, 2628 CD Delft, The Netherlands (e-mail: p.sun-1@tudelft.nl).

Digital Object Identifier 10.1109/TNSM.2021.3082283

This paper is organized as follows. In Section II, we introduce some basic concepts and definitions in network controllability proposed in [1]. In Section III, we analyse the role of critical links in network controllability. In Section IV, we compare the robustness of controllability for three cases: random attack, random attack under protection and targeted attack. In Section V, we propose analytic approximations for the number of controllable nodes N_c in these three cases and measure the accuracy of our approximations. Section VI concludes the paper.

II. Reachability-Based Robustness of Controllability

A. Controllability of Networks

Most real systems are driven by nonlinear dynamics. However, linear dynamics is a first step towards addressing the controllability of real-world systems and provides sufficient controllability conditions for particular nonlinear cases [1]. A system is controllable if it can be driven from any initial state to any desired final state by proper variable inputs in finite time [13]. We consider a linear, time-invariant system $G(A, B)$, which is described by:

$$\frac{d\boldsymbol{x}(t)}{dt} = \boldsymbol{A}\boldsymbol{x}(t) + \boldsymbol{B}\boldsymbol{u}(t) \tag{1}$$

where the $N \times 1$ vector $\boldsymbol{x}(t) = (x_1(t), x_2(t), \ldots, x_N(t))^T$ denotes the state of the system with N nodes at time t. The weighted matrix $\boldsymbol{A}$ is an $N \times N$ matrix which describes the network topology (graph) and the interaction strength between the components as nodes in the graph. In this work, we assume that there are no self-links in the networks, which is also assumed in [1], [14]. The $N \times M$ matrix B is the input matrix which identifies the $M \leq N$ driver nodes controlled by outside input signals. The $M \times 1$ vector $\boldsymbol{u}(t) = (u_1(t), u_2(t), \ldots, u_M(t))^T$ is the input signal vector. A driver node $j \in \{1, \ldots, M\}$ has an input signal u_j that is externally fed in $u_j(t)$. In [6] and [12], the jth external source which generates the input signal u_j is named control node or input node, while the N nodes in the network are named state nodes. Since each driver node is injected by an input signal generated by an external control node, the number of driver nodes equals the number of control nodes.

The linear system defined by equation (1) is controllable, if and only if the $N \times NM$ controllability matrix

$$\boldsymbol{C} = \left(\boldsymbol{B}, \boldsymbol{A}\boldsymbol{B}, \boldsymbol{A}^2\boldsymbol{B}, \ldots, \boldsymbol{A}^{N-1}\boldsymbol{B}\right) \tag{2}$$

has full rank, i.e., $rank(\boldsymbol{C}) = N$. Criterion (2) is called Kalman's controllability rank condition [15]. The rank of matrix $\boldsymbol{C}$ provides the dimension of the controllable subspace of the system. The input matrix $\boldsymbol{B}$, which determines the location of driver nodes, needs to be chosen properly to assure that the controllability matrix $\boldsymbol{C}$ has full rank.

So far, most of the existing studies on the robustness of controllability have measured the increase in the minimum number N_d of driver nodes required as a proxy for the reduction in controllability due to a failure. This indirect approach of measuring robustness is referred to as control-based robustness. The robustness of network controllability from the perspective of reachability is also considered by a few authors, see [6] and [12]. Besides, the control-based robustness analysis of network controllability assumes that the network operator has the capability to attach any amount of additional control signals to the nodes in the network. However, network operators normally have limited budget and resources in real life, which constrains the ability to deploy external controls. Based on these considerations, we focus on the reachability-based robustness of controllability, which determines the maximum number N_c of nodes that are still under control when failure or attack occurs, during which the number N_{d0} of driver nodes remains the same [12]. For the reachability-based controllability, there are two cases, namely free control and fixed control [6]. In the free control case, only the number N_{d0} of driver nodes remains the same, but the set of driver nodes can vary. In the fixed control case, both the number and the set of driver nodes are fixed during attacks or failures. In this paper, we only consider the free control case and delegate the fixed control to future research. For convenience, we use the term reachability to represent reachability-based controllability.

B. R-Value and Challenges

We inherit the framework and some definitions proposed for network robustness [16], [17] to investigate the robustness of reachability. The robustness of a given network determined by a service and an underlying topology is quantified by a robustness value, referred to as the R-value [16]. The R-value is normalized to the interval [0, 1]. Thus, $R = 1$ reflects complete functionality in an network without failures, and $R = 0$ corresponds to the complete absence functionality in a severely damaged network. The R-value can be a metric, which is related to network topology and service, such as the size of the giant component [18], the effective graph resistance [19] and network efficiency [20]. In this paper, we use the normalized maximum number of controllable nodes $n_c = N_c/N$ as the R-value. The number N_c of controllable nodes satisfies $N_{d0} \leq N_c \leq N$, thus $N_{d0}/N \leq n_c \leq 1$.

An elementary challenge is an event that changes the network and thus changes the R-value. We assume that a sequence of changes does not coincide in time. In this paper, we confine an elementary challenge to a link removal in a failure process. A perturbation is a series of m elementary changes, characterized by a sequence of m corresponding R-values $\{R[k]\}_{0<k\leq 1}$, where $k = m/L$ is the fraction of removed links, $m \in \{1, \ldots, L\}$ is the number of removed links and L is the number of links in the network. In this paper, we choose the maximum number n_c of controllable nodes as the R-value and observe the impact of link removal on n_c. As shown in Figure 1, the maximum number n_c of controllable nodes has a decreasing trend as links are removed one by one.

C. Robustness Envelopes

As discussed in the previous part, any realization of failure processes can be expressed as a sequence of R-values denoted $\{R[k]\}_{0<k\leq 1}$ where k is the fraction of removed links and $k \in \{1/L, 2/L, \ldots, 1\}$. Assuming that the nature of the

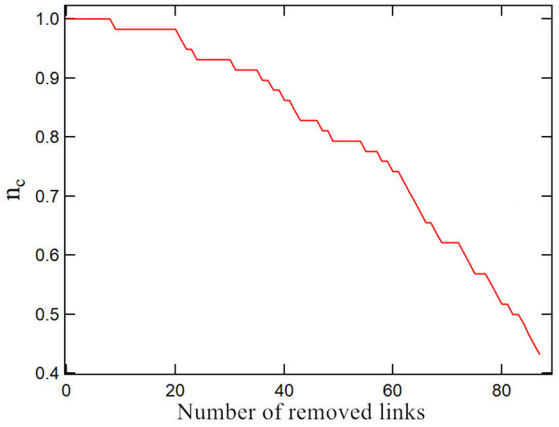


Fig. 1. The impact of link removal on the normalized maximum number n_c of controllable nodes in a communication network DFN (German optical backbone X-WiN network) with $N = 58$ and $L = 87$.

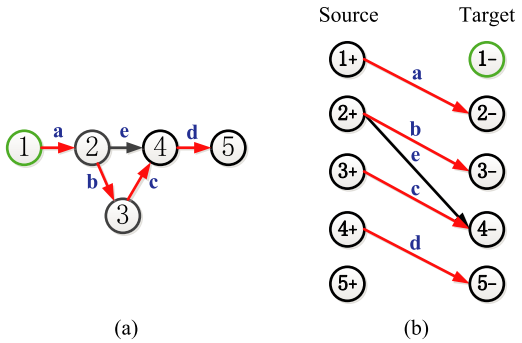


Fig. 2. Driver nodes and matching links in a directed network G. (a) An example network G with $N = 5$ nodes and $L = 5$ directed links. Matching links are denoted in red. Unmatched nodes are denoted in green. (b) The corresponding bipartite graph with 10 nodes and 5 links. By using the Hopcroft-Karp algorithm, a maximum set of matching links $\{a, b, c, d\}$ can be found in the bipartite graph. The target nodes $\{2-, 3-, 4-, 5-\}$ of the matching links are matched nodes. The other target node $\{1-\}$ is an unmatched node, therefore it is also driver node.

failures is unknown and they occur independently, $R[k]$ is a random variable and can be described by its probability density function (pdf). The pdf of this $R[k]$ is computed using all subsets of $\lfloor kL \rfloor$ links in all possible perturbations. The envelope for a network G is constructed using all $R[k]$ for $k \in \{1/L, 2/L, \ldots, 1\}$, where boundaries are given by the extreme R-values

$$R_{\min}[k] \in \{\min(R[1/L]), \min(R[2/L]) \ldots, \min(R[1])\}, \tag{3}$$

$$R_{\max}[k] \in \{\max(R[1/L]), \max(R[2/L]) \ldots, \max(R[1])\}, \tag{4}$$

which gives the worst- and best-case of robustness metrics for a network after a given number of challenges [17]. Besides, the expected R-value resulting from $\lfloor kL \rfloor$ perturbations

$$R_{avg}[k] \in \{E(R[1/L]), E(R[2/L]) \ldots, E(R[1])\}. \tag{5}$$

Since $R[k]$ defines a probability density function, we are interested in the percentiles of $R[k]$

$$R_{\theta\%}[k] \in \{R_{\theta\%}[1/L], R_{\theta\%}[2/L] \ldots, R_{\theta\%}[1]\} \tag{6}$$

where $R_{\theta\%}[k]$ are the points at which the cumulative distribution of $R[k]$ crosses $\frac{\theta}{100}$, namely if $R_{\theta\%}[k] = t$, then $\Pr[R[k] \leq t] = \frac{\theta}{100}$. We refer to $R_{\theta\%}[k]$ as a θ-percentile and define $R_{0\%}[k] = R_{min}[k]$, $R_{100\%}[k] = R_{max}[k]$.

