

Optimal design and defense of networks under link attacks

Christophe Bravard, Liza Charroin, Corinne Touati

► **To cite this version:**

Christophe Bravard, Liza Charroin, Corinne Touati. Optimal design and defense of networks under link attacks. Working paper GATE 2015-19. dernière révision octobre 2016. 2015. <halshs-01176928v3>

HAL Id: halshs-01176928

<https://halshs.archives-ouvertes.fr/halshs-01176928v3>

Submitted on 7 Dec 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

WP 1519 – July 2015, revised October 2016

Optimal design and defense of networks under link attacks

Christophe Bravard, Liza Charroin, Corinne Touati

Abstract:

Networks facilitate the exchange of goods and information and create benefits. We consider a network with n complementary nodes, i.e. nodes that need to be connected to generate a positive payoff. This network may face intelligent attacks on links. To study how the network should be designed and protected, we develop a strategic model inspired by Dziubiński and Goyal (2013) with two players: a Designer and an Adversary. First, the Designer forms costly protected and non-protected links. Then, the Adversary attacks at most k links given that attacks are costly and that protected links cannot be removed by her attacks. The Adversary aims at disconnecting the network shaped by the Designer. The Designer builds a protected network that minimizes her costs given that it has to resist the attacks of the Adversary. We establish that in equilibrium the Designer forms a minimal 1-link-connected network which contains only protected links, or a minimal $(k + 1; n)$ -link-connected network which contains only non-protected links, or a network which contains one protected link and $(n-1)(k+1)/2$ non-protected links. We also examine situations where the Designer can only create a limited number of protected links and situations where protected links are imperfect, that is, protected links can be removed by attacks with some probabilities. We show that if the available number of protected links is limited, then, in equilibrium, there exists a network which contains several protected and non-protected links. In the imperfect defense framework, we provide conditions under which the results of the benchmark model are preserved.

Keywords:

Attacks on links, Network defense, Network design

JEL codes:

D74, D85

Optimal Design and Defense of Networks Under Link Attacks*

Christophe Bravard[†] Liza Charroin[‡] Corinne Touati[§]

October 7, 2016

Abstract

Networks facilitate the exchange of goods and information and create benefits. We consider a network with n complementary nodes, i.e. nodes that need to be connected to generate a positive payoff. This network may face intelligent attacks on links. To study how the network should be designed and protected, we develop a strategic model inspired by Dziubiński and Goyal (2013) with two players: a Designer and an Adversary. First, the Designer forms costly protected and non-protected links. Then, the Adversary attacks at most k links given that attacks are costly and that protected links cannot be removed by her attacks. The Adversary aims at disconnecting the network shaped by the Designer. The Designer builds a protected network that minimizes her costs given that it has to resist the attacks of the Adversary. We establish that in equilibrium the Designer forms a minimal 1-link-connected network which contains only protected links, or a minimal $(k + 1, n)$ -link-connected network which contains only non-protected links, or a network which contains one protected link and $\lceil (n - 1)(k + 1)/2 \rceil$ non-protected links. We also examine situations where the Designer can only create a limited number of protected links and situations where protected links are imperfect, that is, protected links can be removed by attacks with some probabilities. We show that if the available number of protected links is limited, then, in equilibrium, there exists a network which contains several protected and non-protected links. In the imperfect defense framework, we provide conditions under which the results of the benchmark model are preserved.

JEL Classification: D74, D85.

Key Words: Attacks on links, Network defense, Network design.

1 Introduction

Networks can be seen as communication structures. They are composed of nodes and links, where links represent the flow of information. Networks represent a crucial feature in our society, and are of particular interest in different fields such as military defense, telecommunication or computer networks. Some networks can be damaged by natural disasters or intelligent attacks. Attacks can affect nodes (agents, computers, telecommunication antennas, ...) or links (roads, communications flows, ...), and may disconnect a network.¹ In this paper, we examine a model where attacks target links. To illustrate the type of situations we model, consider a firm which has several production units (nodes of the network). Each production unit produces a part of the product and the pieces are assembled by a given production unit. The links of the network allow the parts of the product to be transferred among the units. If one unit is not connected to the rest of the units, its production cannot be transferred and the production

*This research has been supported by a grant of the French National Research Agency (ANR, FELIS Grant, ANR-14-CE28-0010-01). It was performed within the framework of the LABEX CORTEX (ANR-11-LABX-0042) of Université de Lyon, within the program Investissements d'Avenir (ANR-11-IDEX-007) operated by the French National Research Agency (ANR). Funding for this project was provided by a grant from Région Rhône-Alpes. The authors thank Pascal Billand, Hans Haller, Stéphane Sémirat, Jacques Durieu and an anonymous referee for their constructive comments and suggestions.

[†]Université Grenoble 2, UMR 1215 GAEL, F38000 Grenoble, France; CNRS, email: christophe.bravard@univ-grenoble-alpes.fr.

[‡]Univ Lyon, ENS de Lyon, GATE L-SE UMR 5824, F-69342 Lyon, France. E-mail: charroin@gate.cnrs.fr.

[§]Inria, F38330 Montbonnot Saint-Martin, France. email: corinne.touati@inria.fr.

¹A network is connected if no set of nodes is isolated from the others.

has no value. Recall that during the Second World War, the production units for the weapons (nodes) were buried, so they were impossible to target, and attacks had to target the roads (links) in order to destroy the production process of the enemy. Therefore, the issue was to design a communication network between the production units that the enemy could not disconnect.

Our goal is to examine how to design and protect the network in an optimal way, such that the network remains connected after an intelligent link attack.² We say that a network is designed and protected in an optimal way if the costs associated with the design and the protection of the network are minimized.

We consider a two-stage game with two players: a Designer (D) and an Adversary (A).

- Stage 1. The Designer moves first and chooses both a set of protected, and a set of non-protected links. Protected links cannot be removed by the attacks of the Adversary.
- Stage 2. After observing the protected network (strategy) formed by the Designer, the Adversary attacks the network by allocating attacks to specific links. Since the attacks are costly, the Adversary has an incentive to attack at most k links.

Creating protected and non-protected links is costly for the Designer. The benefits obtained by the Designer at the end of the game depend on the connectivity of the residual network, that is, the network obtained after the attack of the Adversary. If the residual network is connected, then the Designer wins the game: her benefits are equal to 1 and the benefits of the Adversary are 0. If the residual network is not connected, then the Adversary wins the game: her benefits are equal to 1 and the benefits of the Designer are 0. The payoffs obtained by the players are equal to the difference between their benefits and the costs associated with their strategies.³

We are interested in the Sub-game Perfect Equilibrium (SPE) of the two-stage game. We assume that the cost of protected links and non-protected links are sufficiently low so that the Designer has some profitable strategies which allow the residual network to be connected. First, we provide for each number of protected links, the minimal number of non-protected links that the Designer has to form in order to prevent the Adversary from disconnecting the network as well as a method to construct a solution network. Second, we establish that only three polar non-empty networks may arise in equilibrium in the benchmark model.

1. A minimal $(k + 1, n)$ -link-connected network which contains no protected links.⁴
2. A minimal $(1, n)$ -link-connected network which contains $n - 1$ protected links.
3. A network which contains one protected link and $\lceil (n - 1)(k + 1)/2 \rceil$ non-protected links.

The first family of networks constitutes the optimal strategy of the Designer when the cost of forming non-protected links is sufficiently low relative to that of forming protected links. The second one is the optimal strategy when the cost of forming non-protected links is sufficiently high relative to that of forming protected links. The third one is optimal for intermediate relative costs (cost of a protected link / cost of a non-protected link) when the number of nodes is odd and the number of attacks is even.

Additionally to the benchmark model described above, we study some variations of the game to develop a larger understanding of optimal design of protected networks. We take into account two types of limitations concerning protections. First, we consider that D cannot create as many protected links as in the benchmark model.⁵ Then, we consider a framework where each protected link has a probability π to be removed when it is attacked by A .⁶

In the framework where the number of protected links available for D is limited, we show that for intermediate relative costs, the optimal strategy of D consists in designing a network which contains both

²Note that an intelligent attack can also be seen as the worst case scenario.

³If we take again our military example, and assume that node $i - 1$ is the supplier of node i , then the Designer has to maintain a path between each pair of nodes $i - 1$ and i to obtain some end products. In other words, the residual network has to be connected to allow some production.

⁴A network g , which contains n nodes, is a minimal $(k + 1, n)$ -link-connected network, if it is not possible to disconnect it by removing k links, and such that there is no network which cannot be disconnected by removing k links and contains a smaller number of links.

⁵If we take again our military example, the Designer may not have enough resources to protect the whole network.

⁶Despite the effort of the Designer (of the army) to protect the communication flow, the Adversary (the enemy) may still be able to succeed in destroying protected links with some probabilities.

protected links and non-protected links. In the framework where protected links are removed by attacks with some probabilities, we provide conditions under which the results obtained in our benchmark model are preserved.

We now relate our paper to the existing literature on networks. This literature has become broader in the recent years (*Jackson* [18], *Goyal* [10] and *Vega-Redondo* [24]). The two seminal papers on the formation of social and economic networks are the paper of *Jackson and Wolinsky* [19] and the paper of *Bala and Goyal* [3]. *Bala and Goyal* [4] and *Haller and Sarangi* [14] introduce imperfectly reliable links in the *Bala and Goyal* [3] model. *Bala and Goyal* [4] show that, for certain ranges of linking cost and probability of failure, the equilibrium network is at least $(2, n)$ -link-connected, i.e. any two nodes are connected by at least two paths. *Haller and Sarangi* [14] extend the model of *Bala and Goyal* [4] by allowing heterogeneity in probabilities of link failure. These authors model random link failure but not an intelligent attack that seeks to interrupt the communication flow. In the present paper, we study the robustness of a network that must be designed and protected to resist an intelligent attack on links.

A growing literature on attacked networks studies the optimal strategy of a Designer whose network is under node attack. *Dziubinski and Goyal* [9](DG) study the optimal design and defense of networks under an intelligent attack. In their framework, there are two players: the Designer and the Adversary; the Designer can form links between n nodes, and protect these nodes to ensure their survival. The model we propose is close to the model of DG, with the following major differences:

- The Adversary attacks nodes in the DG’s framework while she attacks links in our framework;
- In our framework, the Designer wins the game if every node is able to communicate with any other node in the residual network. In the DG’s framework, the Designer wins the game if the residual network is connected regardless of the number of nodes removed by the Adversary. Thus, our setting is based on the complementarity of nodes while DG assume that nodes are substitutable.

DG show that in an SPE, the Designer protects 0 or 1 node. If the Designer protects 0 nodes, then she designs a minimal $(k + 1, n)$ -node-connected network.⁷ We obtain the same type of networks when the Designer uses no protection. At first sight, this result seems intriguing since the Adversary attacks nodes in DG’s paper and links in our paper. However, a minimal $(k + 1, n)$ -node-connected network defined in DG is also a network that contains the minimal number of links and resists the Adversary who attacks links. In DG’s paper, if the Designer protects nodes, she designs a star network⁸ and protects 1 node, the central node. In our framework, when D uses protections, she designs either a network which contains 1 protected link and $\lceil (n - 1)(k + 1)/2 \rceil$ non-protected links, or a network which contains $n - 1$ protected links. The results differ because in our framework every node needs to be connected with any other node in the residual network. Moreover, we establish that if we limit the number of available protections, then there exist optimal strategies for D where she designs networks which contain several protected and non-protected links. This result follows the fact that the number of non-protected links that each protected link allows the Designer to save is not constant.

DG examine imperfect defense through an example. They assume that the protections used by the Designer can fail when they are attacked by the Adversary. More precisely, an attack on an unprotected target always destroys the target, and an attack on a protected target destroys the target with a positive probability. A recent independent work of *Landwehr* [21] extends the analysis of imperfect defense. It shows that for a certain range of protection cost and link formation cost, strategies that use both protections and several links are equilibria.

Hoyer and De Jaegher [17] consider a framework where the Designer has to shape the network and form enough links in order to resist the attacks. In this framework, the Designer cannot protect specific parts of the network. The authors study the optimal way to design a network under link or node deletion with various cost levels. They show that if the costs of forming links are low, a regular network⁹ with a sufficient number of links is the optimal network for the Designer. If costs are high and links are attacked, then a star network is optimal for the Designer. The difference with our paper (except for the fact that they do not use protected links) is that in our framework, nodes are complementary and the Designer

⁷A minimal $(k + 1, n)$ -node-connected network is a network, which contains n nodes, that cannot be disconnected by removing k nodes, and such that there is no network which cannot be disconnected by removing k nodes and contains a smaller number of links.

⁸A star network is a network where one node, the central one, is linked with all other nodes, and other nodes are only linked with the central node.

⁹A network where all nodes have the same number of links.

cannot sacrifice any node to minimize her costs. *Haller* [13] extends the model of *Hoyer and De Jaegher* [17] by adding the possibility for two nodes to be connected by more than one link. In that case, it is harder for the Adversary to disconnect the network. Allowing multiple links between nodes can be seen as a different way to protect a connection between specific nodes than ours.

A part of the literature on attacked networks examines the role played by the contagion of attacks in networks. *Goyal and Vigier* [11] extend the work of DG by allowing the contagion of attacks (or threats). They find that the star network with a protected central node remains an equilibrium network. *Cabrales, Gottardi and Vega-Redondo* [6] and *Baccara and Bar-Isaac* [2] study the contagion of attacks in networks respectively in the field of financial firms where a financial risk can spread between connected firms and in the field of criminal networks where connectivity increases vulnerability because of external threats.¹⁰ *Cerdeiro, Dziubinski and Goyal* [7] and *Acemoglu, Malekian and Ozdaglar* [1] identify nodes to players. Specifically, *Cerdeiro, Dziubinski and Goyal* [7] propose a three-stage game. First, the Designer chooses the network. Second, each player observes the network and chooses independently and simultaneously if she invests in protection or not. Third, the Adversary observes the protected network and chooses the players to infect. In *Acemoglu, Malekian and Ozdaglar* [1] nodes/players are connected in a random network. Players have to invest in protection to be immune. Their investment depends on their links and the probability of being infected in the random network. This model allows to examine for instance the impact of a contagious disease on the individual behavior. These papers are different from the present one for two reasons. First, we study a framework where an attack on a link can remove only this specific link. Indeed, literature on contagious attacks reflects situations such as epidemics or virus spreading while our paper is focused on the study of specific link removal (for military strategies for instance). Second, in our model nodes cannot influence the architecture of the network by their decision.¹¹

The rest of the paper is organized as follows. In section 2, we present the model setup. In section 3, we present our main results. In section 4, we extend our model by examining a framework where the number of protected links available for the Designer is limited, and a framework where protected links have some probabilities to be removed by an attack. In section 5, we conclude.

2 Model setup

To simplify the notations, we set $\llbracket a, b \rrbracket = \{i \in \mathbb{N}, a \leq i \leq b\}$. Moreover, $\lfloor x \rfloor$ and $\lceil x \rceil$ are respectively the largest integer smaller or equal to x and the smallest integer larger or equal to x , and $\text{abs}(x) = \max\{-x, x\}$. Further, for every set X , $\#X$ is its cardinality.

Network. For any integer $n > 4$, let $N = \llbracket 1, n \rrbracket$ and $L(N)$ be the set of *unordered* pairs of N , i.e. $L(N) = \{(i, j) \in N \times N, i \neq j\}$. Throughout the paper, the elements of N are referred to as *nodes* while those of $L(N)$ are called *links*. An unordered pair $(i, j) \in L(N)$ is thus a link said to join nodes i and j and the link is denoted by ij . We introduce the notion of *protected network* as a triplet $g = (N, E_P, E_{NP})$ with $E_P \subseteq L(N)$, $E_{NP} \subseteq L(N)$ and $E_P \cap E_{NP} = \emptyset$. We call *protected links* the elements of E_P and *non-protected links* the elements of E_{NP} . Let G be the set of all protected networks. The significance of this refinement on the links will be made explicit in the two-player game formulation. To simplify the notations, we let $p = \#E_P$. In the rest of the paper, we will interchangeably use the term network or protected network.

¹⁰*McBride and Hewitt* [22] study the best way to dismantle a criminal network with imperfect information on its architecture. There also exists a literature which examines the particular cases of terrorist attacks, transportation network security, and homeland security (see *Brown, Carlyle, Salmeron and Wood* [5], *Tambe* [23], and *Hong* [16]).

¹¹Additionally to economics, several fields investigate problems close to the one we deal with. In an early graph theoretic work, *Harary* [15] exhibits a family of (k, n) -node-connected networks with a total number of links that is minimal. This family of networks is crucial to establish our results. *Groetschel, Monma and Stoer* [12] study a model where a firm has to prevent a communication network to be disconnected given that there exist possibilities of communication failure. As some connections may be interrupted, the firm has to design the least costly network that guarantees the best service for the consumers. Moreover, there also exists a literature on the design of survivable networks (see the survey of *Kerivin and Mahjoun* [20]) in Computer Science. *Cunningham* [8] studies network security and considers a model where the Designer allocates a different number of defense units to each link. A defended link has a level of resistance that depends on the number of defense units the Designer has allocated to it. The Adversary allocates attack units to remove a link. A link is removed if more attack units than defense units have been allocated to this link. The author proposes an algorithm which exhibits how some links have to be reinforced in order to protect the network.

For any network g , let $E_P(g)$ (respectively $E_{NP}(g)$) refer to the set of protected (respectively non-protected) links of g , and $E(g) = E_P(g) \cup E_{NP}(g)$. If there exists a link between i and j in g (i.e. if $ij \in E(g)$), then i and j are called *adjacent*. For each node i , $d_i(g)$ is its *degree* in g , that is the number of links incident to i in g : $d_i(g) = \#\{ij \in E(g)\}$. A *path* between two nodes i_0 and i_q of a network g is a finite alternating sequence of nodes and *distinct* links: $i_0, i_0i_1, i_1, i_1i_2, i_2, \dots, i_{q-1}i_q, i_q$ where $i_\ell \in N$ for all $\ell \in \llbracket 0, q \rrbracket$ and $i_\ell i_{\ell+1} \in E(g)$ for all $\ell \in \llbracket 0, q-1 \rrbracket$. A *cycle* is a path where $i_0 = i_q$. Finally, a network $g = (N, E_P, E_{NP})$ is *connected* if there exists a path between any two nodes $i, j \in N$. We say that network $g' = (N', E'_P, E'_{NP})$ is a subnetwork of $g = (N, E_P, E_{NP})$ if $N' \subseteq N$, $E'_P \subseteq E_P$ and $E'_{NP} \subseteq E_{NP}$. Subnetwork $g' = (N', E'_P, E'_{NP})$ is a component of network g if g' is connected and if there is no connected subnetwork $g'' = (N'', E''_P, E''_{NP})$ of g , with $g'' \neq g'$ and such that $N' \subseteq N''$ or $E(g') \subseteq E(g'')$. By convention, a node $i \in N$ such that $d_i(g) = 0$ is a component.

Two-player game. The players are the Designer (D) and the Adversary (A). We consider a two-stage game where D plays first and A moves at the second stage. Given N , a strategy s^D for D is (identified with) a protected network (N, E_P^D, E_{NP}^D) . In other words, D chooses to create some links from $L(N)$ and to protect some of them:

$$s^D = (N, E_P^D, E_{NP}^D), \quad E_P^D \subseteq L(N), \quad E_{NP}^D \subseteq L(N), \quad \text{and} \quad E_P^D \cap E_{NP}^D = \emptyset.$$

A strategy for the Adversary is a mapping that assigns to each protected network g a subset of links E^A of $E(g)$. In other words, A chooses to attack some links of g :

$$s^A : \begin{cases} G \rightarrow 2^{L(N)}, \\ g \mapsto s^A(g) = E^A, \end{cases} \quad \text{with} \quad s^A(g) \subseteq E(g).$$

Residual network and benefits. At the first stage D designs a protected network s^D . Then, the attack of A leads to a second protected network of the form $g^R = (N, E_P^D, E_{NP}^D \setminus E^A)$, which we call *residual network*. Note that, by construction, g^R is a subnetwork of s^D . The benefits of D are given by

$$\phi(g^R) = \begin{cases} 1, & \text{if } g^R \text{ is connected,} \\ 0, & \text{otherwise.} \end{cases} \quad (1)$$

Network and costs. We assume that attacking a link has a unitary cost c_A . Therefore, the cost of the Adversary associated with $s^A(g) = E^A$ is

$$c^A(E^A) = c_A \#E^A, \quad (2)$$

where $c_A \in [1/(n-3), 1]$.¹² Note that the cost of a strategy is less than 1 if and only if the Adversary attacks less than $k = \lfloor 1/c_A \rfloor$ links.

Similarly, both protected and non-protected links are costly to create: each protected link has a strictly positive cost $c_P > 0$ and each non-protected link has a strictly positive cost $c_L > 0$. We assume that $c_P > c_L$. The cost of a strategy s^D of the Designer is thus:

$$c^D(s^D) = c_P \#E_P^D + c_L \#E_{NP}^D. \quad (3)$$

If the cost of creating protected or non-protected links is too large, then D cannot use a strategy where she forms protected or non-protected links. Therefore, to obtain non trivial results, we assume that the costs of creating protected and non-protected links are sufficiently low: $c_P < 1/(n-1)$ and $c_L < 1/(n(n-1)/2)$.¹³

¹²It will be clear in the following that if $c_A < 1/(n-3)$, then A may attack at least $n-2$ links. Due to our assumptions on the cost of protected and non-protected links and on the payoff function of D , if A can attack $n-2$ links, then in equilibrium, the only protected network without protected links that D may design is the complete network. Moreover, if A can attack strictly more than $n-2$ links, then in equilibrium D cannot design any protected network without protected links.

¹³It will be clear in the following that when $c_P < 1/(n-1)$ and $c_L < 1/(n(n-1)/2)$, then D builds a non-empty protected network in equilibrium.

Payoffs. The payoff of the Designer for choosing s^D when the Adversary responds with s^A is:

$$\Pi^D(s^D, s^A(s^D)) = \phi(g^R) - c^D(s^D). \quad (4)$$

Since $c_P < 1/(n-1)$, D obtains a strictly positive payoff when she designs a network with $n-1$ protected links and 0 non-protected links and the residual network is connected. Similarly, since $c_L < 1/(n(n-1)/2)$, D obtains a strictly positive payoff when she designs a network with 0 protected links and $n(n-1)/2$ non-protected links and the residual network is connected.

The payoff of the Adversary is

$$\Pi^A(s^D, s^A(s^D)) = 1 - \phi(g^R) - c^A(s^A(s^D)). \quad (5)$$

If there exist two strategies of A that lead to the same payoff, A chooses the one having the highest value of $\#E^A$.¹⁴

In a nutshell, in our framework the objective of the Designer is to obtain a connected residual network at a minimal cost. The objective of the Adversary is to obtain a residual network that is disconnected. Note that A does not attack strictly more than $n-3$ links. Indeed, A obtains $\Pi^A(s^D, s^A(s^D)) \leq 1 - \lfloor 1/c_A \rfloor \#E^A < 0$, when $\#E^A > n-3$, while A obtains a payoff equal to zero when she attacks no links. We now provide some illustrations where equation (1) captures the benefits of D . Suppose that D has n production units identified to nodes. Let y_i be the output of production unit i , and δ_i be such that $\delta_i = 1$ if there is a path between $i \in \llbracket 2, n \rrbracket$ and production unit $i-1$, and $\delta_i = 0$ otherwise. Here, production unit $i-1$ can be interpreted as the *unique* supplier of production unit i . We assume that $y_1 = \gamma$, $\gamma > 0$, and $y_i = \delta_i y_{i-1}$ for $i \in \llbracket 2, n \rrbracket$. If the total output obtained by D from the production units is $Y = y_n$, then the total output function is in line with the benefits function of D . The same conclusion occurs if we assume $Y = \min_{i \in N} \{y_i\}$ or $Y = \prod_{i \in N} (y_i)^{\rho_i}$ with $\rho_i > 0$.

We now provide another example. Let nodes be identified to cities and links be identified to communication flows between cities. Public authorities may have an incentive to maintain communication between all the cities when some communication flows are broken because of a natural disaster or a strategic attack. Indeed, if some cities are isolated from the others, then it is difficult for the public authorities to rescue inhabitants of these cities.

Sub-game Perfect Equilibrium (SPE). An SPE is a pair (s_*^D, s_*^A) that prescribes the following strategic choices. At Stage 2, A plays a best response $s_*^A(s^D)$ to $s^D \in G$:

$$s_*^A(s^D) \in \operatorname{argmax}_{X \subseteq E(s^D)} \{\Pi^A(s^D, X)\}.$$

Note that $s_*^A(s^D) \subseteq E_{NP}^D$ since attacks cannot remove protected links. Let $g_*^R(s^D)$ be the residual network obtained when D plays strategy s^D and A plays a best response to s^D , that is $s_*^A(s^D)$. Given the best response outcome $g_*^R(s^D)$, D achieves payoff $\phi(g_*^R(s^D)) - c^D(s^D)$ when choosing s^D . At stage 1, D plays s_*^D such that

$$s_*^D \in \operatorname{argmax}_{X \in G} \{\Pi^D(X, s_*^A(X))\}.$$

Specific architectures. The *empty network* is the network which contains no links. A *tree* is a connected and acyclic network. A network g which contains n nodes is a (κ, n) -link-connected network if any subnetwork g' obtained from g by removing $\kappa-1$ links is connected, and there exists a subnetwork g' obtained from g by removing κ links that is not connected. Let $\mathcal{G}(\kappa, n)$ be the set of minimal (κ, n) -link-connected networks with n nodes, i.e. if $g \in \mathcal{G}(\kappa, n)$, then there does not exist a (κ, n) -link-connected network, g' , such that $\#E(g') < \#E(g)$. It is easy to see that every node i of a network $g \in \mathcal{G}(\kappa, n)$ satisfies $d_i(g) \geq \kappa$, as otherwise it could be separated by removing all links incident to i . Consequently, the number of links in a minimal (κ, n) -link-connected network, $\kappa \geq 2$, is at least $\lceil n\kappa/2 \rceil$. As was shown by Harary [15], this condition is also sufficient. The proof of this result is constructive – Harary describes how to obtain a family of solution graphs when $\kappa \geq 2$. The minimal (κ, n) -link-connected

¹⁴In particular, note that the strategy \emptyset for the Adversary leads to a payoff that equals zero. If $1/c_A$ is an integer, then there may exist a strategy such that $\#E^A = 1/c_A = k$ that disconnect the network. That strategy also has a payoff that equals zero and is chosen by the Adversary according to the tie-breaking rule (if no strategy with $\#E^A < k$ can disconnect the network).

networks described by Harary are called (κ, n) -Harary-networks. To give the reader some idea of what (κ, n) -Harary-networks look like, we provide some examples in Figure 1 with 5 nodes. For full description of the construction the interested reader is referred to Harary [15].

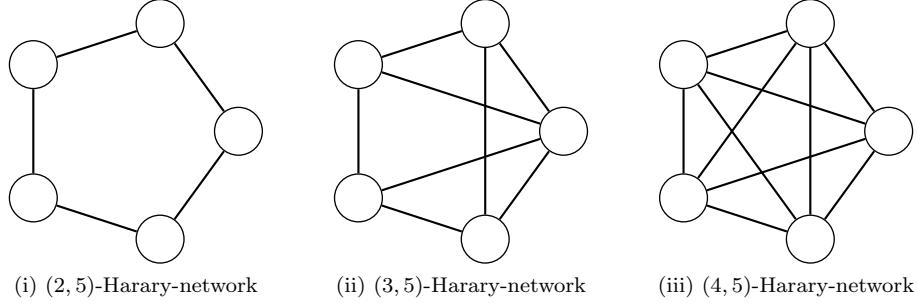


Figure 1: Examples of (κ, n) -Harary-networks.

Specific strategies. We define specific strategies that play a crucial role in the rest of the paper. $(N, \emptyset, \emptyset)$ is the *empty* network. Let $S_{p,k}^D$ be the set of protected networks that (i) cannot be disconnected by k attacks, (ii) contain p protected links, and (iii) contain a minimal number of non-protected links. Protected networks which belong to the same set induce the same costs for D . Protected networks associated with two specific values of p play a crucial role in our analysis.

- $S_{0,k}^D$ is the set of protected networks which contain no protected link, and which are minimal $(k+1, n)$ -link-connected networks. For instance, $(k+1, n)$ -Harary-networks belong to $S_{0,k}^D$.
- $S_{n-1,k}^D$ is the set of protected networks $s^D = (N, E_P, E_{NP})$ which only contain protected links ($E_{NP} = \emptyset$) and such that (N, E_P, \emptyset) is a tree.¹⁵

3 Model Analysis

Our first result provides, for any number of protected links, the minimal number of non-protected links that the Designer has to form in order to prevent the Adversary from disconnecting the network.

To establish the first result, for any pair $(n, k) \in \mathbb{N} \times \llbracket 1, n-3 \rrbracket$, we set $p_1(k, n)$ and $p_2(k, n)$ as follows:

$$\begin{cases} \Delta = (3k+5)^2 - 8n(k+1) \\ p_1(k, n) = \left\lfloor \frac{4n-3k-5-\sqrt{\Delta}}{8} \right\rfloor + 1 & \text{if } \Delta \geq 0, & \text{and } p_1(k, n) = -1 \text{ otherwise,} \\ p_2(k, n) = \left\lfloor \frac{4n-3k-5+\sqrt{\Delta}}{8} \right\rfloor - 1 & \text{if } \Delta \geq 0, & \text{and } p_2(k, n) = -1 \text{ otherwise.} \end{cases}$$

When no confusion is possible, we let $p_1 = p_1(k, n)$ and $p_2 = p_2(k, n)$.

Proposition 1 *Suppose that A attacks exactly k links in an optimal way. Let $n_1(p, k) = \left\lceil \frac{(n-p)(k+1)}{2} \right\rceil$ and $n_2(p, k) = (n-2p)(k+1+p) + (n-1)p - \frac{n(n-1)}{2}$. For $s^D \in S_{p,k}^D$,*

¹⁵Indeed, if the subnetwork of s^D , (N, E_P, \emptyset) , is a tree and $E_{NP} = \emptyset$, then $s^D \in S_{n-1,k}^D$ cannot be disconnected by k attacks. Otherwise – that is if (N, E_P, \emptyset) contains a cycle – the network is not connected unless $\#E_{NP}^D > 0$. Therefore, that network does not satisfy condition (iii) and thus is not in $S_{n-1,k}^D$.

$$\#E_{NP}^D = \begin{cases} n_1(p, k), & \text{for } p \in \llbracket 0, n-2 \rrbracket \setminus \llbracket p_1(k, n), p_2(k, n) \rrbracket, \\ n_2(p, k), & \text{for } p \in \llbracket 0, n-2 \rrbracket \cap \llbracket p_1(k, n), p_2(k, n) \rrbracket, \\ 0, & \text{for } p = n-1. \end{cases} \quad (6)$$

So, if D forms p protected links and A attacks k links, then the optimal cost function associated with the pair (p, k) is

$$C^*(p, k) = c_L \#E_{NP}^D + c_P p, \text{ with } s^D \in S_{p,k}^D. \quad (7)$$

Proof The proof is given in Appendix. \square

Let us provide the intuition of Proposition 1. If D forms $n-1$ protected links, then there exists a set of strategies (i.e. a set of protected networks) for D that allows to resist k attacks without non-protected links, $S_{n-1,k}^D$. Otherwise, let D form $p \in \llbracket 0, n-2 \rrbracket$ protected links and build a protected network $s^D = (N, E_P, E_{NP})$ in $S_{p,k}^D$.

First, (N, E_P, \emptyset) is acyclic. Indeed, if s^D contains a cycle, then there exists a protected link, say ij , that can be removed without altering the fact that s^D resists k attacks. Hence, it is possible for D to remove the protected link ij and replace a non-protected link $i'j'$ by the protected link $i'j'$, and so reduce the number of non-protected links.

Second, consider the following sequence of networks: $g_0 = (N, \emptyset, \emptyset)$ and for any $\ell \in \llbracket 1, p \rrbracket$, $g_\ell = (N, E_P(g_{\ell-1}) \cup ij, \emptyset)$ for some $ij \in E_P \setminus E_P(g_{\ell-1})$. Hence $g_p = (N, E_P, \emptyset)$. Since there is no cycle in (N, E_P, \emptyset) , then, for any ℓ , the extra link of $E_P(g_\ell)$ allows to merge two components of $E_P(g_{\ell-1})$. Since g_0 has n components, then by an immediate recurrence, (N, E_P, \emptyset) has exactly $n-p$ components.

Third, observe that each component of (N, E_P, \emptyset) has to be incident to at least $k+1$ non-protected links, otherwise A can disconnect the protected network s^D with k attacks. Since there are $n-p$ such components, $n_1(p, k)$ is the minimal number of non-protected links to form. In the following example, we illustrate that there exist some values of p and k for which $n_1(p, k)$ non-protected links are sufficient to resist k attacks.

Example 1 Suppose $N = \llbracket 1, 10 \rrbracket$, $k = 6$ and $p = 5$. We describe a strategy s^D where $\#E_{NP}^D = n_1(5, 6)$. Consider the networks of Figure 2 and s^D the protected network such that g' , g'' , g''' are all subnetworks of s^D , with $E_P^D = E_P(g''')$ and $E_{NP}^D = E_{NP}(g') \cup E_{NP}(g'')$. Note that subnetwork g' is a complete network and g'' is a $(3, 5)$ -Harary-network. Finally, we observe that each component in (N, E_P, \emptyset) is incident to at least 7 non-protected links and there is no possibility for A to disconnect s^D with 6 attacks.

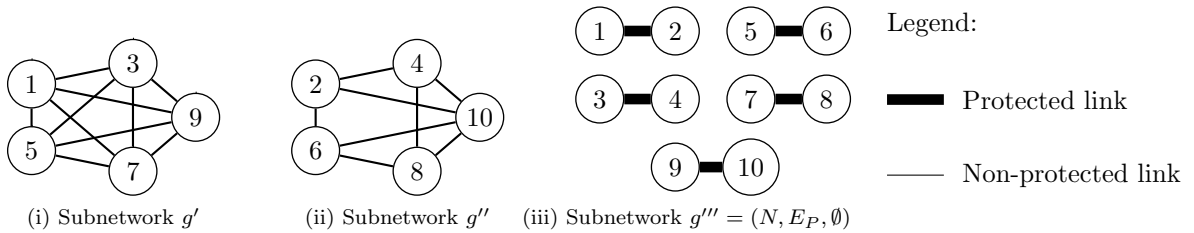


Figure 2: Subnetworks associated with Example 1: the solution network is (N, E_P, E_{NP}) with $E_{NP} = E_{NP}(g') \cup E_{NP}(g'')$.