We apply the envelope to present the influence of the failure process on a network [16], [17]. The envelope profiles the pdf of the random variables of the R-value, which is the probability of a random variable to fall within a particular region. The area of the envelope can be regarded as the variation of the robustness impact of a certain series of challenges, which quantifies the uncertainty or the amount of risk due to perturbations. The effectiveness of attack strategies can also be measured by comparing with the worst-, best- and average performance provided by robustness envelopes.

III. Analysis of Critical Links

Liu *et al.* [1] proved that the minimum number N_d of driver nodes needed for structural controllability, where the external signals are injected to control the directed network, can be obtained through the "maximum matching" of the network. Define the source node of a directed link as the node from which the link originates and the target node as the node where the link terminates. A maximum matching of a directed network is a maximum set of links that do not share source or target nodes [21], which is illustrated in Figure 2(a). Such links are coined "matching links". Target nodes of matching links are matched nodes and the other nodes are unmatched nodes. In order to find the maximum number of matching links, so as to determine the minimum number N_d of driver nodes, a directed network G with N nodes and L links can be converted into a bipartite graph $B_{N,N}$ with $2N$ nodes and L links, as shown in Figure 2(b). A maximum matching in a bipartite graph can be obtained efficiently by the Hopcroft-Karp algorithm [22]. The unmatched nodes in a maximum matching constitute a minimum set of driver nodes.

A. The Role of Critical Links in Maximum Matching

Links in a network can be classified into three categories: critical, redundant, and ordinary [1]. A link is critical if its removal increases the minimum number of driver nodes N_d by 1 to remain in full control of the system. A link is redundant if it never belongs to a maximum matching. A link is ordinary if it is neither critical nor redundant. In Figure 2(a), link a, b, c and d (highlighted in red) are critical links, the removal of any one of them will increase the number of driver nodes by 1, while link e is redundant. The influence of the removal of critical links can be explained by the maximum matching. As shown in 2(b), all the critical links a, b, c, d belong to the maximum matching of size 4. If any one of them is removed, there is no alternative link to take its place in the maximum matching. Thus, a new unmatched node will appear and the number of driver nodes will increase by 1. Besides, critical links are conditional and should be updated during attacks. For example, link c is no longer a critical link in the resulting network after link b is removed.

In our previous work [11], we proposed closed-form analytic approximations for the minimum number N_d of driver nodes needed to fully control networks, where links are removed according to both random and targeted attacks.

B. The Role of Critical Links in the Structure of Control

Parekh *et al.* [12] found the control structure which consists of a backbone of directed paths, called stems, each driven by an independent control. These paths can then control cycles that are inherently self-regulatory. However, ultimately these stems dictate the need for controls: There must be one control node for each stem in the system in order to guarantee that all nodes in the network are controllable (reachable). In this paper, we use the algorithm proposed in [6] to find the control structure in the network.

1) Determine the number *M* of control nodes by the maximum matching introduced in Section III-A.
2) Preprocess the network by adding the fixed number of control nodes and then placing links from each control node to every state node, after which there are *N* nodes and *E* links in the network. Then, for all $i, j = 1, \ldots, N$ and $k = 1, \ldots, M$:
   a) Split the nodes into a pair of positive and negative nodes $x_i \Rightarrow x_i^+, x_i^-$, $u_k \Rightarrow u_k^+, u_k^-$.
   b) Add unit-weight links (x_i^+, x_j^-) and (u_k^+, x_j^-) if the link (x_i, x_j) and (u_k, x_j) exist in the network, respectively.
   c) Add zero-weight links (x_i^+, x_i^-) and (u_k^+, u_k^-).
   d) Add zero-weight links (x_i^+, u_k^-).
   e) Add a weight $W \geq E$ to all links.
3) Use the weighted maximum matching on the bipartite graph generated by step 2 and find a set of matched links. The control structure is then formed by mapping the matched links in the original network.

Finally, the number of controllable nodes equals the number of matched nodes in the control structure. Although the concept of critical links was first proposed in control-based controllability analysis which focuses on the number N_d of driver nodes, critical links also play an important role in reachability-based controllability. We found that the removal of a critical link usually decreases the number of controllable nodes by 1 in most cases when the network is sparse.In this paper, we use the concept of critical links to derive analytical approximations for the decrease in the number N_c of controllable nodes upon link removal.

IV. Number of Controllable Nodes Under Attacks

In this section, we analyze the normalized number of controllable nodes for three different attack scenarios: (a) random attack, (b) targeted attack and (c) random attack under protection. In a random attack, links are removed from the network uniformly at random. In a targeted attack, we assume that the attacker knows the location of critical links and removes critical links uniformly at random. After all critical links are removed, the attacker randomly removes other links. In a random attack under protection, the network operator takes measures to protect the critical links such that only non-critical links are removed randomly.

TABLE I
Properties of the 10 Considered Communication Networks

Networks	N	L	$E[D]$	N_{d0}	L_c
DFN	58	87	3.0	25	14
Colt	153	177	2.3	81	38
Deltacom	113	161	2.8	37	43
GtsCe	149	193	2.6	58	49
TataNld	145	186	2.6	52	48
UsCarrier	158	189	2.4	53	66
Cogentco	197	243	2.5	71	72
Uninett2010	74	101	2.7	26	27
Kdl	754	895	2.4	272	287
Web [24]	643	2280	7.0	324	108

We compare the normalized number n_c of controllable nodes for these three attacks in 10 sparse communication networks [23], [24]. Table I presents the properties of the 10 communication networks: the number *N* of nodes, the number *L* of directed links, the initial minimum number N_{d0} of driver nodes and the number L_c of critical links. For a directed network, the average degree $E[D] = 2L/N$, which also equals the sum of the mean out- and in-degree per node. The first 8 networks are small. The other two networks are relatively large, which are an order larger than the average size of the other 8 communication networks. Besides, the last network has a higher average degree, which is more than twice that of the other networks. The number L_c of critical links can be determined by applying the Hopcraft-Karp algorithm *L* times, by considering all *L* networks that are obtained by removing exactly one link from the original network. As expected, Figure 3 shows that random attack under protection performs the best among the three attack scenarios in maintaining the reachability of the networks. Moreover, we also conclude from Figure 3.

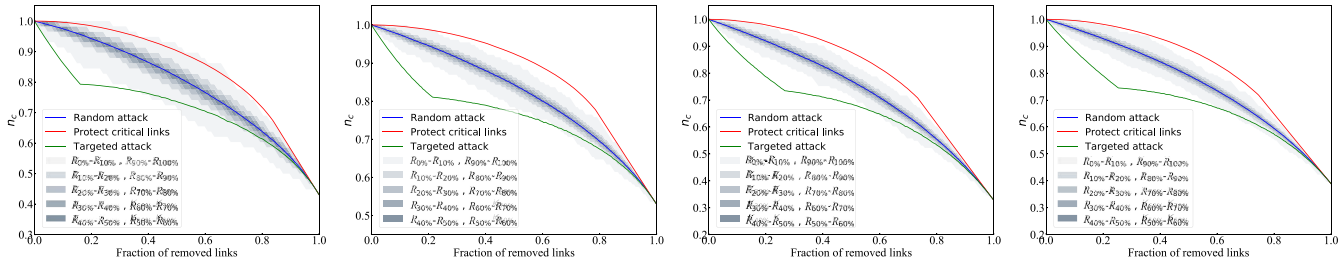
1) In the case of random attack under protection, the slope of the decrease in n_c is almost 0 in the beginning for all networks. This emphasizes the importance of protecting critical links.