Recall that each component of (N, E_P, \emptyset) has to be incident to at least $k+1$ non-protected links in order to resist k attacks. This fact implies that for some parameters p and k , networks in $S_{p,k}^D$ contains $n_2(p, k)$ non-protected links. We illustrate these cases through the following example.

Example 2 Suppose $N = \llbracket 1, 10 \rrbracket$, $k = 7$, and $p = 2$. We assume that D forms a protected network s^D in $S_{p,k}^D$, with a protected link between nodes 1 and 2 and between nodes 3 and 4. Then s^D contains 8

components. We denote by \mathcal{C} the component which contains nodes 1 and 2, and by \mathcal{C}' the component which contains nodes 3 and 4. Each of these components has to be incident to 8 non-protected links. Networks g' and g'' given in Figure 3 are subnetworks of s^D , with $E_P^D = E_P(g')$ and $E_{NP}^D = E_{NP}(g') \cup E_{NP}(g'')$. Each node $a \in \llbracket 5, 10 \rrbracket$ can form at most 5 non-protected links with each other. Hence each of them has to form at least 3 non-protected links with nodes in $\llbracket 1, 4 \rrbracket$. Altogether, nodes 5, 6, 7, 8, 9 and 10 have to form a total of $3 \times 6 = 18$ non-protected links with nodes 1, 2, 3 and 4. As for components \mathcal{C} and \mathcal{C}' , they should be incident to a total of 16 non-protected links. Hence, D cannot form a protected network with $n_1(2, 7)$ protected links that resists k attacks. More precisely, a network that cannot be disconnected by A with 7 attacks has to contain $18 + (5 \times 6)/2 = 33$ non-protected links while $n_1(2, 7) = 32$.

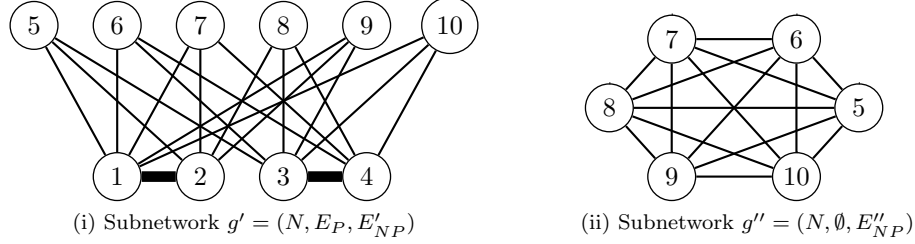


Figure 3: Subnetworks associated with Example 2. The solution is (N, E_P, E_{NP}) with $E_{NP}(g) = E'_{NP} \cup E''_{NP}$.

In Example 2, we have assumed that D uses protected links to form two components of size two in (N, E_P, \emptyset) . Now, assume instead that (N, E_P, \emptyset) has a unique component of size strictly greater than 1. For instance, D forms protected links between nodes 1 and 2 and between nodes 2 and 3. Then, each node $a \in \llbracket 4, 10 \rrbracket$ can form at most 6 non-protected links with other nodes in $\llbracket 4, 10 \rrbracket$. Consequently, since $k = 7$ each node $a \in \llbracket 4, 10 \rrbracket$ has to form at least 2 non-protected links with nodes in $\llbracket 1, 3 \rrbracket$. The component which contains nodes 1, 2 and 3 is incident to at least 14 non-protected links. It follows that D forms $21 + 14 = 35$ non-protected links instead of 33 links in Example 2. This example illustrates how D designs (N, E_P, \emptyset) in order to minimize the costs of forming links.

We now generalize Example 2 to provide some intuition for $p_1(k, n)$ and $p_2(k, n)$. Consider a protected network in $S_{p,k}^D$ where each component of (N, E_P, \emptyset) contains either one or two nodes. There are thus $n - 2p$ components of size 1 and p components of size 2. We observe that components of size 1 need to be incident to at least $k + 1$ non-protected links. Since the number of links between a component of size 1 and other components of size 1 is at most $(n - 2p - 1)$, the total number of non-protected links between components of size 1 and those of size 2 is at least equal to $(n - 2p)((k + 1) - (n - 2p - 1))$. Moreover, to minimize the number of links, the total number of the non-protected links incident to components of size 2 should be equal to $(k + 1)p$. Let x_1 and x_2 be the roots (when they exist) of $(n - 2x)((k + 1) - (n - 2x - 1)) = (k + 1)x$. Since $p_1(k, n) = \lfloor x_1 \rfloor + 1$ and $p_2(k, n) = \lceil x_2 \rceil - 1$, the number of non-protected links required to resist k attacks is given by $n_2(p, k)$ when $p \in \llbracket 0, n - 2 \rrbracket \cap \llbracket p_1(k, n), p_2(k, n) \rrbracket$.

We now characterize the SPE according to the costs of forming protected and non-protected links.

Proposition 2 *Let the payoff functions be given by equations (4) and (5), and let (s_*^D, s_*^A) be an SPE.¹⁶*

1. Suppose that $n(k + 1) \bmod 2 = 0$.

¹⁶The case of equality follows the same pattern, that is:

1. If $n(k + 1) \bmod 2 = 0$ and $\frac{c_P}{c_L} = \left(\frac{n-1}{n-1}\right) \left(\frac{k+1}{2}\right)$, then $s_*^D \in S_{n-1,k}^D \cup S_{0,k}^D$ and $s_*^A(s_*^D) = \emptyset$.
2. Suppose that $n(k + 1) \bmod 2 = 1$.
 - (a) If $\frac{c_P}{c_L} = \left(\frac{n-1}{n-2}\right) \left(\frac{k+1}{2}\right)$, then $s_*^D \in S_{n-1,k}^D \cup S_{1,k}^D$ and $s_*^A(s_*^D) = \emptyset$.
 - (b) If $\frac{k+2}{2} = \frac{c_P}{c_L}$, then $s_*^D \in S_{0,k}^D \cup S_{1,k}^D$ and $s_*^A(s_*^D) = \emptyset$.

(a) If $\frac{c_P}{c_L} < \binom{n}{n-1} \binom{k+1}{2}$, then $s_*^D \in S_{n-1,k}^D$ and $s_*^A(s_*^D) = \emptyset$.

(b) If $\binom{n}{n-1} \binom{k+1}{2} < \frac{c_P}{c_L}$, then $s_*^D \in S_{0,k}^D$ and $s_*^A(s_*^D) = \emptyset$.

2. Suppose that $n(k+1) \bmod 2 = 1$.

(a) If $\frac{c_P}{c_L} < \binom{n-1}{n-2} \binom{k+1}{2}$, then $s_*^D \in S_{n-1,k}^D$ and $s_*^A(s_*^D) = \emptyset$.

(b) If $\binom{n-1}{n-2} \binom{k+1}{2} < \frac{c_P}{c_L} < \frac{k+2}{2}$, then $s_*^D \in S_{1,k}^D$ and $s_*^A(s_*^D) = \emptyset$.

(c) If $\frac{k+2}{2} < \frac{c_P}{c_L}$, then $s_*^D \in S_{0,k}^D$ and $s_*^A(s_*^D) = \emptyset$.

Proof Let $\Delta = (3k+5)^2 - 8n(k+1)$, by straightforward calculations, we have $\left\lfloor \frac{4n-3k-5-\sqrt{\Delta}}{8} \right\rfloor \geq 0$. Hence, we have either $p_1(k,n) = p_2(k,n) = -1$, or $p_1(k,n) \geq 1$. Therefore, from Proposition 1, if D builds a protected network in $S_{0,k}^D$, then her cost equals $C^*(0,k) = n_1(0,k)c_L$. Similarly, by Proposition 1, $C^*(n-1,k) = (n-1)c_P$. Moreover, by Lemma 3, we know that $n_2(p,k) \geq n_1(p,k)$ when $p \in \llbracket p_1(k,n), p_2(k,n) \rrbracket$. Hence by Proposition 1, for all $p \in \llbracket 1, n-2 \rrbracket$, $C^*(p,k) \geq n_1(p,k)c_L + pc_P$. We prove successively the two parts of the proposition.

1. Let $n(k+1) \bmod 2 = 0$. For $p \in \llbracket 1, n-2 \rrbracket$, $C^*(p,k) - C^*(0,k) \geq p(c_P - c_L(k+1)/2) > 0$, if $c_P/c_L > (k+1)/2$. Moreover, $C^*(n-1,k) - C^*(0,k) > 0$ if $c_P/c_L > \binom{n}{n-1} \binom{k+1}{2}$. Assume that $c_P/c_L > \binom{n}{n-1} \binom{k+1}{2}$. Then, $c_P/c_L > (k+1)/2$ and $s_*^D \in S_{0,k}^D$.

For $p \in \llbracket 1, n-2 \rrbracket$, $C^*(p,k) - C^*(n-1,k) > 0$ if $c_P/c_L < \binom{n-p}{n-p-1} \binom{k+1}{2}$. Assume that $c_P/c_L < \binom{n}{n-1} \binom{k+1}{2}$. Then, $c_P/c_L < \binom{n-p}{n-p-1} \binom{k+1}{2}$ for all $p \in \llbracket 1, n-2 \rrbracket$. Consequently, $s_*^D \in S_{n-1,k}^D$.

By assumption, $C^*(0,k) < 1$ and $C^*(n-1,k) < 1$. It follows that D has an incentive to build a protected network in $S_{0,k}^D$ or $S_{n-1,k}^D$ since she obtains benefits equal to 1. Hence, A cannot disconnect s_*^D with k attacks. Consequently, $s_*^A(s_*^D) = \emptyset$.

2. Let $n(k+1) \bmod 2 = 1$. By Proposition 1, we know that:

$C^*(0,k) - C^*(1,k) > 0$ if $c_P/c_L < (k+2)/2$,

$C^*(1,k) - C^*(n-1,k) > 0$ if $c_P/c_L < \binom{n-1}{n-2} \binom{k+1}{2}$, and

$C^*(0,k) - C^*(n-1,k) > 0$ if $c_P/c_L < \frac{1}{n-1} \left\lceil \frac{n(k+1)}{2} \right\rceil$.

By using the same argument as in the previous point, we establish that for $p \in \llbracket 2, n-2 \rrbracket$, if $s_*^D \in S_{p,k}^D$, then s_*^D is not an optimal strategy for D . Assume that $c_P/c_L < \binom{n-1}{n-2} \binom{k+1}{2}$. Then

$c_P/c_L < \binom{n}{n-1} \binom{k+1}{2}$ and $s_*^D \in S_{n-1,k}^D$. Assume that $c_P/c_L > \binom{n-1}{n-2} \binom{k+1}{2}$. There are two possible cases. If $c_P/c_L < (k+2)/2$, then $C^*(1,k) - C^*(n-1,k) < 0$ and $C^*(1,k) - C^*(0,k) < 0$, and $s_*^D \in S_{1,k}^D$. If $c_P/c_L > (k+2)/2$, then $C^*(0,k) - C^*(n-1,k) < 0$ and $C^*(0,k) - C^*(1,k) < 0$, and $s_*^D \in S_{0,k}^D$.

Since D has an incentive to build a network s_*^D in $S_{0,k}^D$, $S_{1,k}^D$ or $S_{n-1,k}^D$, A cannot disconnect s_*^D with k attacks. Consequently, $s_*^A(s_*^D) = \emptyset$. □

Let us provide the intuition of Proposition 2. Note that if D builds a protected network that A cannot disconnect with k attacks (which is the maximal number of attacks that A has an incentive to choose), then A attacks no link in an SPE since each attack is costly. Also, due to the costs of forming protected and non-protected links, in an SPE D always has an incentive to build a protected network that A cannot disconnect.

We now compare the costs of protected networks that belong to different sets $S_{p,k}^D$, $p \in \llbracket 0, n-1 \rrbracket$, with respect to the values of c_L and c_P .

First, we consider point 1 of Proposition 2: $n(k+1) \bmod 2 = 0$. In Figure 4, we draw lines $(d_{p,p'})$ whose slopes $\sigma_{p,p'}$ can be interpreted as the *average number* of non-protected links that each protected link allows to save between a protected network in $S_{p,k}^D$, $p \in \llbracket 0, n-1 \rrbracket$, and a protected network that belongs to $S_{p',k}^D$, $p' \neq p$. We draw four such lines:

- Line $(d_{0,p'})$ whose slope $\sigma_{0,p'}$ shows the average saving of non-protected link *per protected link* between networks in $S_{0,k}^D$ and those of $S_{p',k}^D$, for $p' \in \llbracket 1, n-2 \rrbracket \setminus \llbracket p_1, p_2 \rrbracket$, and similarly:
- Line $(d_{0,\tilde{p}})$ of slope $\sigma_{0,\tilde{p}}$ between networks in $S_{0,k}^D$ and in $S_{\tilde{p},k}^D$, with $\tilde{p} \in \llbracket 0, n-2 \rrbracket \cap \llbracket p_1, p_2 \rrbracket$
- Line $(d_{0,n-1})$ of slope $\sigma_{0,n-1}$ between networks in $S_{0,k}^D$ and in $S_{n-1,k}^D$.
- Line $(d_{n-1,p'})$ of slope $\sigma_{n-1,p'}$ between networks in $S_{n-1,k}^D$ and in $S_{p',k}^D$, with $p' \in \llbracket 1, n-2 \rrbracket$.

Observe that for any $p' \in \llbracket 1, n-2 \rrbracket \setminus \llbracket p_1, p_2 \rrbracket$ and any $\tilde{p} \in \llbracket 0, n-2 \rrbracket \cap \llbracket p_1, p_2 \rrbracket$, $\text{abs}(\sigma_{0,\tilde{p}}) < \text{abs}(\sigma_{0,p'})$. Similarly, for any $p' \in \llbracket 1, n-2 \rrbracket \setminus \llbracket p_1, p_2 \rrbracket$, $\text{abs}(\sigma_{0,p'}) < \text{abs}(\sigma_{0,n-1})$. Moreover, for any $p \in \llbracket 1, n-2 \rrbracket$, $\text{abs}(\sigma_{0,n-1}) < \text{abs}(\sigma_{n-1,p'})$.

Suppose $c_P/c_L > \text{abs}(\sigma_{0,n-1})$. Then, costs of forming links with a strategy in $S_{0,k}^D$ are lower than the costs of forming links with a strategy in $S_{n-1,k}^D$. Moreover, $c_P/c_L > \text{abs}(\sigma_{0,n-1}) > \text{abs}(\sigma_{0,p'}) > \text{abs}(\sigma_{0,\tilde{p}})$, with $p' \in \llbracket 1, n-2 \rrbracket \setminus \llbracket p_1, p_2 \rrbracket$ and $\tilde{p} \in \llbracket 0, n-2 \rrbracket \cap \llbracket p_1, p_2 \rrbracket$. It follows that the costs of forming links are minimized for strategies in $S_{0,k}^D$. Conversely, suppose $c_P/c_L < \text{abs}(\sigma_{0,n-1})$. Then $c_P/c_L < \text{abs}(\sigma_{n-1,p'})$, for $p' \in \llbracket 1, n-2 \rrbracket$. It follows that the costs of forming links are minimized for strategies in $S_{n-1,k}^D$. Finally, note that $\text{abs}(\sigma_{0,n-1}) = \binom{n}{n-1} \binom{k+1}{2}$.

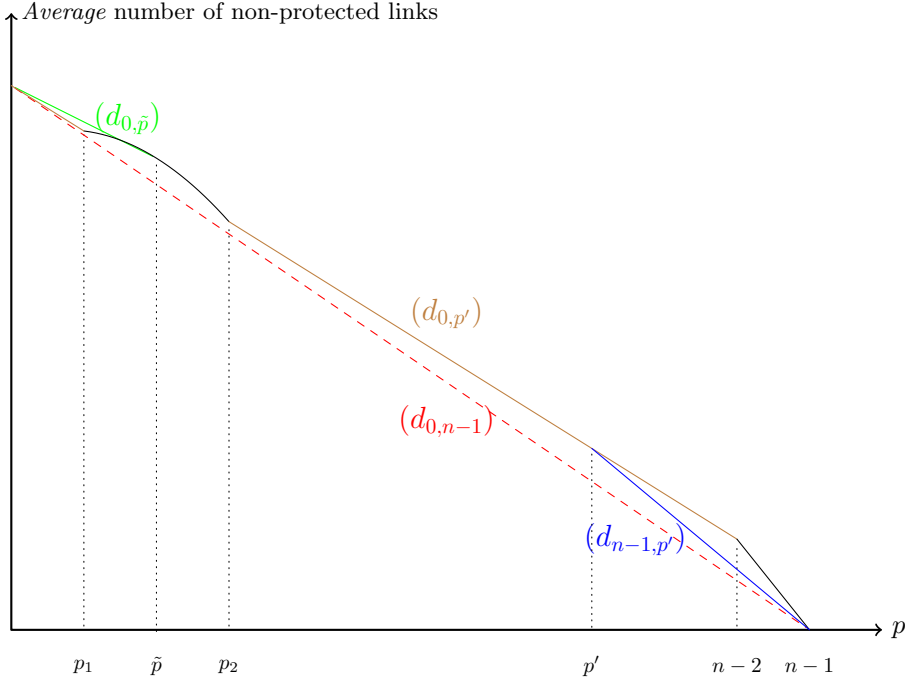


Figure 4: Intuition of Proposition 2 when $n(k+1) \bmod 2 = 0$.