2) The targeted attack is the most harmful: when the fraction of removed links is smaller than the fraction of critical links, the decrease in n_c is almost linear in the fraction of removed links. When all the critical links are removed, the slope of the decrease in n_c is almost 0 in all 8 networks. Considering the set of critical links is determined from the initial network, this indicates that the set of critical links of a network does not significantly change during the attack process when the fraction of removed links is small.

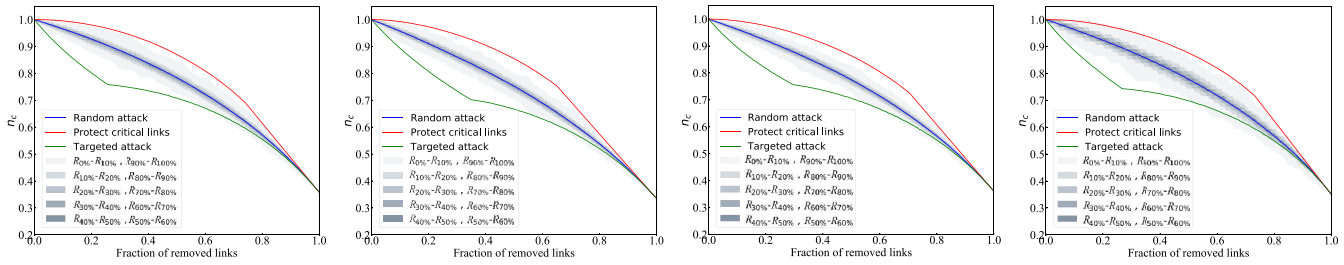
3) The performance of random attack is between targeted attack and random attack under protection. After all links are removed, the normalized number of controllable nodes equals N_{d0}/N.

4) Critical links have a significant impact on the number N_c of controllable nodes upon link removal, which plays a key role to derive analytical approximations for the number N_c of controllable nodes.

We also use robustness envelopes to evaluate the effectiveness of the three attack strategies. As shown in Figure 3, the curves for random attack under protection are quite close to the



(a) DFN (b) Colt (c) Deltacom (d) GtsCe



(e) TataNld (f) UsCarrier (g) Cogentco (h) Uninett2010

Fig. 3. Performance of the normalized number n_c of controllable nodes as a function of the fraction of removed links l for three attack scenarios. The results for each fraction l is based on 1000 simulations. Each envelope of the challenges for the normalized number n_c of controllable nodes is based on 10^4 realizations. In order to compare the scenario for random attack under protection with the other two scenarios in the same sub-figure, we remove critical links uniformly at random after all the other links are removed.

boundaries represented by the 90-percentile $R_{90\%}[k]$ among all networks, which means that random attack under protection outperforms 90% realizations of random attack. The curves for targeted attack are much lower than the lower bound of envelopes especially when the fraction of removed links is small in all networks, which again underlines the harm of targeted attack.

V. Approximations for the Number of Controllable Nodes

In the previous section, we compared the number of controllable nodes for the three attack scenarios by using a large amount of simulations. In this section, we deduce analytical approximations to quantify the robustness of reachability, expressed in terms of the normalized number n_c of controllable nodes, for the three attack scenarios. Then, we evaluate the accuracy of the analytical approximations in 8 small networks, 2 large networks as well as more sparse communication networks. Lastly, we also use synthetic networks to measure the performance of our analytical approximations. Our approximations will be based upon the concept of critical links introduced in [1].

A. Number of Controllable Nodes Under Random Attacks

1) *The Fraction l of Removed Links is Less Than the Fraction l_c of Critical Links:* Given a network with N nodes and L links, the initial number N_c of controllable nodes equals N. The number L_c of critical links can be determined by the method we introduced in Section IV.

As discussed in Section III-B, the number N_c of controllable nodes decreases by at least one when a critical link is removed. However, we found that the number N_c of controllable nodes only decreases by one for every critical link that is removed in each of the 10 sparse communication networks in Table I. Thus, we heuristically assume that after removing a critical link, the number N_c of controllable nodes decreases by one. If we denote the number of removed links by m, then the fraction of removed links $l = \frac{m}{L}$, while the fraction of critical links l_c satisfies $l_c = \frac{L_c}{L}$. We consider the case $l \leq l_c$, where m links are removed uniformly at random under the condition that the number of removed links obeys $m \leq L_c$. Now assume that of these m links i links are critical ($i \leq m$) and, hence, $m - i$ links are non-critical. We assume that the set of critical links is nearly unchanged when the fraction of removed links is small. Invoking the fact that after removing a critical link, the number N_c of controllable nodes decreases by one, thus, when i critical links are iteratively removed one by one, the number N_c of controllable nodes decreases by one in each iteration. For the $m - i$ removed non-critical links, the number N_c of controllable nodes remains the same based on our assumption that the set of critical links is unchanged when the fraction of removed links is small. Since there are $\binom{L_c}{i}$ possible ways to choose i critical links from L_c critical links and there are $\binom{L-L_c}{m-i}$ possible ways to choose $m-i$ non-critical links from $L - L_c$ non-critical links, the contribution to the decrease in N_c is $i\binom{L_c}{i}\binom{L-L_c}{m-i}$. The average decrease N_c^* of the number N_c of controllable nodes after randomly removing m links, is the sum of this expression for all $i = 1, 2, \ldots, m$ divided by $\binom{L}{M}$.

$$N_c^* = \frac{\sum_{i=1}^{m} i\binom{L_c}{i}\binom{L-L_c}{m-i}}{\binom{L}{m}} \tag{7}$$

Using $i\binom{L_c}{i} = L_c\binom{L_c-1}{i-1}$ and Vandermonde's formula $\sum_{j=0}^{k}\binom{a}{j}\binom{b}{k-j} = \binom{a+b}{k}$ for any number a and b, we

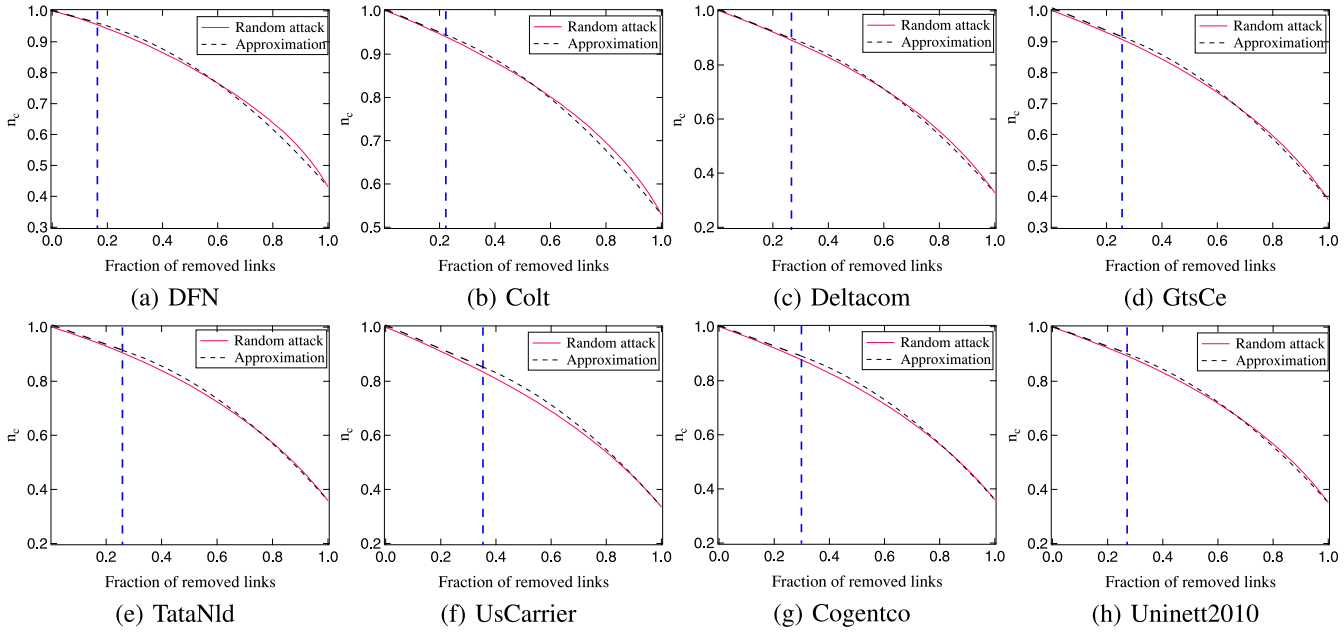


Fig. 4. The normalized maximum number of controllable nodes n_c as a function of the fraction of removed links l in communication networks under random attacks. The results for each fraction l is based on 1000 simulations. In each sub-figure, the vertical dashed line marks the position where $l = l_c$.

obtain $L_c \sum_{i=0}^{m-1} \binom{L_c-1}{i}\binom{L-L_c}{m-1-i} = L_c\binom{L-1}{m-1}$. Finally, dividing this expression by $\binom{L}{m}$, leads to the average decrease of controllable nodes

$$N_c^* = lL_c \tag{8}$$

When the fraction of removed links is less than, or equal to l_c, we obtain

$$N_c = N - lL_c \tag{9}$$

We then normalize the number N_c of controllable nodes to the fraction $\frac{N_c}{N}$ of the minimum number of controllable nodes and denote the obtained approximation as $n_{c,rand}$,

$$n_{c,rand} = \frac{N - lL_c}{N} \tag{10}$$

2) *The Fraction l of Removed Links is Larger Than the Fraction l_c of Critical Links:* Considering that in most cases l_c is quite small, we also estimate the normalized maximum number n_c of controllable nodes when the fraction l of removed links is larger than the fraction l_c of critical links. For $l \geq l_c$, we heuristically propose a simple closed-form approximation for $n_{c,rand}$:

$$n_{c,rand} = al^2 + bl + c \tag{11}$$

where the parameters a, b and c will be determined by boundary conditions. For the first two boundary conditions we assume that, for $l = l_c$, Eq. (11) has the same value and the same derivative as Eq. (10). This leads to the equations $N - l_c L_c = N(al_c^2 + bl_c + c)$ and $-L_c = N(2al_c + b)$, respectively. Finally, if we remove all links, i.e., $l = 1$, only N_{d0} nodes can be controlled. This gives the boundary condition $N_{d0}/N = a + b + c$. Solving for a, b and c and combining with the approximation Eq. (10), we obtain the following approximation for $n_{c,rand}$ for all values of l:

$$n_{c,rand} = \begin{cases} \frac{N - lL_c}{N} & l \leq l_c \\ al^2 + bl + c & l \geq l_c \end{cases} \tag{12}$$

with, $a = -\frac{N - N_{d0} - L_c}{N(l_c - 1)^2}$, $b = -L_c/N - 2al_c$, and $c = (N_{d0} + L_c)/N + a(2l_c - 1)$. Eq. (12) respresents a closed-form approximation for n_c, which only depends on N, L, N_{d0} and L_C. The computational complexity of the approximation is $O(\sqrt{N}L^2)$, which is needed for the computation of L_c.

We compare the approximation Eq. (12) with simulation results for the 8 communication networks. Figure 4 illustrates that the approximation both under- and overestimates the value of n_c. For moderate values of the fraction of removed links, the approximation exhibits a very good fit for the communication networks. For some networks, such as Deltacom, GtsCe, TataNld and Uninett2010, our approximation Eq. (12) fits well with the simulation results regardless of the fraction of removed links.

The performance of our approximations are also measured by three performance indicators:

1) r^* denotes the absolute value of the relative error at $l = 0.2$. We choose the value 0.2 reflecting a relatively large fraction in terms of link-based failures or attacks.

2) l^* represents the smallest value of l, where the relative error between the approximation and the simulated mean exceeds 5%.

3) γ denotes the fraction of the interval $[0, l_c]$ for which the absolute value of the relative error between the approximation and the mean simulated value does not exceed 5%. The value of γ is computed by K different values of the fraction of removed links, i.e., $v_1, v_2, \ldots, v_K$, are evenly determined in the interval $[0, l_c]$. Let $n_c^*(v_i)$ and $n_c(v_i)$ denote the mean simulated n_c and the approximation (10) at the fraction of removed links $l = v_i$, respectively. Thus, in terms of the

TABLE II
PERFORMANCE INDICATORS FOR THE APPROXIMATION $n_{c,rand}$ FOR THE 8 COMMUNICATION NETWORKS

Networks	r^*	l^*	γ
DFN	0.78%	0.87	100%
Colt	1.23%	0.81	100%
Deltacom	1.08%	1.00	100%
GtsCe	1.20%	1.00	100%
TataNld	0.45%	1.00	100%
UsCarrier	1.17%	1.00	100%
Cogentco	0.98%	1.00	100%
Uninett2010	1.24%	1.00	100%

TABLE III
PERFORMANCE INDICATORS FOR THE APPROXIMATION $n_{c,crit}$ FOR THE 8 COMMUNICATION NETWORKS

Networks	r^*	l^*	γ
DFN	4.53%	0.23	100%
Colt	4.88%	0.21	97.81%
Deltacom	7.52%	0.18	67.42%
GtsCe	5.06%	0.19	74.80%
TataNld	10.11%	0.12	46.51%
UsCarrier	3.26%	0.22	63.04%
Cogentco	8.75%	0.18	60.81%
Uninett2010	8.86%	0.19	71.08%

indicator function 1_x that equals 1 if the condition x is true, otherwise it is zero,

$$\gamma = \frac{\sum_{i=1}^{K} \mathbf{1}_{\left|\frac{n_c^*(v_i)-n_c(v_i)}{n_c^*(v_i)}\right| \leq 5\%}}{K}.$$

Table II gives the three performance indicators for Eq. (12). As shown in the table, when the fraction of removed links is less than 0.2, the absolute relative error between Eq. (12) and the simulated mean is less than 5% for all 8 networks. For most networks, such as Deltacom, GtsCe, TataNld, UsCarrier, Cogentco and Uninett2010, Eq. (12) still fits well with simulation results regardless of the fraction of removed links. For the worst performing networks, DFN and Colt, 87% and 81% of the links can be removed before the absolute relative error exceeds 5%, respectively. When the fraction of removed links is less than the fraction l_c of critical links, the absolute value of the relative error between the approximation and the mean simulated value is always less than 5%. Thus, γ equals 100% for all networks.

B. Number of Driver Nodes Under Targeted Attacks

1) *The Fraction l of Removed Links is Smaller Than the Fraction l_c of Critical Links:* We assume that, as long as the number of removed links $m \leq L_c$, the removal of each link decreases the number N_c of controllable nodes by one. Consequently, when the number of removed links is smaller than L_c (the fraction l of removed links is smaller than l_c), the approximation for the minimum number N_c of driver nodes decreases linearly with the fraction l of removed links. When the number of removed links equals the number L_c of critical links, the minimum number N_c of driver nodes equals $N - L_c$. Thus, when the fraction l of removed links is no more than the fraction l_c of critical links, we obtain the following approximation for n_c:

$$n_{c,crit} = \frac{N - lL}{N}. \tag{13}$$

2) *The Fraction l of Removed Links is Larger Than the Fraction l_c of Critical Links:* We now construct an approximation when the number of removed links is larger than L_c (the fraction l of removed links is larger than l_c), in a similar way as in the previous section. Again assuming that for $l \geq l_c$ it holds that n_c is quadratic in l, we obtain $n_{c,crit} = dl^2 + el + f$. Boundary conditions are now obtained from the assumptions that the parabola passes through $(1, N_{d0}/N)$ and $(l_c, (N - L_c)/N)$ and has a zero derivative at the latter point. This leads to the following approximation for n_c for all values of l:

$$n_{c,crit} = \begin{cases} \frac{N-lL}{N} & l \leq l_c \\ dl^2 + el + f & l \geq l_c \end{cases} \tag{14}$$

with, $d = -\frac{N - N_{d0} - l_c L}{N(l_c - 1)^2}$, $e = -2dl_c$, and $f = N_{d0}/N + d(2l_c - 1)$.

In Figure 5, we compare our approximation Eq. (14) with simulation results. Simulation results show that the difference in the curve trend at $l = l_c$, is due to the fact that until $l = l_c$ only critical links are targeted causing a faster descent in the number of controllable nodes. We observe that the approximation Eq. (14) fits well with simulation results when the fraction of removed links is sufficiently small in these communication networks. In some networks, such as DFN and UsCarrier, Eq. (14) is close to simulation results even when the fraction of removed links is relatively large. When the fraction of removed links is getting larger, the difference between our approximation Eq. (14) and simulation results is relatively large. However, approximation Eq. (14) always seems to overestimate the impact of targeted attack on the normalized maximum number n_c of controllable nodes, hence, approximation Eq. (14) can be considered a worst-case approximation.