We now consider the second part of Proposition 2: $n(k+1) \bmod 2 = 1$. The intuition is similar to the case where $n(k+1) \bmod 2 = 0$ except for protected networks which belong to $S_{0,k}^D$ and $S_{1,k}^D$. Consequently, we focus only on three sets of protected networks: $S_{0,k}^D$, $S_{1,k}^D$ and $S_{n-1,k}^D$. In Figure 5, the slope $\sigma_{0,1}$ of $(d_{0,1})$, can be interpreted as the number of non-protected links that the protected link allows to save between a protected network in $S_{0,k}^D$ and a protected network in $S_{1,k}^D$. The same interpretation is valid for the slope $\sigma_{1,n-1}$ of $(d_{1,n-1})$, which relates a protected network in $S_{1,k}^D$ and a protected network in $S_{n-1,k}^D$, and for the slope $\sigma_{0,n-1}$ of the line $(d_{0,n-1})$, which relates a protected network in $S_{0,k}^D$ and a protected network in $S_{n-1,k}^D$.

Suppose that $c_P/c_L < \text{abs}(\sigma_{1,n-1})$. Then, $c_P/c_L < \text{abs}(\sigma_{0,n-1})$, and networks in $S_{n-1,k}^D$ have a minimal link formation cost. Conversely, suppose that $c_P/c_L > \text{abs}(\sigma_{1,n-1})$. If $k < n-3$,¹⁷ then there

¹⁷For $k = n-3$, $\frac{k+2}{2} = \binom{n-1}{n-2} \binom{k+1}{2}$, and thus the case 2.(b) of Proposition 2 never occurs.

are two possibilities. If $c_P/c_L > \text{abs}(\sigma_{0,1})$, then protected networks in $S_{0,k}^D$ minimize the cost of forming links. If $c_P/c_L < \text{abs}(\sigma_{0,1})$, then protected networks in $S_{1,k}^D$ minimize the cost of forming links. Finally, note that $\text{abs}(\sigma_{0,1}) = (k+2)/2$, and $\text{abs}(\sigma_{1,n-1}) = \binom{n-1}{n-2} \binom{k+1}{2}$.

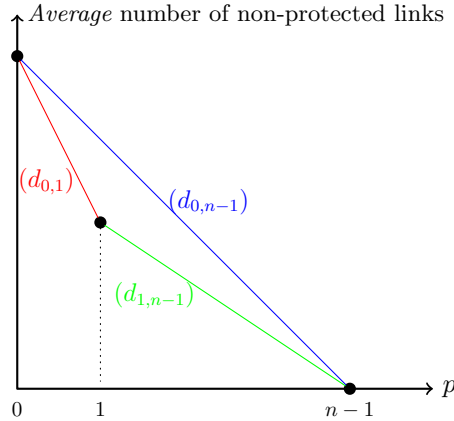


Figure 5: Intuition of Proposition 2 when $n(k+1) \bmod 2 = 1$ and $k < n-3$.

We now compare the results obtained in our framework where A attacks links and the results obtained in DG's framework where A attacks nodes (Proposition 1 in [9]). Recall that in DG's paper, the non-empty networks formed by D at equilibrium are either a star network with a protected central node, or a minimal $(k+1, n)$ -node-connected network without any protection. Hence, (s^D, \emptyset) , with $s^D \in S_{0,k}^D$, is an SPE when the cost of links (non-protected links in our case) is sufficiently low relative to the cost of protection in both frameworks.

However, when A attacks nodes, D uses at most one protected node at equilibrium. The role played by the protections is different since in DG, if D builds a star network, one protection is sufficient to protect the network and resist any attack of A . By contrast, in our framework, D may use more than one protected link in an SPE: indeed, when the cost of protected links is sufficiently low relative to the cost of non-protected links, D designs a $(1, n)$ -link-connected network which contains $n-1$ protected links. Protecting a network under link-attack is thus more costly than protecting a network under node-attack. This is because our framework calls for the survival of every node, a requirement which does not hold in the DG's framework.

4 Limited number of protected links and imperfectly protected links

In this section, we consider two potential types of restrictions on the protection of the network for the Designer. First, we consider the case where D can only use a limited number of protected links and we focus on situations where this number is smaller than $p_2(k, n)$. Then, we consider a framework where links are imperfectly protected and can be removed by the Adversary with some a priori known probability $\pi \in (0, 1)$.

4.1 Limited number of protected links

We examine a framework where the maximal number of protected links, \bar{p} , that D can form is strictly smaller than $n-1$. More precisely, we are interested in the case where $\bar{p} \in \llbracket p_1(k, n), p_2(k, n) \rrbracket$.¹⁸

Proposition 2 establishes that for $n(k+1) \bmod 2 = 0$, there exists no SPE in which D uses both protected and non-protected links. In contrast, when $\bar{p} \in \llbracket p_1, p_2 \rrbracket$, there exist values where the SPEs are

¹⁸This interval is non empty if $p_2 > p_1$ which implies that $\Delta > 0$, that is $n < \frac{(3k+5)^2}{8(k+1)}$.

of the form (s^D, \emptyset) with $s^D \in S_{p_1-1, k}^D$. Note that networks in $S_{p_1-1, k}^D$ contain both protected and non-protected links when $p_1 > 1$.¹⁹ The following proposition gives a condition on the values of the parameters upon which such situations arise. To simplify the analysis, we assume that $(k+1) \bmod 2 = 0$.

Proposition 3 *Assume that $(k+1) \bmod 2 = 0$ and $n < (3k+5)^2/(8(k+1))$. Assume further that $p_1 > 1$, $p_2 - p_1 \geq 2$ and that the maximal number of protected links is $\bar{p} \in \llbracket p_1, p_2 \rrbracket$. Let (s_*^D, s_*^A) be an SPE. There exists $\varepsilon > 0$ such that if $(k+1)/2 - \varepsilon < c_P/c_L < (k+1)/2$, then $s_*^D \in S_{p_1-1, k}^D$ and $s_*^A(s_*^D) = \emptyset$.*

Proof If $p_1 > 1$, for all $p \in \llbracket 1, p_1 - 1 \rrbracket$, $C^*(0, k) - C^*(p, k) = p(\frac{k+1}{2}c_L - c_P)$. If $c_P/c_L < (k+1)/2$, then $C^*(0, k) - C^*(p, k) > 0$. Therefore, if $c_P/c_L < (k+1)/2$, then $\arg \min_{p \in \llbracket 0, p_1 - 1 \rrbracket} \{C^*(p, k)\} = p_1 - 1$.

Now, let $p \in \llbracket p_1, \bar{p} \rrbracket$. We have $C^*(p, k) - C^*(p_1 - 1, k) = (p - p_1 + 1)c_P + (n_2(p, k) - n_1(p_1 - 1, k))c_L$. Consider $\varepsilon = \frac{k+1}{2} - \max_{p \in \llbracket p_1, \bar{p} \rrbracket} \left\{ \frac{n_1(p_1 - 1, k) - n_2(p, k)}{p - p_1 + 1} \right\}$. Since $\llbracket p_1 - 1, \bar{p} \rrbracket$ is a discrete non-empty set, then ε is well defined. If $(k+1)/2 - \varepsilon < c_P/c_L$, then $C^*(p, k) - C^*(p_1 - 1, k) > ((p - p_1 + 1)((k+1)/2 - \varepsilon) + n_2(p, k) - n_1(p_1 - 1, k))c_L \geq ((p - p_1 + 1)\frac{n_1(p_1 - 1, k) - n_2(p, k)}{p - p_1 + 1} + n_2(p, k) - n_1(p_1 - 1, k))c_L = 0$. Thus, $\arg \min_{p \in \llbracket p_1 - 1, \bar{p} \rrbracket} \{C^*(p, k)\} = p_1 - 1$.

It remains to show that $\varepsilon > 0$. If $p \in \llbracket p_1 - 1, \bar{p} \rrbracket$, then $(n - 2p)((k+1) - (n - 2p - 1)) > (k+1)p$. Moreover, by straightforward calculations if $(k+1) \bmod 2 = 1$ and $(n - 2p)((k+1) - (n - 2p - 1)) > (k+1)p$, then $n_2(p, k) > n_1(p, k)$. Thus

$$\frac{n_1(p_1 - 1, k) - n_2(p, k)}{p - p_1 + 1} < \frac{n_1(p_1 - 1, k) - n_1(p, k)}{p - p_1 + 1} \leq \frac{(k+1)(n - p_1 + 1 - n + p)}{2(p - p_1 + 1)} = \frac{k+1}{2}.$$

Since this holds for any $p \in \llbracket p_1, \bar{p} \rrbracket$, then $\varepsilon > 0$.

Finally, since (i) $s_*^D \in S_{p_1-1, k}^D$, (ii) any network in $S_{p_1-1, k}^D$ cannot be disconnected with k attacks and (iii) attacks are costly for A , then $s_*^A(s_*^D) = \emptyset$. \square

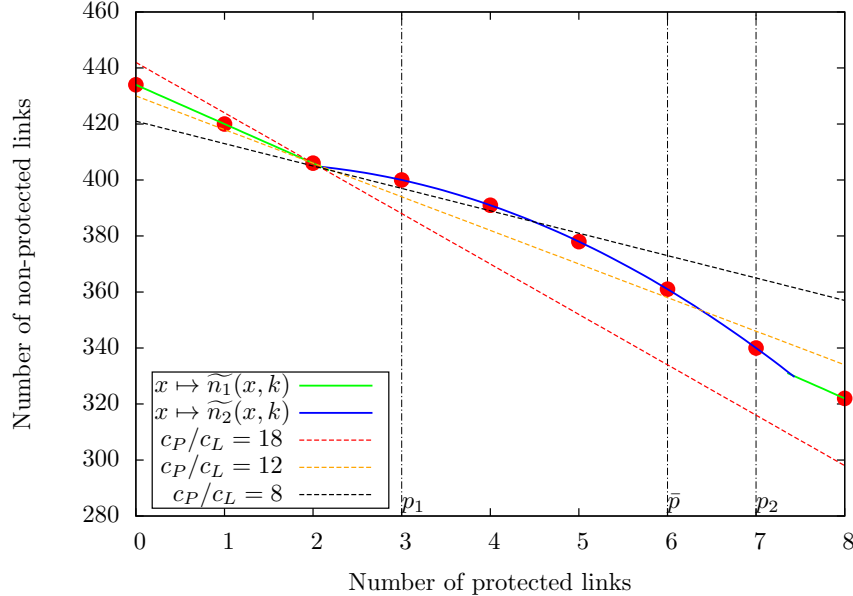


Figure 6: Example with $n = 31$, $k = 27$. The optimal strategy for D depends on the value of c_P/c_L .

¹⁹Indeed, note that $p_1 \leq \frac{4n}{8} + 1 \leq n - 1$ since $n \geq 4$ and therefore $S_{p_1-1, k}^D$ always contains non-protected links.

We illustrate this result with an example:

Example 3 Suppose that $n = 31$ and $k = 27$. Then, $p_1 = 3$ and $p_2 = 7$. Let $\bar{p} = 6$. Figure 6 shows the number of non-protected links in networks of $S_{p,k}^D$ as a function of the number of protected links (large dots in red).

Let us consider the following function:

$$\widetilde{E}_{NP} : x \mapsto \begin{cases} \widetilde{n}_1(x) = \frac{(n-x)(k+1)}{2} & \text{if } x \in [0, p_1] \cup (p_2, n-2], \\ \widetilde{n}_2(x) = (n-2x) \left(k+1 - \frac{n-2x-1}{2} \right) & \text{otherwise.} \end{cases}$$

Note that since $(k+1) \bmod 2 = 0$, then for any $p \in \llbracket 1, n-2 \rrbracket$ we have $\widetilde{n}_1(p) = n_1(p, k)$, $\widetilde{n}_2(p) = n_2(p, k)$ and $\widetilde{E}_{NP}(p) = \#E_{NP}(s^D)$ for any $s^D \in S_{p,k}^D$. Therefore the functions \widetilde{n}_1 , \widetilde{n}_2 and \widetilde{E}_{NP} can be interpreted as the natural continuous extensions of $p \mapsto n_1(p, k)$, $p \mapsto n_2(p, k)$ and $p \mapsto \#E_{NP}(s^D)$, $s^D \in S_{p,k}^D$, respectively. The functions \widetilde{n}_1 and \widetilde{n}_2 are plotted in green and blue plain lines respectively in Figure 6.

Let us consider a value of c_P and c_L . The lines of equal costs (isocost lines) have slope $-c_P/c_L$, thus have equations of the type

$$y = c c_L - p c_P, \quad (\text{Iso})$$

with $c c_L$ being the y -coordinate of the y -intercept. The value of c corresponds to the associated cost for the Designer normalized by c_L .

The optimal cost for D corresponds to the smallest value c such that the associated element of (Iso) intersects with the plot of \widetilde{E}_{NP} for some integer value p . Then, the optimal strategies of D belong to $S_{p,k}^D$.

Figure 6 shows the optimal strategies of D for three different values of c_P/c_L , namely 18, 12 and 8 (in dashed lines):

For large values of c_P/c_L (value 18 in Figure 6) the slope of the line of (Iso) is larger (in absolute value) than that of \widetilde{n}_1 and thus the optimal strategy for D is obtained with $p = 0$ protected links. In other words, if $c_P/c_L > \frac{k+1}{2}$, then the optimal strategies for D belong to $S_{0,k}^D$.

For small values of c_P/c_L (value 8 in Figure 6) the slope of the line (Iso) is low, hence favoring strategies with maximal values of p . In Figure 6, one can see that for $c_P/c_L = 8$, the optimal strategy for the Designer is obtained for $p = \bar{p}$ protected links.

For intermediate values of c_P/c_L (value 12 in Figure 6) the optimal strategy for the Designer is obtained when using $p_1 - 1$ protected links, which is the inflection point of \widetilde{E}_{NP} .

DG [9] show that when A attacks nodes, there exist situations where the optimal strategy of D is a star network with a protected central node. In this case, D uses both node protections and link creations to protect her network. In our framework, D may use both protected and non-protected links to protect her network if the number of protected links available to D belongs to $\llbracket p_1, p_2 \rrbracket$. This result is a consequence of the discontinuity in the number of non-protected links that each protected link allows the Designer to save (given that D builds a network that resists k attacks).

4.2 Imperfectly protected links

We now assume that each protected link has a probability $\pi \in (0, 1)$ to be removed when it is attacked by A . Let $g = (N, E_P, E_{NP})$ be an (imperfectly) protected network, and E^A an attack over the links of g . In the benchmark model, g^R is obtained by removing the links of E_{NP} that are targeted by A , i.e. $g^R = (N, E_P, E_{NP} \setminus E^A)$. Now, a realization of the attack, g^ϱ , is a subnetwork of g^R of the form $g^\varrho = (N, E_\varrho, E_{NP} \setminus E^A)$ with $E_\varrho \subseteq E_P \setminus E^A$. We illustrate these networks in the following example.

Example 4 Suppose $N = \llbracket 1, 5 \rrbracket$, $E_{NP}(g) = \{13, 15, 25, 34, 35, 45\}$, $E_P(g) = \{12, 24\}$ and $E^A = \{12, 34\}$ (see Figure 4). The subnetwork g^R obtained when removing the non-protected links that are attacked (i.e. $E^A \cap E_{NP}$), is drawn in Figure 7(ii). The two possible realizations are drawn in Figures 7(iii) and 7(iv). Note that g_1^ϱ occurs with probability $1 - \pi$, and g_2^ϱ occurs with probability π .

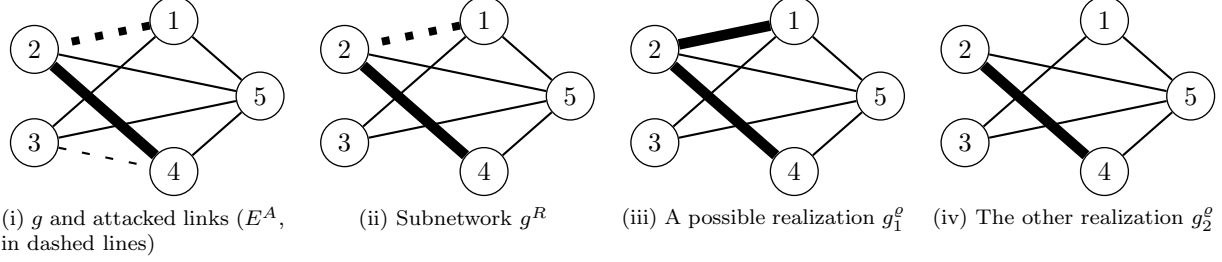


Figure 7: Networks of Example 4. Thick lines represent protected links and dashed lines represent links that are targeted by A .