Comparing with the targeted attack, we quantify the performance of the approximation Eq. (14) in Table III. For DFN and Colt, Eq. (14) is a very good approximation when the fraction of removed links is less than l_c. Eq. (14) performs the best for DFN, 23% of the links can be removed before the absolute relative error exceeds 5%. Eq. (14) does not perform well for TataNld.

C. Number of Driver Nodes Under Random Attacks With Protection

For this scenario, we assume that a fraction of links l_c is protected, then we can only attack a fraction $1 - l_c$ of the links. We now construct an approximation for the number $N_{c,prot}$ of controllable nodes when the attack is random under protection. We heuristically assume that the fraction $n_{c,prot}$ of controllable nodes is quadratic in l, we obtain $n_{c,prot} = pl^2 + ql + r$. Boundary conditions are now obtained from the assumptions that the parabola passes through $(1, N_{d0}/N)$ and (0, 1) and has a zero derivative at the latter point. This leads to the following

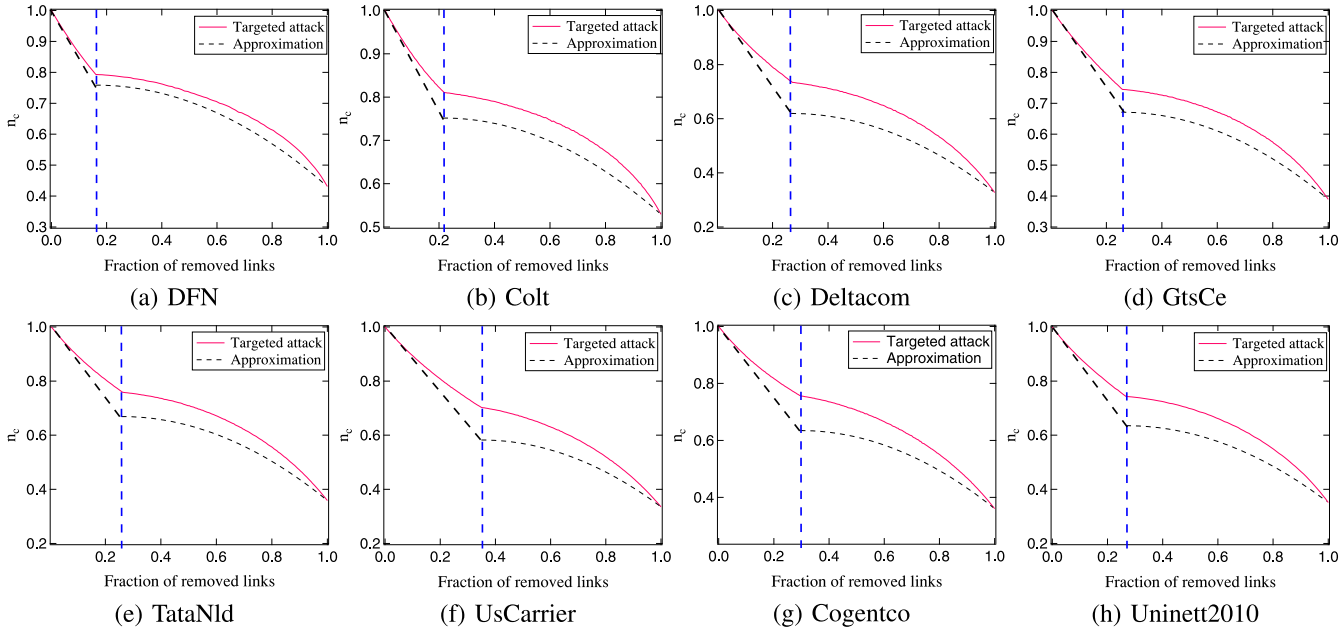


Fig. 5. The normalized maximum number of controllable nodes n_c as a function of the fraction of removed links l in communication networks under targeted attacks. The results for each fraction l is based on 1000 simulations. In each sub-figure, the vertical dashed line marks the position where $l = l_c$.

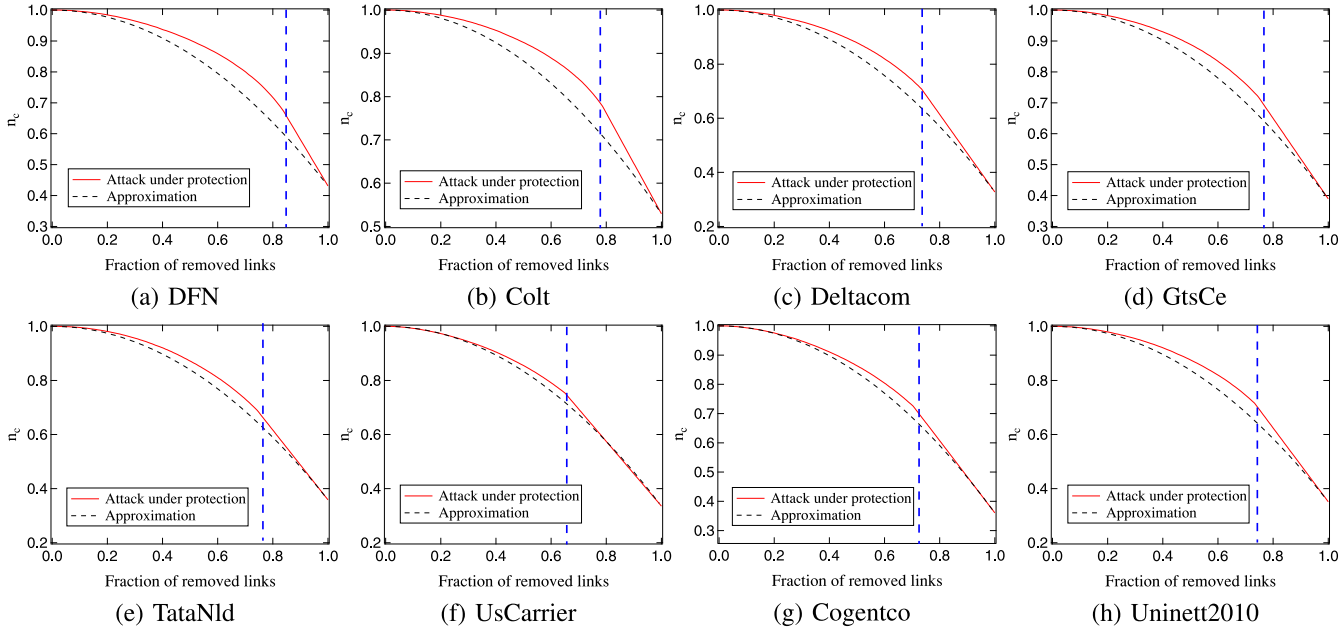


Fig. 6. The normalized maximum number of controllable nodes n_c as a function of the fraction of removed links l in communication networks under random attacks under protection. The results for each fraction l is based on 1000 simulations. In each sub-figure, the vertical dashed line marks the position where $l = 1 - l_c$. In order to compare the simulation results for random attack under protection with our approximation in the same sub-figure, we remove critical links uniformly at random after all the other links are removed.

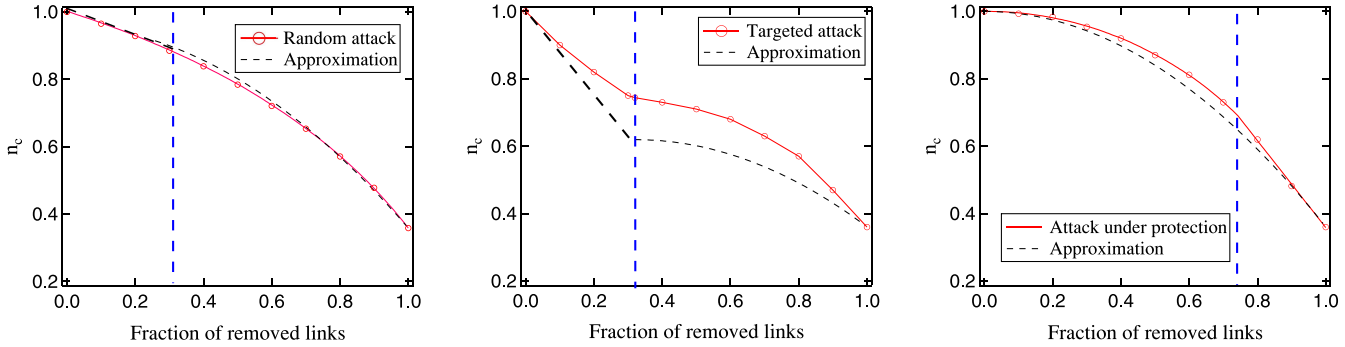
approximation for n_c for all values of l:

$$n_{c,prot} = pl^2 + ql + r \tag{15}$$

with, $p = N_{d0}/N - 1$, $q = 0$, and $r = 1$.

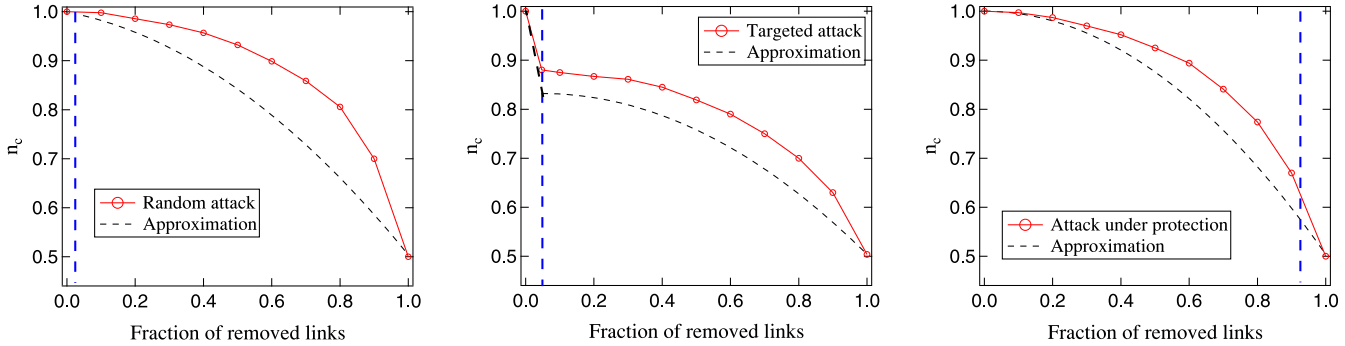
We compare the approximation Eq. (15) with simulation results for the 8 communication networks. The fraction l of removed links in our approximation Eq. (15) is from 0 to 1. However, only a fraction $1 - l_c$ of links are removed in the simulation for this scenario. Thus, we still remove critical links uniformly at random after all non-critical links are removed, in order to compare the simulation results and our approximation Eq. (15) in the same interval [0, 1]. Figure 6 shows that for moderate values of the fraction of removed links, the approximation exhibits an excellent fit for simulation results. For some networks, such as TataNld, UsCarrier and Cogentco, our approximation Eq. (15) fits well with the simulation results regardless of the fraction of removed links.

Similarly, the performance of our approximation Eq. (15) is measured by three performance indicators. As shown in Table IV, when the fraction of removed links is less than 0.2,



(a) Random attack (b) Targeted attack (c) Random attack with protection

Fig. 7. Performance of the normalized number n_c of controllable nodes as a function of the fraction of removed links l for three attack scenarios in the Kdl network. The results for each fraction l is based on 1000 simulations. For random attack and targeted attack, the vertical dashed line marks the position where $l = l_c$. For random attack with protection, the vertical dashed line marks the position where $l = 1 - l_c$.



(a) Random attack (b) Targeted attack (c) Random attack with protection

Fig. 8. Performance of the normalized number n_c of controllable nodes as a function of the fraction of removed links l for three attack scenarios in the Web network. The results for each fraction l is based on 1000 simulations. For random attack and targeted attack, the vertical dashed line marks the position where $l = l_c$. For random attack with protection, the vertical dashed line marks the position where $l = 1 - l_c$.

TABLE IV
PERFORMANCE INDICATORS FOR THE APPROXIMATION $n_{c,prot}$ FOR THE 8 COMMUNICATION NETWORKS

Networks	r^*	l^*	γ
DFN	2.65%	0.28	100%
Colt	1.89%	0.32	100%
Deltacom	1.15%	0.37	100%
GtsCe	1.09%	0.38	100%
TataNld	0.84%	0.42	100%
UsCarrier	0.32%	0.62	100%
Cogentco	0.56%	0.58	100%
Uninett2010	1.22%	0.47	100%

the absolute relative error between Eq. (15) and the simulated mean is less than 5% for all 8 sparse communication networks. For some networks, such as UsCarrier and Cogentco, even when the fraction of removed links is large (0.62 and 0.58, respectively), Eq. (15) still fits well with simulation results. Even for the worst performing network, DFN, 28% of the links can be removed before the absolute relative error exceeds 5%.

D. Verification by Large Networks

We use the last two large networks, Kdl and Web, from Table I to further evaluate the accuracy of our approximations. The simulations setting is slightly different from the previous part. The fraction l of removed links is ranging from 0.1 to 1 with a step 0.1, considering the high computational complexity by using all fraction l ranging from 0 to 1. As shown in Figure 7, the approximation Eq. (12) for the random attack and Eq. (15) for the random attack with protection perform well in estimating the fraction n_c of driver nodes. We also find that the approximation Eq. (14) for targeted attack fits well with simulation results when the fraction of removed links is sufficiently small. Though the approximation Eq. (14) does not perform well when the fraction of removed links is large, approximation Eq. (14) can be considered a worst-case approximation. Considering the Kdl network and the 8 small networks have similar average degree, the above observation implies that the size of the network does not significantly influence the performance of our approximations. By contrast, Figure 8(a) and (c) show that the approximation Eq. (12) for the random attack and Eq. (15) for the random attack with protection do not perform well for the Web network which has a larger average degree than the above networks. In a network with a higher average degree, there are more alternate matchings which make it more likely for the critical links to change as links are removed. As a result, our approximations do not perform well since our assumption is that the set of critical links is nearly unchanged when the fraction of removed links is small. However, since most communication networks are sparse [25], [26], [27], we can expect that our approximations are applicable for most communication networks.

TABLE V
PERFORMANCE INDICATORS FOR THE APPROXIMATION n_c FOR KDL

Types of attacks	r^*	γ
Random attack	2.36%	100%
Targeted attack	7.54%	53.65%
Random attack with protection	3.67%	92.84%

TABLE VI
PERFORMANCE INDICATORS FOR THE APPROXIMATION n_c FOR WEB

Types of attacks	r^*	γ
Random attack	13,78%	100%
Targeted attack	9.21%	54.16%
Random attack with protection	5.25%	100%

TABLE VII
PERFORMANCE INDICATORS FOR THE APPROXIMATION n_c FOR 200 COMMUNICATION NETWORKS

Types of attacks	r^*	l^*	γ
Random attack	4.26%	0.47	98.47%
Targeted attack	10.44%	0.11	47.81%
Random attack with protection	3.52%	0.23	99.36%

We then quantify the performance of the approximations for the network Kdl and Web in Tables V and VI, respectively. Results show that the network Web has larger r^* values than the network Kdl in all three attacks, which also indicates that our approximation perform better in the networks with lower average degree.

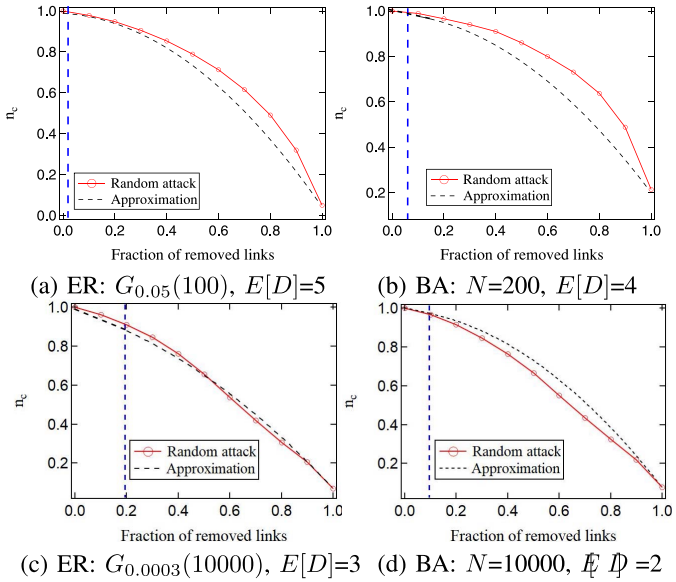
E. Verification by More Communication Networks

We further use the dataset available at a specialized database - the Internet Topology Zoo [23] to select more communication networks and verify the accuracy of our approximations. The networks in the dataset initially are not directed, however, we use the information available in two attributes, i.e., source node and target node, to make these networks directed. After excluding networks with extremely small size, we have 200 communication networks.