Let g^e be a realization and $\lambda(g^e|g^R, E^A)$ be the probability that g^e is realized given g^R and E^A . We have

$$\lambda(g^e|g^R, E^A) = \prod_{\substack{ij \in E_P(g^e), \\ ij \in E_P^D \cap E^A}} (1 - \pi) \prod_{\substack{ij \notin E_P(g^e), \\ ij \in E_P^D \cap E^A}} \pi.$$

The expected benefits obtained by D , $\mathbb{E}\phi(g^R, E^A)$, when she builds a protected network g and A chooses E^A is

$$\mathbb{E}\phi(g^R, E^A) = \sum_{\substack{g^e = (N, E_\theta, E_{NP}^D \setminus E^A) \\ E_\theta \subseteq E_P^D}} \lambda(g^e|g^R, E^A) \phi(g^e).$$

We assume that the costs incurred by D when she chooses a strategy are given by equation (3), the costs incurred by A when she chooses a strategy are given by equation (2). The expected payoffs obtained by D , $\mathbb{E}\Pi^D$, is the difference between the expected benefits and the costs of forming protected and non-protected links:

$$\mathbb{E}\Pi^D(s^D, s^A(s^D)) = \mathbb{E}\phi(g^R, E^A) - c^D(s^D). \quad (8)$$

Finally, the expected payoffs obtained by A are

$$\mathbb{E}\Pi^A(s^D, s^A(s^D)) = 1 - \mathbb{E}\phi(g^R, E^A) - c^A(s^A(s^D)). \quad (9)$$

Recall that $c_A \geq 1/(n-3)$.

Proposition 4 *Let the payoff functions be given by equations (8) and (9). Suppose that $\pi < c_A$. Then, results provided in Proposition 2 are preserved.*

Proof Let $s_*^D = (N, E_P, E_{NP})$ and let $s_*^A(s_*^D) = E^A$ with $E^A = (E_P^A, E_{NP}^A)$, where $E_P^A \subseteq E_P$ and $E_{NP}^A \subseteq E_{NP}$.

If A can disconnect the protected network s_*^D with $E^A = (\emptyset, E_{NP}^A)$, then her best response is to not attack any protected links, i.e. $E_P^A = \emptyset$, since attacks are costly.

If A cannot disconnect the protected network s_*^D with (\emptyset, E_{NP}^A) , then (E_P^A, E_{NP}^A) should disconnect the network with a strictly positive probability (otherwise A would not be playing a best response). The highest probability to disconnect network (N, E_P, E_{NP}) occurs when the deletion of any protected link implies that (N, E_P, E_{NP}) is disconnected. This probability is $1 - (1 - \pi)^{\#E_P^A}$. Since $1 - \pi \in (0, 1)$, by Taylor's expansion $1 - (1 - \pi)^{\#E_P^A} = \sum_{\ell=0}^{\infty} \binom{\#E_P^A}{\ell} (-1)^{\ell+1} \pi^\ell$. Then by Leibniz's rule on alternating series $1 - (1 - \pi)^{\#E_P^A} \leq \pi \#E_P^A$. Hence, A disconnects the network with a probability lower or equal to $\pi \#E_P^A$ when she attacks the links in E^A . So, the expected benefits associated with E^A are at most $\pi \#E_P^A$ with a cost of at least $c_A \#E_P^A$. Hence, if $E_P^A \neq \emptyset$ and $\pi < c_A$, then the expected payoff associated to E^A is (strictly) negative and thus A does not play a best response. Therefore $E_P^A = \emptyset$.

Since no optimal strategy of A targets any protected link, the situation is equivalent to the one examined in Proposition 2. \square

Proposition 4 examines situations where probability π is low relative to c_A , the cost of attacking links. We now examine other situations through an example.

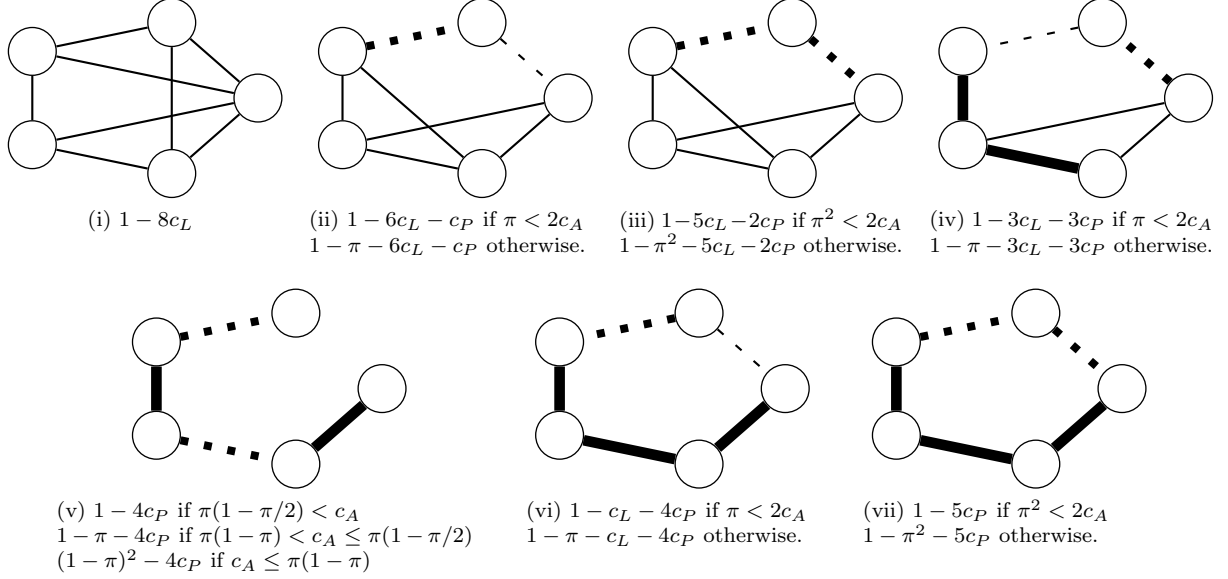


Figure 8: Networks of Example 5: $n = 5$, $k = 2$. Optimal strategies for D for different numbers of protected links and when A attacks (up to) 2 links. The captions of the figures give the Designer's expected benefit.

Example 5 Suppose that $N = \llbracket 1, 5 \rrbracket$ and $k = 2$, thus the maximum number of attacks that A has an incentive to do is 2. To simplify the analysis, we assume that $c_P > (3/2)c_L$. Figure 8 shows the networks that maximize the expected payoff of D for different numbers of protected links $p = 1, \dots, 5$ and all of A potential best response attacks. As in the previous figures, the protected links are represented by thick lines and dashed lines identify the links potentially attacked by A . The captions represent the Designer's expected benefit. Depending on the value of c_A , the Adversary may attack 0, 1 or 2 links leading to different values of D 's expected benefit.

Note that the (3,5)-Harary-network (in $S_{0,k}^D$ represented in Figure 8(i)) contains 0 protected links and 8 non-protected links and is fully protected against any attack of 2 links. Therefore, all other strategies that can be optimal for D induce networks which contain at most 7 links, which in turn implies that there exists at least one node of degree lower or equal to 2. Further, the expected benefit associated to any network for the Designer is always bounded by the probability of its weakest node to be disconnected. Hence, the expected benefits of any protected network not in $S_{0,k}^D$ are lower or equal to $1 - \pi^2$. Figure 8(vii) shows a network with such expected benefits and 5 protected links. Since $c_P > c_L$, D has no incentive to form networks which contains 6 protected links or more. Based on these observations, Figure 8 contains all potential optimal strategies of D under imperfectly protected links.

Let us now focus on the optimal strategies of A . Networks 8(ii)-8(iv) and 8(vi)-8(vii) can only be disconnected if at least two links fail. Therefore the Adversary has no incentive to attack only 1 link in these networks. Further, since $c_P > (3/2)c_L$, we have $3c_L + 3c_P > 6c_L + c_P$, and thus the strategy depicted in 8(iv) is never an optimal strategy for D . Further, suppose that for some value of π , the strategy of D depicted in Figure 8(iii) is optimal for D , then it has a greater or equal expected payoff than that of Figure 8(vii) and thus $5c_L \leq 3c_P$. But, then its expected payoff is no more than $1 - 5c_L - 10c_L/3 < 1 - 8c_L$ and thus is strictly dominated by strategy depicted in Figure 8(i) which is a contradiction. Thus, building the network of Figure 8(iii) is never an optimal strategy. It follows that only strategies depicted in Figures 8(i), 8(ii), 8(v), 8(vi) 8(vii) are candidate to be optimal for D .

Let us now compare these strategies with that of the case of perfectly protected links. From Proposition 2, since $n(k+1) \bmod 2 = 1$, three potential SPEs could occur, resulting in the networks of Figure 8(i), 8(ii) and 8(v). However, as in this case $k = n - 3$, then protected networks in $S_{1,k}^D$ (depicted in Figure 8(ii)) are never optimal when $\pi = 0$ (as explained in Footnote 17). For small values of π , the networks of Figure 8(i) and 8(v) can occur. For instance for $\pi = c_A = 0.1$ and $c_P = 0.2$ and $c_L = 0.075$, then the networks of $S_{0,k}^D$ are optimal, while for $c_P = 0.12$ and $c_L = 0.075$ the networks of $S_{n-1,k}^D$ are optimal.

Note that in that cases, A does not attack any link in SPEs.

Suppose that $c_P/c_L > (k+2)/2$, then for perfectly protected links (i.e. $\pi = 0$), building the network of Figure 8(i) is an equilibrium strategy for D and an equilibrium strategy for A is to attack no link. Observe now that the expected payoff of D associated with this network is not modified when π changes, while her expected payoff associated with all other networks drawn in Figure 8(v) decreases with π and reaches negative values. Consequently, given c_P and c_L there exists a probability $\bar{\pi}$ such that for $\pi > \bar{\pi}$ strategies in $S_{0,k}^D$ are optimal for D .

Moreover, for $\pi = 0.45$, $c_P = 0.113$, and $c_L = 0.075$ and $c_A = 0.2$, the network of Figure 8(vii) is induced by an optimal strategy for D given that A chooses an optimal attack. Further, for $c_P = 0.12$, $c_L = 0.075$, $\pi = 0.3$ and $c_A = 0.2$, the network of Figure 8(vi) is induced by an optimal strategy for D given that A plays an optimal strategy. Note that D never builds these protected networks in our benchmark model.²⁰

Example 5 establishes three main insights. First, $S_{0,k}^D$ are the unique optimal strategies when π is sufficiently high. Second, there exist situations where the Designer's best strategy is to build a network where each node is incident to m protected links, with $m = k$. Note that since $c_P > c_L$, D has no incentive to form a protected network where each node is incident to $k+1$ protected links. Third, there exist optimal strategies for D where nodes which belong to the same component in (N, E_P, \emptyset) are linked with a non-protected link (see Figure 8(vi)) in an optimal strategy for D .

DG [9] examine the impact of imperfect defense in a framework where D protects nodes instead of links. They use an example and provide two insights. First, there exist parameters such that the SPEs obtained in the perfect defense model remain equilibria in the imperfect defense model, namely the empty network, the center protected star, and the minimal $(k+1, n)$ -node-connected networks. Second, they establish that richer strategies than those played by D in the perfect defense model may appear in equilibrium. In particular, for some parameters an optimal strategy for D is to protect several nodes and create a network which generalizes the center protected star network, or to design a $(2, n)$ -node-connected network and to protect all the nodes.

It is worth noting that imperfect defense has the same type of impact in the framework of DG and in our framework. First, if the probability of successful attacks π is sufficiently high and the cost of forming non-protected links is sufficiently low, then strategies in $S_{0,k}^D$ are the unique optimal strategies for D . Second, the set of strategies candidate to be an equilibrium is larger in the imperfect defense framework than in the perfect defense framework. In particular, for sufficiently high π , D has an incentive to use more protections than in a situation where $\pi = 0$: there exist parameters where D protects all the nodes in DG's framework, and there exist parameters where D designs a network where each node is incident to k protected links in our framework. Third, in both frameworks it is difficult to obtain general results when imperfect defense is introduced. However, Landwehr [21] provides equilibrium strategies for D when the number of attacks is very small. In particular, he establishes that if $k = 2$, then there exist 6 types of strategies that D may play in equilibrium according to the value of π , c_P , and c_L .

5 Conclusion

In this paper, we have studied the optimal way to design and protect a network under strategic link attacks. In our benchmark model, the number of protected links available for the Designer is not bounded, and protected links cannot be removed by the Adversary. Our main findings in this model are the following. In equilibrium, three types of networks may arise according to the value of the parameters of the model (which are the number of nodes and the costs of link creation and attack). First, if the relative cost of protections (cost of a protected link/cost of a non-protected link) is low comparatively to the number of attacks, then D forms a $(1, n)$ -link-connected network which contains only protected links. Second, if the relative cost of protection is high in regards to the number of attacks, then the Designer forms a minimal $(k+1, n)$ -connected network which contains only non-protected links. Third, for intermediate relative costs of protection, there exist situations where the Designer forms a network which contains one protected link and $\lceil (n-1)(k+1)/2 \rceil$ non-protected links. To sum up, in this paper

²⁰Note that in Figure 8(vii), each node is incident to 2 protected links and $k = 2$. Interestingly, DG [9] and Landwehr [21] establish that in models with imperfect defense, there exist parameters where D designs a $(2, n)$ -Harary-networks in equilibrium.

we provide the minimal costs that D incurs to protect her network against an intelligent attack (i.e. the worst attack).

We have also examined a framework where the number of protected links available for the Designer is limited. In that case, we have established that for intermediate relative costs, the Designer forms a network which contains several protected and non-protected links. Finally, we have discussed the case of imperfectly protected links. We cannot provide a full characterization of the SPEs in the imperfect defense model, but we have given conditions under which results obtained in the framework with perfect defense are preserved. Moreover, we have established through an example that the set of equilibria is larger in the framework with imperfect defense links than in the framework with perfect defense.

In this paper, we have assumed that the Designer incurs the same costs if she forms protected links that are adjacent and if they are not adjacent. It would be interesting to examine a situation where it is more costly for the Designer to form protected links that are not adjacent. As we explained after Example 2, if D protects adjacent links, it can lead her to form strictly more non-protected links than in the optimal strategies described in Proposition 1.

Adding constraints on the location of protected links can be applied in different contexts. Indeed, it is more costly for a company to protect some cables (by reinforcing them or replacing them with new equipments) in different locations. For instance, the company has to send several teams of workers to protect cables which are far from each other instead of a single team when they are close to each other.

Appendix I: Proof of Proposition 1

The proof of Proposition 1 is organized in the following way. We first establish that the subnetwork (N, E_P, \emptyset) of $s^D \in S_{p,k}^D$ is acyclic (Lemma 1). In Appendix A, first given a number of nodes and a number of components, we provide an alternative optimization problem whose optimum corresponds to the optimal Designer's strategy (Lemma 2). Second, we provide a lower bound on the number of non-protected links in a protected network $s^D \in S_{p,k}^D$ (Lemma 3). In Appendices B and C, we provide the solutions of the optimization problem, both in terms of value (i.e. the minimum number of non-protected links) as well as a constructive method for the Designer to obtain an optimal set of non-protected links according to the number of protected links and the number of attacks.

The following lemma will allow us to establish that if D forms p protected links in $s^D \in S_{p,k}^D$, then s^D contains $n - p$ components.

Lemma 1 *Let $s^D = (N, E_P, E_{NP})$, $s^D \in S_{p,k}^D$. The subnetwork (N, E_P, \emptyset) of s^D is acyclic.*

Proof If $p = n - 1$, the result holds from Footnote 15. Otherwise, to introduce a contradiction, suppose that the subnetwork (N, E_P, \emptyset) of $s^D \in S_{p,k}^D$ contains a cycle. Then, there exists a link $ij \in E_P$ such that $(N, E_P \setminus \{ij\}, E_{NP})$ cannot be disconnected by an optimal attack of A . Moreover, since $p \in \llbracket 1, n - 2 \rrbracket$ and g^R cannot be disconnected by an optimal attack of A , we have $E_{NP} \neq \emptyset$. Let $i'j' \in E_{NP}$. Network $(N, E_P \setminus \{ij\} \cup \{i'j'\}, E_{NP} \setminus \{i'j'\})$ contains p protected links and $\#E_{NP} - 1$ non-protected links, a contradiction. \square

A An equivalent optimization formulation

The equivalent problem formulation relies on the concept of multigraph and contraction of networks which we now develop.

Multigraphs. A *multigraph* is a graph where multiple links and loops are allowed. Formally, an (undirected) multigraph \hat{g} is an ordered triplet $(\hat{N}, \hat{E}, \hat{\psi})$ consisting of a non-empty set of nodes, \hat{N} , a set of links, \hat{E} , disjoint with \hat{N} , and an incidence function $\hat{\psi} : \hat{E} \rightarrow \hat{N}^2$ that associates to each link an *unordered* pair of nodes of \hat{g} . If e is a link and i and j are nodes such that $\hat{\psi}(e) = (i, j)$, then e is said to join i and j .²¹

²¹By definition a simple graph does not contain a loop, that is a link joining a node to itself; neither does it contain multiple links, that is, several links joining the same two nodes. Therefore, it is a multigraph for which $\hat{\psi}$ is injective and for which there is no $e \in E$ such that $\hat{\psi}(e) = (i, i)$ with $i \in \hat{N}$.

We define the adjacency matrix $\mathcal{M}(\hat{g})$ of a multigraph $\hat{g} = (\hat{N}, \hat{E}, \hat{\psi})$ as $\forall (a, b) \in \hat{E}^2$, $\mathcal{M}_{a,b}(\hat{g}) = \#\{e \in \hat{E} : \hat{\psi}(e) = (a, b)\}$, i.e. the number of links between nodes a and b in \hat{g} . Note that the adjacency matrix of an undirected multigraph is symmetric. We note $|\cdot|$ the sum of elements of a matrix or a vector, that is, for any matrix $\mathcal{A} \in \llbracket 1, a \rrbracket \times \llbracket 1, b \rrbracket$, $|\mathcal{A}| = \sum_{i \in \llbracket 1, a \rrbracket, j \in \llbracket 1, b \rrbracket} \mathcal{A}_{i,j}$.