For each attack strategy, we calculate the values of the three performance indicators for all 200 communication networks and then get the average value for each indicator. As shown in Table VII, the approximation Eq. (12) for the random attack and Eq. (15) for the random attack with protection perform well in estimating the fraction n_c of driver nodes. For the targeted attack, the approximation Eq. (14) fits well with simulation results when the fraction of removed links is sufficiently small.

F. Verification by Synthetic Networks

In this section, we test our approximations on two types of synthetic networks, the directed Erdős-Rényi (ER) random network $G_p(N)$ and the Barabási-Albert (BA) scale-free network $BA(N, M_0, M)$. When generating the directed Erdős-Rényi random network $G_p(N)$ with N nodes, the probability that every node has an outbound link to the other nodes is p. We generate the scale-free network $BA(N, M_0, M)$ by using the Barabási-Albert (BA) model, where N is the number of nodes, M is the number of out-going links for each new node added to the current network. We assume that initially the network consists of a complete digraph on M_0 nodes, where M_0 equals M. In the initial complete digraph, every pair of distinct nodes is connected by a pair of unique links (one in each direction). New nodes are added to the network one at a time. Each new node is connected to M existing nodes with a probability that is proportional to the number of links that the existing nodes already have.



(a) ER: $G_{0.05}(100)$, $E[D]$=5 (b) BA: N=200, $E[D]$=4

(c) ER: $G_{0.0003}(10000)$, $E[D]$=3 (d) BA: N=10000, $E[D]$=2

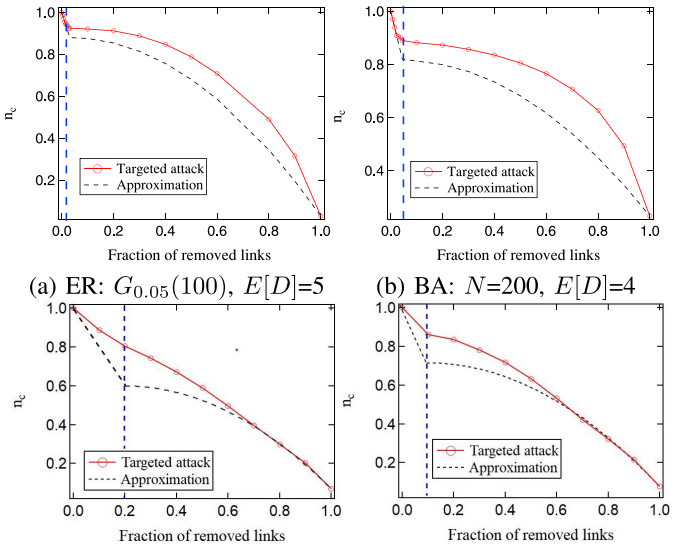
Fig. 9. The normalized maximum number of controllable nodes n_c as a function of the fraction of removed links l in synthetic networks under random attacks. The results for each fraction l is based on 10000 simulations.

In our simulations, we generate Erdős-Rényi (ER) random networks $G_p(N)$ with $N = 100$, $p = 0.05$ and $N = 10000$, $p = 0.0003$, Barabási-Albert (BA) networks with $N = 200$, $M = M_0 = 2$ and $N = 10000$, $M = M_0 = 1$. Figure 9 shows that the approximation Eq. (12) for the random attack performs well in estimating the fraction n_c of controllable nodes in both types of synthetic networks when the fraction of removed links is small. Figure 10 shows that the approximation Eq. (14) for the targeted attack performs well as long as the fraction of removed links is sufficiently small. Figure 11 shows that Eq. (15) for the random attack with protection perform well in both types of synthetic networks when the fraction l of removed links is less than the fraction l_c of critical links. For the large ER and BA networks, Eq. (15) fits well with simulation results even when the fraction l of removed links is large. The approximation Eq. (14) does not perform well if the fraction l of removed links is large. However, Eq. (14) can be considered an approximation for the worst-case scenario.

Next we quantify the performance of each approximation for synthetic networks. As shown in Tables VIII, IX, and X, the approximation Eq. (12) for random attack and Eq. (15) for random attack with protection fit well with simulation results even when the fraction l of removed links is relatively large ($l = 0.2$).

VI. CONCLUSION

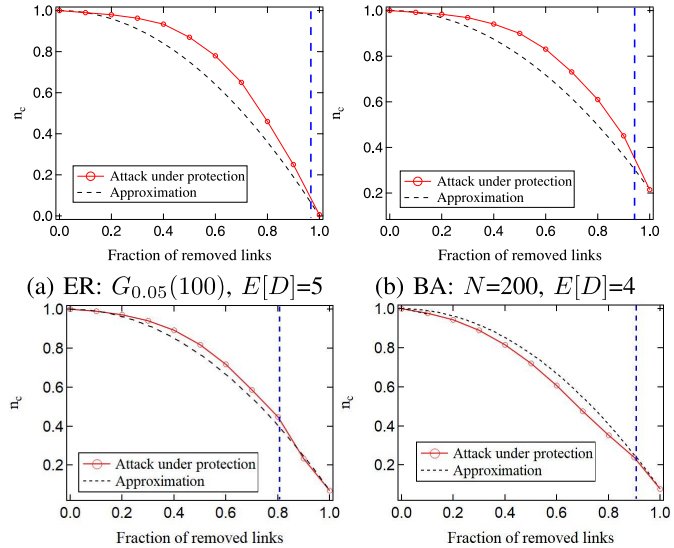
In this study, we analyzed the role of critical links in network controllablity. Simulation results on communication



(a) ER: $G_{0.05}(100)$, $E[D]$=5 (b) BA: N=200, $E[D]$=4

(c) ER: $G_{0.0003}(10000)$, $E[D]$=3 (d) BA: N=10000, $E[D]$=2

Fig. 10. The normalized maximum number of controllable nodes n_c as a function of the fraction of removed links l in synthetic networks under targeted attacks. The results for each fraction l is based on 10000 simulations.



(a) ER: $G_{0.05}(100)$, $E[D]$=5 (b) BA: N=200, $E[D]$=4

(c) ER: $G_{0.0003}(10000)$, $E[D]$=3 (d) BA: N=10000, $E[D]$=2

Fig. 11. The normalized maximum number of controllable nodes n_c as a function of the fraction of removed links l in synthetic networks under random attacks with protection. The results for each fraction l is based on 10000 simulations. In each sub-figure, the vertical dashed line marks the position where $l = 1 - l_c$.

TABLE VIII
PERFORMANCE INDICATORS FOR THE APPROXIMATION $n_{c,rand}$ FOR SYNTHETIC NETWORKS

Types of networks	r^*	γ
ER: $G_{0.05}(100)$	3.27%	100%
BA: N=200, $E[D]$=4	6.78%	100%
ER: $G_{0.0003}(10000)$	8.95%	85.54%
BA: N=10000, $E[D]$=2	7.63%	96.23%

networks have suggested analytical closed-form approximations for the number N_c of controllable nodes. We derived closed-form approximations for the number N_c of controllable nodes as a function of the fraction of removed links, for random attacks, targeted attacks and random attack under protection. Both for random and targeted attacks, our approximation is linear in the fraction l of removed links when this fraction is smaller than the fraction of critical links. When the fraction of removed links is larger than the fraction of critical links, our approximation is quadratic in l. We validated our approximation through simulations on sparse communication networks and synthetic networks. Both for random attacks and random attacks under protection, our approximations for these two cases are always very good, as long as the fraction of removed links is smaller than the fraction of critical links. For some cases, the approximation is still accurate for larger fractions of removed links. For targeted attack, our approximation performs well as long as the fraction of removed links is sufficiently small, whereas our approximation does not perform well when the fraction of removed links is large. However, the approximation for the targeted attack always serves as a worst-case estimate.

TABLE IX
PERFORMANCE INDICATORS FOR THE APPROXIMATION $n_{c,crit}$ FOR SYNTHETIC NETWORKS

Types of networks	r^*	γ
ER: $G_{0.05}(100)$	12.64%	71.97%
BA: N=200, $E[D]$=4	17.36%	63.91%
ER: $G_{0.0003}(10000)$	23.56%	34.28%
BA: N=10000, $E[D]$=2	16.28%	46.76%

TABLE X
PERFORMANCE INDICATORS FOR THE APPROXIMATION $n_{c,prot}$ FOR SYNTHETIC NETWORKS

Types of networks	r^*	γ
ER: $G_{0.05}(100)$	5.21%	100%
BA: N=200, $E[D]$=4	8.63%	100%
ER: $G_{0.0003}(10000)$	4.16%	100%
BA: N=10000, $E[D]$=2	6.94%	100%

REFERENCES

[1] Y.-Y. Liu, J.-J. Slotine, and A.-L. Barabási, "Controllability of complex networks," *Nature*, vol. 473, no. 7346, pp. 167–173, 2011.