Contractions of networks. Let g be a network. Given a link $ij \in E(g)$, the network $g \circ ij$ is obtained by contracting the link ij ; that is, by merging the two nodes i and j into a single node $\{i, j\}$, and making any node a adjacent to the (new) node $\{i, j\}$ in $g \circ ij$ if and only if a is adjacent to i or j in the network g . In other words, all links, other than those incident to neither i nor j , are included in $E(g \circ ij)$ if and only if they are included in $E(g)$.

For any set $F \subseteq E(g)$, we define the F -contraction of network g and denote by \hat{g}^F the network obtained from g by sequences of link contractions for all links in F . Note that the resulting network does not depend on the order of links contractions. In particular, we are interested in the case where $F = E_P(g)$, that is the contraction over all protected links of the network. We illustrate the E_P -contraction of a network g in Figure 9.

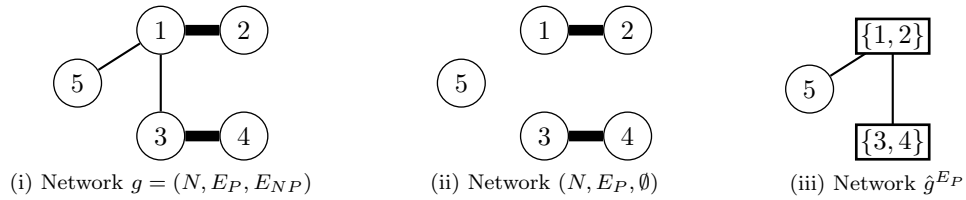


Figure 9: Illustration of the E_P -contraction: the links 12 and 34 are contracted.

Multigraphs and E_P -contractions of networks. Let $g = (N, E_P, E_{NP})$ be a protected network and $\Gamma_1(g), \dots, \Gamma_\ell(g), \dots, \Gamma_\nu(g)$ be the components of the subnetwork (N, E_P, \emptyset) with $\gamma_\ell(g)$ the number of nodes of the component $\Gamma_\ell(g)$. When no confusion is possible, we simplify notations by removing (g) . By construction, the E_P -contraction of g is $\hat{g}^{E_P} = (\llbracket 1, \nu \rrbracket, \hat{E}_g, \hat{\psi}_g)$ with $\hat{E}_g = \{e_{ij} : ij \in E_{NP}(g)\}$ and

$$\forall (a, b) \in \llbracket 1, \nu \rrbracket^2, \forall i \in \Gamma_a(g), \forall j \in \Gamma_b(g), \quad \hat{\psi}_g(e_{ij}) = (a, b).$$

Note that a protected network g induces one and only one E_P -contraction \hat{g}^{E_P} (up to ordering). However the converse is not true: a multigraph can be the E_P -contraction of two (or more) distinct protected networks. However, these graphs have the same number of non-protected links (which is given by $\sum_{a,b \in \llbracket 1, \nu \rrbracket^2} \mathcal{M}_{a,b}(\hat{g}^{E_P})$), and the same minimum number of protected links (which is equal²² to $n - \nu$). Therefore, all protected networks resulting in a given E_P -contraction have the same minimal cost.

An optimal strategy for the Designer is the choice of vector $(\Gamma_\ell(g))_{1 \leq \ell \leq \nu}$ and matrix $\mathcal{M}(g)$, that is the number and size of the components and the number of non-protected links, under some constraints, which we develop below:

Lemma 2 *For a given number of components ν , an optimal strategy for D is a solution of the following*

²²Since \hat{g}^{E_P} results of the contraction of p links, then $\nu \geq n - p$. The equality is attained when the subnetwork (N, E_P, \emptyset) contains no cycle.

optimization problem:

$$\begin{array}{l}
\min_{\substack{\gamma \in \mathbb{N}^\nu, \\ \mathcal{M} \in \mathbb{N}^{\nu \times \nu}}} \frac{1}{2} \sum_{i \in [1, \nu]} \sum_{j \in [1, \nu]} \mathcal{M}_{i,j} \quad s.t. \quad \left\{ \begin{array}{l}
\forall (i, j) \in [1, \nu]^2, \mathcal{M}_{i,j} = \mathcal{M}_{j,i}, \quad (\text{CS-1}) \\
\forall (i, j) \in [1, \nu]^2, \mathcal{M}_{i,j} \leq \gamma_i \gamma_j, \quad (\text{CS-2}) \\
\forall i \in [1, \nu], \mathcal{M}_{i,i} \leq \gamma_i(\gamma_i - 1)/2 - (\gamma_i - 1), \quad (\text{CS-3}) \\
\forall I \subseteq [1, \nu], \sum_{i \in I} \sum_{j \in [1, \nu] \setminus I} \mathcal{M}_{i,j} \geq k + 1, \quad (\text{CS-4}) \\
n = \sum_{i \in [1, \nu]} \gamma_i. \quad (\text{CS-5})
\end{array} \right.
\end{array}$$

Proof Consider any matrix $\mathcal{M} = (\mathcal{M}_{i,j})_{i \in [1, \nu], j \in [1, \nu]}$. Build $\hat{N} = [1, \nu]$, $\hat{E} = \cup_{(i,j) \in \hat{N} \times \hat{N}} \hat{E}_{ij}$, with $\hat{E}_{ij} = \{e_{ij}^1, \dots, e_{ij}^{\mathcal{M}_{i,j}}\}$, and $\hat{\psi}(e) = (i, j)$ if and only if $e \in \hat{E}_{ij}$. The triplet $(\hat{N}, \hat{E}, \hat{\psi})$ is an (undirected) multigraph if and only if the links are undirected (constraint (CS-1)).

In turns, this multigraph is the E_P -contraction of a protected network $g = (N, E_P, E_{NP})$, such that (N, E_P, \emptyset) has ν components of size $\gamma_1, \gamma_2, \dots, \gamma_\nu$, if and only if all nodes of g belong to exactly one component (constraint (CS-5)), if each node in one component is connected to any other node in a different component by at most one link (constraint (CS-2)), and if the network does not contain any loop. The latter requires that each node in a component Γ_i is linked with at most $\gamma_i - 1$ nodes which belong to Γ_i . From Lemma 1, each component of size γ_i contains exactly $\gamma_i - 1$ protected links (since it is connected and does not contain any loops). This is reflected in constraint (CS-3).

The goal of the Designer is to minimize her number of non-protected links, which are given by $\frac{1}{2} \sum_{i \in [1, \nu]} \sum_{j \in [1, \nu]} \mathcal{M}_{i,j}$.

Finally, no component of (N, E_P, \emptyset) should be vulnerable to an attack of A , that is, every component of (N, E_P, \emptyset) should be incident to at least $k + 1$ non-protected links in g : this means that $\forall i \in [1, \nu], \sum_{j \in [1, \nu] \setminus \{i\}} \mathcal{M}_{i,j} \geq k + 1$. This should also hold for any group of components, as reflected by constraint (CS-4). \square

This formulation allows us to directly derive a lower bound on the number of (non-protected) links that are necessary in the construction of a network that resists k attacks. We will show in the subsequent paragraphs that this bound can be reached under some assumptions on n, ν and k (Lemmas 5 and 9).

Lemma 3 *The number of non-protected links induced by any strategy is at least $\underline{n}_1 = \left\lceil \frac{\nu(k+1)}{2} \right\rceil$.*

Proof This is a direct consequence of constraint (CS-4). Indeed, for any i , eq. (CS-4) implies that $\sum_{j \in [1, \nu], j \neq i} \mathcal{M}_{i,j} \geq k + 1$. Therefore, $\sum_{i \in [1, \nu]} \sum_{j \in [1, \nu]} \mathcal{M}_{i,j} \geq \sum_{i \in [1, \nu]} (k + 1) = \nu(k + 1)$. \square

In the rest of the proof, we provide, for each number of components ν , the optimal value of the optimization problem given in Lemma 2 as well as the corresponding optimal vector γ and matrix \mathcal{M} . The constructions will rely heavily on the following definitions:

$$\alpha_1 = \left\lfloor \frac{n}{\nu} \right\rfloor, \quad \alpha_2 = n \bmod \nu, \quad \beta_1 = \left\lfloor \frac{k+1}{\nu-1} \right\rfloor, \quad \text{and } \beta_2 = (k+1) \bmod (\nu-1). \quad (11)$$

Note that by construction $n = \alpha_1 \nu + \alpha_2$ and $k + 1 = \beta_1 (\nu - 1) + \beta_2$. Also, one specific vector of component sizes, $\bar{\gamma} = (\bar{\gamma}_1, \dots, \bar{\gamma}_\nu)$, will come in handy in the proofs. Consider

$$\bar{\gamma}_i = \begin{cases} \alpha_1 & \text{if } i \leq \nu - \alpha_2, \text{ and} \\ \alpha_1 + 1 & \text{otherwise.} \end{cases} \quad (12)$$

Then $\sum_{i \in \llbracket 1, \nu \rrbracket} \bar{\gamma}_i = \alpha_1(\nu - \alpha_2) + (\alpha_1 + 1)\alpha_2 = \alpha_1\nu + \alpha_2 = n$, which satisfies constraint (CS-5). Roughly speaking, if we create ν almost equally sized components over n nodes, then there will be $\nu - \alpha_2$ components of size $\alpha_1 = \lfloor \frac{n}{\nu} \rfloor$ and α_2 components of size $\alpha_1 + 1$ (this is reflected in the definition of $\bar{\gamma}$). Consider one of the ν components. Recall that we have to satisfy (CS-4). If we distribute $k + 1$ links between this component and the $(\nu - 1)$ others in a way as balanced as possible, then $\nu - 1 - \beta_2$ components will be incident to $\beta_1 = \lfloor \frac{k+1}{\nu-1} \rfloor$ links and β_2 components will be incident to $\beta_1 + 1$ links to the considered component.

We distinguish two cases depending on the values of α_1 and β_1 , that lead to different constructions and optimal numbers of non-protected links. These two cases depend on the average size of the components relatively to the number of required links.

B Solution of the case where $\alpha_1^2 \geq \beta_1 + 1$.

Intuitively, in this case, the average size of the components is high enough compared to the number of required links. Thus, we are able to form a sufficient number of links incident to each component without contradicting constraint (CS-2). In this case, an optimal matrix \mathcal{M} that reaches the lower bound given by Lemma 3 can be built. It relies on the matrices of (a, b) -Harary-networks $\mathcal{H}(a, b)$ (for any $a < b$). In the following paragraph, we give the general formula of the adjacency matrices obtained by the construction of Harary [15]. Consider the squared matrices $\mathcal{D}(b)$ of size b which are defined by $\mathcal{D}(b)_{i,j} = 1$ if $(j - i) \bmod b = 1$ and 0 otherwise. Similarly, the squared matrices $\mathcal{E}(b)$ of size b which are defined by $\mathcal{E}(b)_{i,j} = 1$ if $j - i = 1$ and 0 otherwise.

That is,

$$\mathcal{D}(b) = \begin{pmatrix} 0 & 1 & & & \\ & 0 & 1 & & \\ & & & \ddots & \\ & 0 & & & 0 & 1 \\ 1 & & & & & 0 \end{pmatrix} \quad \text{and} \quad \mathcal{E}(b) = \begin{pmatrix} 0 & 1 & & & \\ & 0 & 1 & & \\ & & & \ddots & \\ & 0 & & & 0 & 1 \\ & & & & & 0 \end{pmatrix}.$$

Note that the powers of $\mathcal{D}(b)$ and $\mathcal{E}(b)$ have a specific structure. They satisfy, for any $a \in \mathbb{N}$, $\mathcal{D}(b)_{i,j}^{(a)} = 1$ if $(j - i) \bmod b = a$ and 0 otherwise and similarly $\mathcal{E}(b)_{i,j}^{(a)} = 1$ if $j - i = a$ and 0 otherwise.

Let us also define the matrix $\mathcal{F}(a, b) = \mathcal{D}(b)^{(a)} + (\mathcal{D}(b)^{(a)})^\top$ (with \cdot^\top the transpose operator). Note that if $0 < 2a < b$, then $\mathcal{F}(a, b)$ is an adjacency matrix of a network.²³ Finally, $\mathcal{G}(a, b) = \mathcal{E}(b)^{(a)} + (\mathcal{E}(b)^{(a)})^\top$ is also the adjacency matrix of a network for any $a, b > 0$.

Then, following the processes provided by Harary [15], when $a \in \llbracket 2, b-1 \rrbracket$, a matrix of a (a, b) -Harary-network is given by:

$$\mathcal{H}(a, b) = \begin{cases} \sum_{i=1}^{a/2} \mathcal{F}(i, b) & \text{if } a \bmod 2 = 0, \\ \sum_{i=1}^{\lfloor a/2 \rfloor} \mathcal{F}(i, b) + \mathcal{G}\left(\left\lfloor \frac{b}{2} \right\rfloor, b\right) & \text{if } a \bmod 2 = 1. \end{cases}$$

We also consider the situation where $a = 1$, and we assume that matrix $\mathcal{H}(1, b)$ is given by the previous formula, i.e. $\mathcal{H}(1, b) = \mathcal{G}\left(\left\lfloor \frac{b}{2} \right\rfloor, b\right)$. We have the following important lemma which is a very slight extension of Harary [15].

Lemma 4 *If $ab \bmod 2 = 0$ then each node of the network whose adjacency matrix is $\mathcal{H}(a, b)$ has a degree of a . Otherwise, all nodes have a degree of a except for the node $\lfloor b/2 \rfloor + 1$, that has a degree of $a + 1$. Consequently, for any a and b , we have*

$$|\mathcal{H}(a, b)| = \begin{cases} ab & \text{if } ab \bmod 2 = 0, \\ ab + 1 & \text{if } ab \bmod 2 = 1. \end{cases}$$

²³Indeed, \mathcal{F} is symmetric by construction and $\mathcal{F}(a, b)_{i,j} \in \{0, 1\}$. Suppose $\mathcal{F}(a, b)_{i,j} = 2$. Then, by construction $(j - i) \bmod b = a$ and $(i - j) \bmod b = a$. It follows that $((j - i) + (i - j)) \bmod b = 2a \bmod b$. Hence, $2a \bmod b = 0$, a contradiction since $0 < 2a < b$.

Proof For $a > 1$, the result follows from Harary [15]. For $a = 1$, the result follows from the construction of $\mathcal{G}(\lfloor \frac{b}{2} \rfloor, b)$: a link is formed between i and j if and only if $(j - i) = \lfloor \frac{b}{2} \rfloor$. \square

Note that $\mathcal{H}(b-1, b) = \begin{pmatrix} 0 & & & & \\ & 0 & & & \\ & & \ddots & & \\ & & & 0 & \\ 1 & & & & 0 \\ & & & & & 0 \end{pmatrix}$ is the adjacency matrix of a clique.

We can now describe the optimal strategy when $\alpha_1^2 \geq \beta_1 + 1$. An optimal strategy consists in overlapping cliques and a Harary-network.

Lemma 5 *Suppose that $\alpha_1^2 \geq \beta_1 + 1$. Then, any optimal strategy has exactly \underline{n}_1 non-protected links.*

Proof Consider vector $\bar{\gamma}$ as defined in eq. (12), so (CS-5) is satisfied. Intuitively, each node should be incident to $k+1$ links. If $k+1 \geq \nu-1$, multi-links are needed. Since $k+1 = \beta_1(\nu-1) + \beta_2$, we consider the matrix \mathcal{M} which is the sum of the adjacency matrices of β_1 cliques (i.e. $(\nu-1, \nu)$ -Harary-networks) and one (β_2, ν) -Harary-network. More precisely, let \mathcal{M} be given by:

$$\mathcal{M} = \beta_1 \mathcal{H}(\nu-1, \nu) + \mathcal{H}(\beta_2, \nu).$$

Let us first check that \mathcal{M} is an admissible matrix (i.e. that it satisfies the constraints (CS-1)-(CS-4)):

CS-1 is satisfied by construction.

CS-2 Further, $\forall i, j, \mathcal{M}_{i,j} \leq \beta_1 + 1 \leq \alpha_1^2 \leq \bar{\gamma}_i \bar{\gamma}_j$, which complies with constraint (CS-2).

CS-3 By construction Harary-networks do not contain loops. Hence $\forall i, \mathcal{M}_{i,i} = 0$ which satisfies constraint (CS-3).

CS-4 Finally, let $I \subseteq \llbracket 1, \nu \rrbracket$:

- If I is the singleton $\{i\}$, by definition of a Harary-network, $\sum_{j \in \llbracket 1, \nu \rrbracket \setminus \{i\}} \mathcal{M}_{i,j} = \beta_1(\nu-1) + \beta_2 = k+1$.
- Similarly, if $\#I = \nu-1$, then $\exists j \in \llbracket 1, \nu \rrbracket, I = \llbracket 1, \nu \rrbracket \setminus \{j\}$. Then $\sum_{i \in I} \sum_{j \in \llbracket 1, \nu \rrbracket \setminus I} \mathcal{M}_{i,j} = \sum_{i \in \llbracket 1, \nu \rrbracket, i \neq j} \mathcal{M}_{i,j} = k+1$.
- Finally, suppose that $2 \leq \#I \leq \nu-2$. Note that if $2 \leq \#I \leq \nu-2$, then $\nu > 3$. By construction, we have $\sum_{j \in \llbracket 1, \nu \rrbracket} \mathcal{M}_{1,j} = k+1$. We establish that I resists a number of attacks greater or equal to the number of attacks node 1 resists. By definition of an (a, ν) -Harary-network, for $a \in \llbracket 2, \nu-1 \rrbracket$ and $I \subset \llbracket 1, \nu \rrbracket, \#I \in \llbracket 2, \nu-2 \rrbracket$, we have $\sum_{i \in I} \sum_{j \in \llbracket 1, \nu \rrbracket \setminus I} \mathcal{H}(a, \nu)_{i,j} \geq a = \sum_{j \in \llbracket 1, \nu \rrbracket} \mathcal{H}(a, \nu)_{1,j}$ and for $\nu > 3$, $\sum_{i \in I} \sum_{j \in \llbracket 1, \nu \rrbracket \setminus I} \mathcal{H}(\nu-1, \nu)_{i,j} \geq \#I(\nu - \#I) \geq 2(\nu-2) > \nu-1 = \sum_{j \in \llbracket 1, \nu \rrbracket} \mathcal{H}(\nu-1, \nu)_{1,j}$. We now deal with $\mathcal{H}(a, \nu)$ for $a = 1$. We have for $\nu > 3$, $\sum_{i \in I} \sum_{j \in \llbracket 1, \nu \rrbracket \setminus I} \mathcal{H}(\nu-1, \nu)_{i,j} \geq \nu = \sum_{j \in \llbracket 1, \nu \rrbracket} \mathcal{H}(\nu-1, \nu)_{1,j} + \sum_{j \in \llbracket 1, \nu \rrbracket} \mathcal{H}(1, \nu)_{1,j}$. Note that since $k \geq 1$, if $\beta_2 = 1$, then $\beta_1 \geq 1$. Consequently, we have $\sum_{i \in I} \sum_{j \in \llbracket 1, \nu \rrbracket \setminus I} \mathcal{M}_{i,j} = \sum_{i \in I} \sum_{j \in \llbracket 1, \nu \rrbracket \setminus I} (\beta_1 \mathcal{H}(\nu-1, \nu)_{i,j} + \mathcal{H}(\beta_2, \nu)_{i,j}) \geq \sum_{j \in \llbracket 1, \nu \rrbracket} (\beta_1 \mathcal{H}(\nu-1, \nu)_{1,j} + \mathcal{H}(\beta_2, \nu)_{1,j}) = k+1$.

Hence constraint (CS-4) is satisfied.

Let us now compute the number of links induced by this strategy. We have:

$$|\mathcal{M}| = \begin{cases} \beta_1 \nu(\nu-1) + \beta_2 \nu & \text{if } \beta_2 \nu \bmod 2 = 0 \\ \beta_1 \nu(\nu-1) + \beta_2 \nu + 1 & \text{if } \beta_2 \nu \bmod 2 = 1 \end{cases} = \begin{cases} \nu(k+1) & \text{if } \beta_2 \nu \bmod 2 = 0 \\ \nu(k+1) + 1 & \text{if } \beta_2 \nu \bmod 2 = 1. \end{cases}$$

But then, $(k+1) \bmod 2 = (\beta_1(\nu-1) \bmod 2 + \beta_2 \bmod 2) \bmod 2$. Hence $\beta_2 \nu \bmod 2 = 1 \Rightarrow (k+1) \bmod 2 = 1 \Rightarrow \nu(k+1) \bmod 2 = 1 \Rightarrow \left\lceil \frac{\nu(k+1)+1}{2} \right\rceil = \left\lceil \frac{\nu(k+1)}{2} \right\rceil$. Therefore, in all cases, the number of links is \underline{n}_1 , which is optimal from Lemma 3. \square

Hence, when $\alpha_1^2 \geq \beta_1 + 1$, the lower bound on the number of links is attained. An important special case is when $2\nu \leq n$, that is when the number of protections is large. In that case, $\bar{\gamma}_i > 1$, and the components we consider in our construction have a size strictly higher than 1.

Lemma 6 *If $2\nu \leq n$, then $\alpha_1^2 \geq \beta_1 + 1$.*

Thus, if $2\nu \leq n$, then any optimal strategy has exactly \underline{n}_1 non-protected links.

Proof Indeed, $2\nu \leq n \Rightarrow \nu(2\nu - n) \leq 2(2\nu - n) \Rightarrow 2\nu^2 + 2n \leq n\nu + 4\nu$
 $\Rightarrow \nu n - 2\nu \leq 2n\nu - 2n - 2\nu^2 + 2\nu \Rightarrow \nu(n - 2) \leq 2(n - \nu)(\nu - 1)$
 $\Rightarrow \frac{n - 2}{\nu - 1} \leq \frac{2(n - \nu)}{\nu}$.

Hence $\frac{k+1}{\nu-1} \leq \frac{n-2}{\nu-1} \leq 2\frac{n}{\nu} - 2$. Thus, if $2\nu \leq n$, then $\lfloor \frac{n}{\nu} \rfloor \geq 2$ and hence $\lfloor \frac{n}{\nu} \rfloor^2 \geq 2\lfloor \frac{n}{\nu} \rfloor > 2(\frac{n}{\nu} - 1) \geq \frac{k+1}{\nu-1}$. Finally $\lfloor \frac{n}{\nu} \rfloor^2 \geq \lfloor \frac{k+1}{\nu-1} \rfloor + 1$. \square

C Solution of the case where $\alpha_1^2 < \beta_1 + 1$.

Intuitively, in this case, the average size of the components is low compared to the number of required links. However, small sized components cannot bear too many multiple links. Thus, the constraint (CS-2) may be harder to satisfy. More precisely, with the condition that $\alpha_1^2 < \beta_1 + 1$, the construction based on Harary networks of Lemma 5 is no longer valid (it violates constraint (CS-2)).

By Lemma 6, when $\alpha_1^2 < \beta_1 + 1$, we have $2\nu > n$ and thus $\alpha_1 = 1$ and $\alpha_2 = n - \nu$. To provide a lower bound on the minimal number of links required for the network to resist k attacks, we need to establish bounds concerning δ , the number of components of size 1 in a network at equilibrium. More precisely, we successively establish that $\delta \geq 2\nu - n$ and $\delta \leq k + 1$.

Let $\gamma = (\gamma_1, \dots, \gamma_\nu)$ be any (non-decreasing) vector satisfying constraint (CS-5). Then, $\gamma_1 = 1$, otherwise we would have $\sum_{i \in [1, \nu]} \gamma_i \geq 2\nu > n = \sum_{i \in [1, \nu]} \gamma_i$ which is impossible. Let δ be such that for all $i \leq \delta$, $\gamma_i = 1$ and for all $i > \delta$, $\gamma_i > 1$. Hence, δ is the maximal number such that $\gamma_\delta = 1$. Note that $n = \sum_{i \in [1, \nu]} \gamma_i = \delta + \sum_{i \in [\delta+1, \nu]} \gamma_i \geq \delta + 2(\nu - \delta)$. We have $\delta < \nu$ since $\nu < n$ as there exists at least one protected link. Hence, $\delta \geq 2\nu - n$. Finally, $\alpha_1^2 < \beta_1 + 1$ gives $\beta_1 \geq 1$ and thus $k + 1 \geq \nu - 1$ which implies $k + 1 \geq \delta$ by definition of δ as soon as $\nu < n$ (that is, there exists at least 1 protected link).

Up to a reordering of the nodes, any solution matrix \mathcal{M} has the following form:

$$\mathcal{M} = \left(\begin{array}{c|c} \overbrace{\mathcal{A}}^{\delta} & \overbrace{\mathcal{B}}^{\nu - \delta} \\ \hline \mathcal{B}^\top & \mathcal{C} \end{array} \right) \left. \begin{array}{l} \delta \\ \nu - \delta \end{array} \right\} \quad (13)$$

Note that adjacency matrix \mathcal{A} captures links between nodes in $\llbracket 1, \delta \rrbracket$, that is components of size 1. Adjacency matrix \mathcal{B} captures links between nodes in $\llbracket 1, \delta \rrbracket$ and nodes in $\llbracket \delta + 1, \nu \rrbracket$ (we will see below that nodes in $\llbracket \delta + 1, \nu \rrbracket$ are identified with components of size 2). Finally, adjacency matrix \mathcal{C} captures links between nodes in $\llbracket \delta + 1, \nu \rrbracket$.

In the next lemma, we establish that if $\alpha_1^2 < \beta_1 + 1$, then the number of non-protected links may be strictly higher than \underline{n}_1 .

Lemma 7 *If $\alpha_1^2 < \beta_1 + 1$ then the required number of non-protected links is at least:*

- $\underline{n}_2 = (2\nu - n)(k + 1) - \frac{(2\nu - n)(2\nu - n - 1)}{2}$ if $(3\nu - 2n)(k + 1) > (2\nu - n)(2\nu - n - 1)$.
- \underline{n}_1 otherwise.

Proof The bound \underline{n}_1 has been established in Lemma 3. Further, in the case of $\alpha_1^2 \geq \beta_1 + 1$, the optimal bound of \underline{n}_1 can be achieved (Lemma 5) because the elements of each row of \mathcal{M} add up to $k + 1$ with at most one row whose elements add up to $k + 2$.

From eq. (13), constraint (CS-2) imposes that each row of \mathcal{A} adds up to at most $\delta - 1$. Thus, to comply with constraint (CS-4), the elements of each row of \mathcal{B} need to add up to at least $k + 1 - (\delta - 1)$. Thus, the total sum of elements of matrix \mathcal{B} , that is, $|\mathcal{B}|$ is at least $\delta(k + 1 - (\delta - 1))$. Thus, the total number of links required is $\left\lceil \frac{|\mathcal{M}|}{2} \right\rceil \geq \left\lceil \frac{(|\mathcal{A}| + |\mathcal{B}|) + |\mathcal{B}^\top|}{2} \right\rceil \geq \left\lceil \frac{\delta(k + 1) + |\mathcal{B}|}{2} \right\rceil \geq \left\lceil \frac{\delta(k + 1) + \delta(k + 1 - (\delta - 1))}{2} \right\rceil = \delta(k + 1) - \frac{\delta(\delta - 1)}{2}$.

The function $\delta \mapsto \delta(k+1) - \frac{\delta(\delta-1)}{2}$ is concave, quadratic and its maximum is obtained at $k+3/2$; so this function is increasing in the interval $[2\nu-n, k+3/2]$. Since $2\nu-n \leq \delta \leq k+3/2$, then $\lceil \frac{|\mathcal{A}|}{2} \rceil \geq (2\nu-n)(k+1) - \frac{(2\nu-n)(2\nu-n-1)}{2} = \underline{n}_2$.

This bound is attained when the sum of elements of each row of \mathcal{B}^\top is larger than or equal to $k+1$ (hence leading to a zero matrix \mathcal{C}). This happens necessarily when $|\mathcal{B}| \geq (k+1)(\nu-\delta)$, that is $\delta(k+1 - (\delta-1)) \geq (k+1)(\nu-\delta)$, i.e. $(2\delta-\nu)(k+1) \geq \delta(\delta-1)$. Since $\delta \geq 2\nu-n$, this implies that $(3\nu-2n)(k+1) \geq (2\nu-n)(2\nu-n-1)$. We now establish that \underline{n}_1 occurs when $(3\nu-2n)(k+1) = (2\nu-n)(2\nu-n-1)$. Suppose $\nu(k+1) \bmod 2 = 1$ and $(3\nu-2n)(k+1) = (2\nu-n)(2\nu-n-1)$, by straightforward calculations we have $\underline{n}_2 < \underline{n}_1$, a contradiction since by definition $\underline{n}_2 \geq \underline{n}_1$. Moreover, if $\nu(k+1) \bmod 2 = 0$ and $(3\nu-2n)(k+1) = (2\nu-n)(2\nu-n-1)$, then by straightforward calculations we have $\underline{n}_2 = \underline{n}_1$. Finally, by straightforward calculations we have $\underline{n}_2 \geq \underline{n}_1$ when $(3\nu-2n)(k+1) > (2\nu-n)(2\nu-n-1)$. The result follows. \square

We now introduce the last two lemmas. In Lemma 8, we provide a construction that allows us to build exactly \underline{n}_2 non-protected links for resisting k attacks when $(2\nu-n)(2\nu-n-1) < (3\nu-2n)(k+1)$. Similarly, in Lemma 9 we provide a construction that allows us to build exactly \underline{n}_1 links for resisting k attacks when $(2\nu-n)(2\nu-n-1) \geq (3\nu-2n)(k+1)$. The proofs of these lemmas rely on the properties of a type of matrices which we denote as $\bar{\mathcal{Z}} \in \mathbb{N}^{\delta \times (\nu-\delta)}$, $\bar{\mathcal{Z}}_{i,j} \in \{0,1\}$, that are such that:

$$\forall i \in \llbracket 1, \delta \rrbracket, \quad \sum_{j \in \llbracket 1, \nu-\delta \rrbracket} \bar{\mathcal{Z}}_{i,j} = (k+1) - (\nu-1), \text{ and} \quad (14)$$

$$\forall d \in \llbracket 1, \delta \rrbracket, \quad \forall x, y \in \llbracket 1, \nu-\delta \rrbracket, \quad \left| \sum_{i \in \llbracket 1, d \rrbracket} \bar{\mathcal{Z}}_{i,x} - \sum_{i \in \llbracket 1, d \rrbracket} \bar{\mathcal{Z}}_{i,y} \right| \leq 1. \quad (15)$$

In other words, these matrices have δ rows and $\nu-\delta$ columns and elements in $\{0,1\}$. The sum of elements of each row is equal to $(k+1) - (\nu-1)$ (eq. 14) while the partial sums of the columns are balanced, i.e. their values differ by at most 1 (eq. 15). $\bar{\mathcal{Z}}$ is well defined as $(k+1) - (\nu-1) \geq 0$ (since $\beta_1 > 0$ from $\alpha_1^2 < \beta_1 + 1$) and $(k+1) - (\nu-1) \leq \nu-\delta$ (since $k+1 \leq n-1$). Also, note that the size of $\bar{\mathcal{Z}}$ complies with that of \mathcal{B} as defined in eq. (13). We have $|\bar{\mathcal{Z}}| = \delta((k+1) - (\nu-1))$. Matrix $\bar{\mathcal{Z}}$ captures the links formed by the following process. At step 0 there is no link between nodes in $\llbracket 1, \delta \rrbracket$ and nodes in $\llbracket \delta+1, \nu \rrbracket$. At each step $t = 1, \dots, \tau$, we consider the set of nodes in $\llbracket 1, \delta \rrbracket$ which are involved in the lowest number of links formed during the process (links formed at steps $1, \dots, t-1$), and we pick the node, say i_t , with the lowest index in this set. Similarly, we consider the subsets of $\llbracket \delta+1, \nu \rrbracket$ whose size is equal to $k+1 - (\nu-1)$ and such that the sum of links incident to the nodes of these subsets is minimal. We pick one of this subset and call it W_t (for instance the one such that the sum of indices of nodes is minimal). Then we form a link between i_t and each node in W_t . The process stops after each node in $\llbracket 1, \delta \rrbracket$ is involved in $k+1 - (\nu-1)$ links, i.e. $\tau = \delta$.²⁴

In Lemmas 8 and 9, $\bar{\mathcal{Z}}$ allow us to capture some of the links between nodes in $\llbracket 1, \delta \rrbracket$ and nodes in $\llbracket \delta+1, \nu \rrbracket$. Roughly speaking, these links allow us to ensure that nodes in $\llbracket 1, \delta \rrbracket$ have a degree equal to $k+1$ and these links are distributed in a way as balanced as possible between nodes in $\llbracket \delta+1, \nu \rrbracket$.

In the following lemma, we form links for satisfying the condition that each node in $\llbracket 1, \delta \rrbracket$ has a degree equal to $k+1$; these links are captured by adjacency matrices \mathcal{A} and \mathcal{B} . When these links are formed, nodes in $\llbracket \delta+1, \nu \rrbracket$ have degrees strictly higher than $k+1$. Hence we do not form any links between nodes in $\llbracket \delta+1, \nu \rrbracket$, and \mathcal{C} is the zero matrix. Moreover, the number of links required to ensure that each node has a degree at least equal to $k+1$ is strictly higher than \underline{n}_1 . We provide a construction that leads to a network that resists k attacks and contains \underline{n}_2 non-protected links.

Lemma 8 *If $\alpha_1^2 < \beta_1 + 1$ and $(2\nu-n)(2\nu-n-1) < (3\nu-2n)(k+1)$ then, optimal strategies require exactly \underline{n}_2 non-protected links.*

Proof We have already shown in Lemma 7 that \underline{n}_2 is a lower bound on the required number of links. We now show that this bound can be reached by providing a solution adjacency matrix.