[2] Y. Nakayama, K. Mori, K. Takaragi, and S. Domen, "System and method for performing interlocution at a plurality of terminals connected to communication network," U.S. Patent 5 280 583, Jan. 18, 1994.

[3] Y. Lou, L. Wang, and G. Chen, "Toward stronger robustness of network controllability: A snapback network model," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 65, no. 9, pp. 2983–2991, Sep. 2018.

[4] C.-L. Pu, W.-J. Pei, and A. Michaelson, "Robustness analysis of network controllability," *Physica A Stat. Mech. Appl.*, vol. 391, no. 18, pp. 4420–4425, 2012.

[5] S. Nie, X. Wang, H. Zhang, Q. Li, and B. Wang, "Robustness of controllability for networks based on edge-attack," *PLOS ONE*, vol. 9, no. 2, 2014, Art. no. e89066.

[6] J. Thomas, S. Ghosh, D. Parek, D. Ruths, and J. Ruths, "Robustness of network controllability to degree-based edge attacks," in *Proc. Int. Workshop Complex Netw. Appl.*, 2016, pp. 525–537.

[7] S.-M. Chen, Y.-F. Xu, and S. Nie, "Robustness of network controllability in cascading failure," *Physica A Stat. Mech. Appl.*, vol. 471, pp. 536–539, Apr. 2017.

[8] Y. Lou, L. Wang, and G. Chen, "A framework of hierarchical attacks to network controllability," *Commun. Nonlinear Sci. Numer. Simulat.*, vol. 98, p. 105780, Jul. 2021.

[9] X. Yan-Dong, L. Song-Yang, H. Lv-Lin, and B. Liang, "Optimization of robustness of network controllability against malicious attacks," *Chin. Phys. B*, vol. 23, no. 11, 2014, Art. no. 118902.

[10] Z. Zhang, Y. Yin, X. Zhang, and L. Liu, "Optimization of robustness of interdependent network controllability by redundant design," *PLOS ONE*, vol. 13, no. 2, 2018, Art. no. e0192874.
[11] P. Sun, R. E. Kooij, Z. He, and P. Van Mieghem, "Quantifying the robustness of network controllability," in *Proc. 4th Int. Conf. Syst. Rel. Safety (ICSRS)*, 2019, pp. 66–76.
[12] D. Parekh, D. Ruths, and J. Ruths, "Reachability-based robustness of network controllability under node and edge attacks," in *Proc. 10th Int. Conf. Signal Image Technol. Internet-Based Syst. (SITIS)*, 2014, pp. 424–431.
[13] A. Lombardi and M. Hörnquist, "Controllability analysis of networks," *Phys. Rev. E*, vol. 75, no. 5, 2007, Art. no. 056110.
[14] P. Van Mieghem, *Graph Spectra for Complex Networks*. Cambridge, U.K.: Cambridge Univ. Press, 2011.
[15] R. E. Kalman, "Mathematical description of linear dynamical systems," *J. Soc. Ind. Appl. Math. A Control*, vol. 1, no. 2, pp. 152–192, 1963.
[16] P. Van Mieghem *et al.*, "A framework for computing topological network robustness," Fac. Electr. Eng. Math. Comput. Sci., Delft Univ. Technol., Delft, The Netherland, Rep. 20101218, 2010.
[17] S. Trajanovski, J. Martín-Hernández, W. Winterbach, and P. Van Mieghem, "Robustness envelopes of networks," *J. Complex Netw.*, vol. 1, no. 1, pp. 44–62, 2013.
[18] H. Cetinay, K. Devriendt, and P. Van Mieghem, "Nodal vulnerability to targeted attacks in power grids," *Appl. Netw. Sci.*, vol. 3, no. 1, p. 34, 2018.
[19] X. Wang, E. Pournaras, R. E. Kooij, and P. Van Mieghem, "Improving robustness of complex networks via the effective graph resistance," *Eur. Phys. J. B*, vol. 87, no. 9, p. 221, 2014.
[20] Z. He, P. Sun, and P. Van Mieghem, "Topological approach to measure network recoverability," in *Proc. 11th Int. Workshop Resilient Netw. Design Model. (RNDM)*, 2019, pp. 1–7.
[21] Y. Yang and G. Xie, "Mining maximum matchings of controllability of directed networks based on in-degree priority," in *Proc. 35th Chin. Control Conf. (CCC)*, 2016, pp. 1263–1267.
[22] J. E. Hopcroft and R. M. Karp, "An nˆ5/2 algorithm for maximum matchings in bipartite graphs," *SIAM J. Comput.*, vol. 2, no. 4, pp. 225–231, 1973.
[23] S. Knight, H. X. Nguyen, N. Falkner, R. Bowden, and M. Roughan, "The Internet topology zoo," *IEEE J. Sel. Areas Commun.*, vol. 29, no. 9, pp. 1765–1775, Oct. 2011.
[24] R. Rossi and N. Ahmed. (2013). *Network Repository*. [Online]. Available: http://networkrepository.com
[25] C. I. Del Genio, T. Gross, and K. E. Bassler, "All scale-free networks are sparse," *Phys. Rev. Lett.*, vol. 107, no. 17, 2011, Art. no. 178701.
[26] M. S. Uddin, S. T. H. Murshed, and L. Hossain, "Towards a scale free network approach to study organizational communication network," in *Proc. PACIS*, 2010, p. 196.
[27] L. Li, D. Alderson, J. C. Doyle, and W. Willinger, "Towards a theory of scale-free graphs: Definition, properties, and implications," *Internet Math.*, vol. 2, no. 4, pp. 431–523, 2005.

Peng Sun received the bachelor's degree in communication engineering and the master's degree in communication and information system from Shandong University in 2012 and 2017, respectively. He is currently pursuing the Ph.D. degree with the Delft University of Technology.

Robert E. Kooij received the M.Sc. and Ph.D. (*cum laude*) degrees in mathematics from the Delft University of Technology in 1988 and 1993, respectively. Since 1997, he has been employed with TNO ICT (the former KPN Research), the largest ICT Research Institute in the Netherlands. He is active both in national research projects, e.g., EQUANET, as well as in international projects, such as MUSE. He also regularly gives guestlectures on telecommunications with universities throughout the Netherlands. Since November 2005, he has been a part-time Associate Professor with Network Architectures and Services Group, where he works on the NWO Project ROBUNET which deals with the robustness of large scale networks. He was working as a Principal Research Scientist with the Singapore University of Technology and Design from 2018 to 2020. He is currently a part-time Professor with the Network Architectures and Services Group, Delft University of Technology. He has published several papers on quality aspects of voice-over-IP, video, and TCP performance. His work mainly focuses on the quality of multimedia services as perceived by users and on modeling network quality at the packet level. His main research interest is Quality of Service for IP networks such as the Internet.

Piet Van Mieghem (Senior Member, IEEE) received the master's and Ph.D. degrees in electrical engineering from K. U. Leuven, Belgium, in 1987 and 1991, respectively. He has been a Professor with the Delft University of Technology with a Chair of Telecommunication Networks, and a Chairman of the section Network Architectures and Services since 1998. Before joining Delft, he worked with the Interuniversity Micro Electronic Center from 1987 to 1991. From 1993 to 1998, he was a member of Alcatel Corporate Research Center, Antwerp where he was engaged in performance analysis of ATM systems and in network architectural concepts of both ATM networks (PNNI) and the Internet. He was a Visiting Scientist with the Department of Electrical Engineering, MIT, from 1992 to 1993, and a Visiting Professor with the Department of Electrical Engineering, UCLA in 2005; Center of Applied Mathematics, Cornell University in 2009; and Department of Electrical Engineering, Stanford University in 2015. He currently serves on the editorial board of the *Journal of Complex Networks* (Oxford University Press). He was member of the editorial board of the *Computer Networks* from 2005 to 2006, the IEEE/ACM TRANSACTIONS ON NETWORKING from 2008 to 2012, the *Journal of Discrete Mathematics* from 2012 to 2014, and the *Computer Communications* from 2012 to 2015.