Consider the following construction: let $\bar{\gamma}$ be as in eq. (12), that is, the nodes are ordered such that the $\delta = 2\nu-n$ components are of size 1 and the others are of size 2. Moreover, let (i) all components

²⁴We present a formal construction of $\bar{\mathcal{Z}}$ in Appendix II.

of size 1 be connected to all the other components (i.e. both the components of size 1 and 2), (ii) no component of size 2 be connected to any other component of size 2 and (iii) some components of size 2 have two links with a component of size 1. More precisely, we consider a solution matrix \mathcal{M} of the shape given by eq. (13) with:

$$\left\{ \begin{array}{l} \mathcal{A} = \mathcal{H}(\delta - 1, \delta), \\ \mathcal{B} = \underbrace{\left(\begin{array}{c} \mathbf{1} \\ \mathbf{1} \\ \vdots \\ \mathbf{1} \end{array} \right)}_{\nu - \delta} \delta + \bar{\mathcal{Z}}, \text{ and} \\ \mathcal{C} = \underbrace{\left(\begin{array}{c} \mathbf{0} \\ \mathbf{0} \\ \vdots \\ \mathbf{0} \end{array} \right)}_{\nu - \delta} \nu - \delta \quad (\text{i.e. } \mathcal{C} \text{ is the zero matrix}). \end{array} \right. \quad (16)$$

It is important to note that the assumption of Lemma 8 imposes that $3\nu - 2n > 0$. Hence, in this construction, $\delta > \nu - \delta$. Let us show that this construction satisfies the problem constraints altogether with the desired value of $\lceil |\mathcal{M}|/2 \rceil$.

CS-1 is verified, since by construction \mathcal{A} and \mathcal{C} are symmetric.

CS-2 $\forall i, j \in \llbracket 1, \delta \rrbracket$, $\mathcal{A}_{i,j} \leq 1 = \gamma_i \gamma_j$, $\forall i \in \llbracket 1, \delta \rrbracket, j \in \llbracket 1, \nu - \delta \rrbracket$ $\mathcal{B}_{i,j} \leq 2$ and $\forall i, j \in \llbracket 1, \nu - \delta \rrbracket$, $\mathcal{C}_{i,j} = 0 \leq 4$ and thus constraint (CS-2) is satisfied.

CS-3 is verified by construction since $\forall i, \mathcal{A}_{i,i} = \mathcal{C}_{i,i} = 0$.

CS-4 Let $I \subseteq \llbracket 1, \nu \rrbracket$.

- If I is a singleton $\{i\} \subseteq \llbracket 1, \delta \rrbracket$, then $\sum_{j \neq i} \mathcal{M}_{i,j} = (\delta - 1) + (\nu - \delta) + ((k + 1) - (\nu - 1)) = k + 1$.
- If I is a singleton $\{i\} \subseteq \llbracket \delta + 1, \nu \rrbracket$, then $\sum_{j \neq i} \mathcal{M}_{i,j} \geq \delta + \left\lfloor \frac{\delta((k+1) - (\nu - 1))}{\nu - \delta} \right\rfloor = (2\nu - n) + \left\lfloor \frac{(2\nu - n)((k+1) - (\nu - 1))}{n - \nu} \right\rfloor = \left\lfloor (2\nu - n) \frac{(k+1) - (2\nu - 1 - n)}{n - \nu} \right\rfloor > \left\lfloor \frac{(2\nu - n)(k+1) - (k+1)(3\nu - 2n)}{n - \nu} \right\rfloor$ by the lemma's assumption. Hence $\sum_{j \neq i} \mathcal{M}_{i,j} > \left\lfloor \frac{(-\nu + n)(k+1)}{n - \nu} \right\rfloor = k + 1$.
- Otherwise, note that since $\sum_{i \in I} \sum_{j \in \llbracket 1, \nu \rrbracket \setminus I} \mathcal{M}_{i,j} = \sum_{j \in \llbracket 1, \nu \rrbracket \setminus I} \sum_{i \in I} \mathcal{M}_{i,j}$, wlog we can suppose that $\#I \leq \lfloor \nu/2 \rfloor$. (Indeed, if $\#I \geq \lfloor \nu/2 \rfloor$, then consider $J = \llbracket 1, \nu \rrbracket \setminus I$. By construction, $\#J = \nu - \#I \leq \lfloor \nu/2 \rfloor$. Then $\sum_{i \in J} \sum_{j \in \llbracket 1, \nu \rrbracket \setminus J} \mathcal{M}_{i,j} = \sum_{i \in \llbracket 1, \nu \rrbracket \setminus I} \sum_{j \in I} \mathcal{M}_{i,j}$. Thus $\sum_{i \in J} \sum_{j \in \llbracket 1, \nu \rrbracket \setminus J} \mathcal{M}_{i,j} \geq k + 1 \Rightarrow \sum_{i \in I} \sum_{j \in \llbracket 1, \nu \rrbracket \setminus I} \mathcal{M}_{i,j} \geq k + 1$.) Further, since $k + 1 \geq \nu - 1$, then we can suppose that $\#I \leq \lfloor \frac{k}{2} \rfloor + 1$. Then, note that for all $i, j \in \llbracket 1, \nu \rrbracket$, we have $\mathcal{M}_{i,j} \leq 2$. Then, $\sum_{i \in I} \sum_{j \in \llbracket 1, \nu \rrbracket \setminus I} \mathcal{M}_{i,j} \geq \#I((k + 1) - (\#I - 1) \max_{i,j} \mathcal{M}_{i,j})$. Consider function $x \mapsto x((k + 3) - 2x)$. It is concave quadratic and attains its maximum at $(k + 3)/4$. Therefore, its minimum in $[1, \lfloor \frac{k}{2} \rfloor + 1]$ is $1(k + 3 - 2) = k + 1$. Hence $\sum_{i \in I} \sum_{j \in \llbracket 1, \nu \rrbracket \setminus I} \mathcal{M}_{i,j} \geq k + 1$.

Hence constraint (CS-4) is satisfied.

Finally, $|\mathcal{M}| = |\mathcal{A}| + 2|\mathcal{B}| + |\mathcal{C}| = (\delta - 1)\delta + 2\delta((\nu - \delta) + (k + 1) - (\nu - 1)) = (\delta - 1)\delta + 2\delta(1 - \delta + (k + 1)) = \delta(1 - \delta + 2(k + 1))$ which leads to the result by substituting $\delta = 2\nu - n$. \square

In the following lemma, first we form links for satisfying the condition that nodes in $\llbracket 1, \delta \rrbracket$ have a degree exactly equal to $k + 1$; these links are captured by adjacency matrices \mathcal{A} and \mathcal{B} . When these links are formed, nodes in $\llbracket \delta + 1, \nu \rrbracket$ have a degree strictly lower than $k + 1$. Hence we add links between nodes in $\llbracket \delta + 1, \nu \rrbracket$. These links are captured by adjacency matrix \mathcal{C} which is a non zero matrix. In the proof of Lemma 9, we describe a specific way to form links between nodes in $\llbracket \delta + 1, \nu \rrbracket$, and so a way to build adjacency matrix \mathcal{C} . This construction leads to a network that resists k attacks and contains \underline{n}_1 non-protected links.

Lemma 9 *If $\alpha_1^2 < \beta_1 + 1$ and $(2\nu - n)(2\nu - n - 1) \geq (3\nu - 2n)(k + 1)$, then optimal strategies have exactly \underline{n}_1 non-protected links.*

Proof Consider $\bar{\gamma}$ as in eq. (12). We construct a solution matrix \mathcal{M} of the shape of eq. (13) with $\delta = 2\nu - n$, and \mathcal{A} and \mathcal{B} defined as in eq. (16). Recall that adjacency matrix \mathcal{A} captures links between nodes in $\llbracket 1, \delta \rrbracket$, i.e. links between components of size 1, and adjacency matrix \mathcal{B} captures links between nodes in $\llbracket 1, \delta \rrbracket$ and nodes in $\llbracket \delta + 1, \nu \rrbracket$, i.e. links between components of size 1 and components of size 2. We now build matrix \mathcal{C} , which captures the links formed between nodes in $\llbracket \delta + 1, \nu \rrbracket$.

Recall that the matrix $\bar{\mathcal{Z}}$ has the property that the sum of elements between two columns should differ by at most 1 (from eq. (15)). Thus, by construction each column of matrix \mathcal{B} has a sum of either $\lfloor \frac{|\bar{\mathcal{Z}}|}{\nu - \delta} \rfloor + \delta$, or $\lfloor \frac{|\bar{\mathcal{Z}}|}{\nu - \delta} \rfloor + \delta + 1$. Let us denote by K the corresponding set of columns of the first kind:

$$K = \left\{ j, \sum_{i \in \llbracket 1, \delta \rrbracket} \mathcal{B}_{i,j} = \left\lfloor \frac{|\bar{\mathcal{Z}}|}{\nu - \delta} \right\rfloor + \delta \right\}.$$

Note that by construction of $\bar{\mathcal{Z}}$, we have $\#K = (\nu - \delta) - |\bar{\mathcal{Z}}| \bmod (\nu - \delta)$.

Thus, intuitively, so as to resist k attacks, we need to construct matrix \mathcal{C} in such a way that:

- \mathcal{C} is symmetric, all diagonal elements are either 0 or 1, all other elements are 0, 1 or 2,
- the sum of elements of each row in K should be (at least) equal to $k + 1 - \lfloor \frac{|\bar{\mathcal{Z}}|}{\nu - \delta} \rfloor - \delta$ while the sum of elements of each other row should be (at least) equal to $k - \lfloor \frac{|\bar{\mathcal{Z}}|}{\nu - \delta} \rfloor - \delta$.

For ease of notations, let us introduce

$$f = k - \left\lfloor \frac{|\bar{\mathcal{Z}}|}{\nu - \delta} \right\rfloor - \delta.$$

Thus \mathcal{C} is a matrix with minimal value $|\mathcal{C}|$ for which: (i) the sum of elements of each row in K is (at least) equal to $f + 1$ and (ii) the sum of elements of each row not in K is (at least) equal to f , so that each node $i \in \llbracket \delta + 1, \nu \rrbracket$ has a degree of $k + 1$.

We construct \mathcal{C} as the sum of 2 matrices, \mathcal{C}^1 and \mathcal{C}^2 , which we define below. So, we have $\mathcal{C} = \mathcal{C}^1 + \mathcal{C}^2$. In the first one, \mathcal{C}^1 , the sum of elements of each row associated with nodes in K is equal to 1 (except possibly for one node in K , for this node, this sum is equal to 0 or 2) while the sum of elements of other rows is equal to zero. In other words, adjacency matrix \mathcal{C}^1 captures the fact that we add a degree to each node in K (except possibly one). So if we restrict our attention to adjacency matrices \mathcal{B} and \mathcal{C}^1 each node in $\llbracket \delta + 1, \nu \rrbracket$ has the same degree, equal to $\lfloor \frac{|\bar{\mathcal{Z}}|}{\nu - \delta} \rfloor + \delta + 1$ (except possibly for one node which has a degree equal to $\lfloor \frac{|\bar{\mathcal{Z}}|}{\nu - \delta} \rfloor + \delta$).

In the second one, \mathcal{C}^2 each row adds up to f (except possibly one which adds up to $f + 1$). Let us now explain the role played by the adjacency matrix \mathcal{C}^2 . Due to the construction of adjacency matrices \mathcal{B} and \mathcal{C}^1 , we know that we have to add f degrees to each node in $\llbracket \delta + 1, \nu \rrbracket$ (except possibly for one). For these nodes, we use the same method as in Lemma 5: we overlap some cliques and a Harary network. Matrix \mathcal{C}^2 captures this building process.

Construction of \mathcal{C}^1 : Matrix \mathcal{C}^1 is a symmetric matrix which satisfies the following conditions

- (i) all elements are either 0 or 1: $\mathcal{C}_{i,j}^1 \in \{0, 1\}$;
- (ii) each row (resp. column) whose index is not in the set K contains only elements equal to zero: $\forall i, j, \mathcal{C}_{i,j}^1 = 1 \Rightarrow (i \in K \text{ and } j \in K)$;
- (iii) each row whose index is in K admits at most one non-zero element;
- (iv) there exists at most one column with two elements which are not equal to zero.

To construct \mathcal{C}^1 , let σ be an ordering of the elements of K , that is $K = \{\sigma_1, \sigma_2, \dots, \sigma_{\#K}\}$, with $\sigma_1 < \sigma_2 < \dots < \sigma_{\#K}$. Let $\mathcal{J}(i, j, b)$ be the squared matrix of size b whose elements are all zero except

for the one at row i and column j and its symmetric element (at row j and column i) whose value is 1. We set:

$$\mathcal{C}^1 = \sum_{i \in \llbracket 0, \lfloor \frac{\#K}{2} \rrbracket - 1 \rrbracket} \mathcal{J}(\sigma_{2i+1}, \sigma_{2i+2}, \nu - \delta) + \begin{cases} 0 & \text{if } \#K \bmod 2 = 0 \text{ or } (\nu - \delta)f \bmod 2 = 1, \\ \mathcal{J}(\sigma_{\#K}, \sigma_1, \nu - \delta) & \text{otherwise.} \end{cases}$$

\mathcal{C}^1 is symmetric as a sum of symmetric matrices. Note that in the case where $\#K \bmod 2 = 1$ and $f(\nu - \delta) \bmod 2 = 0$, the column of index σ_1 adds up to 2: node σ_1 is incident to 2 links while other nodes in K are incident to 1 link.

Let us briefly comment on the special status of $\mathcal{J}(\sigma_{\#K}, \sigma_1, \nu - \delta)$. In the case where $\#K \bmod 2 = 1$, there exists no network such that each node in K has a degree of 1 while each node not in K has a degree of 0 (since in any network the sum of degrees is always even). In other words, there exists no adjacency matrix such that all rows (resp. columns) in the set K add up to 1 and all rows (resp. columns) out of the set K add up to 0. Therefore, to ensure that each row (resp. column) in K adds up to 1, an additional degree is required. However, in the case where $(\nu - \delta)f \bmod 2 = 1$, we will see below that the matrix \mathcal{C}^2 , whose construction is based on Harary networks contains exactly one node whose degree is $f + 1$. Therefore, in that case, we construct \mathcal{C}^2 in such a way that this extra link is incident to node $\sigma_{\#K}$.

Construction of \mathcal{C}^2 : Matrix \mathcal{C}^2 can be constructed in an analogous way as \mathcal{M} in Lemma 5. Indeed, \mathcal{C}^2 is analogous to the adjacency matrix of a multigraph in which each node (among $\nu - \delta$ nodes) should be incident to f links. Since it may be that $f \geq \nu - \delta - 1$, then multiple links may be required. Therefore, we introduce f_1 and f_2 such that $f = (\nu - \delta - 1)f_1 + f_2$ and $f_2 < \nu - \delta - 1$, i.e.

$$f_1 = \lfloor f / (\nu - \delta - 1) \rfloor, \quad \text{and} \quad f_2 = f \bmod (\nu - \delta - 1).$$

Let us give some intuitions for f_1 and f_2 . Consider one component in $\llbracket \delta + 1, \nu \rrbracket$. If we distribute f links between this component and the $(\nu - \delta - 1)$ others in a way as balanced as possible, then $(\nu - \delta - 1) - f_2$ components will be incident to f_1 links and f_2 components will be incident to $f_1 + 1$ links to the considered component. Consider now the adjacency matrix, \mathcal{X} , which is the sum of the adjacency matrices of f_1 cliques (i.e. $(\nu - \delta - 1, \nu - \delta)$ -Harary networks) and one $(f_2, \nu - \delta)$ -Harary network, that is:

$$\mathcal{X} = f_1 \mathcal{H}(\nu - \delta - 1, \nu - \delta) + \mathcal{H}(f_2, \nu - \delta).$$

Recall that in the case where both $\#K \bmod 2 = 1$ and $(\nu - \delta)f \bmod 2 = 0$, an extra link is required that is adjacent to node $\sigma_{\#K}$. Due to this link, σ_1 has a degree equal to 2 in \mathcal{C}^1 . In \mathcal{M} , σ_1 has a degree equal to $k + 2$. Note that $(\nu - \delta)f \bmod 2 = 1 \Leftrightarrow (\nu - \delta)f_2 \bmod 2 = 1$. Hence, in the case where $\#K \bmod 2 = 0$ and $(\nu - \delta)f \bmod 2 = 1$, in the network whose adjacency matrix is $\mathcal{H}(f_2, \nu - \delta)$, the node of index $\lfloor \frac{\nu - \delta}{2} \rfloor + 1$ has a degree equal to $f_2 + 1$ (from Lemma 4). In \mathcal{M} , this node has a degree equal to $k + 2$. Moreover, in the case where $\#K \bmod 2 = 1$ and $(\nu - \delta)f \bmod 2 = 0$, node $\sigma_{\#K}$ has a degree equal to 0 in \mathcal{C}^1 and in the network whose adjacency matrix is $\mathcal{H}(f_2, \nu - \delta)$, the node of index $\lfloor \frac{\nu - \delta}{2} \rfloor + 1$ has a degree equal to $f_2 + 1$. Therefore, when $\#K \bmod 2 = 1$, consider as matrix \mathcal{C}^2 the matrix obtained from \mathcal{X} by interchanging indices $\sigma_{\#K}$ and $\lfloor \frac{\nu - \delta}{2} \rfloor + 1$ when $\#K \bmod 2 = 1$.²⁵ Note that, by construction and the permutation of indices, if $\#K \bmod 2 = 1$ and $(\nu - \delta)f_2 \bmod 2 = 1$, then node $\sigma_{\#K}$ has the same degree as the other nodes.

Let us now show that this construction satisfies the optimization problem's constraints:

CS-1 is satisfied, since by construction \mathcal{A} and \mathcal{C} are both symmetric.

²⁵Let us define formally the required permutation:

$$\forall i, \tilde{\sigma}_i = \begin{cases} i & \text{if } i \neq \sigma_{\#K} \text{ and } i \neq \lfloor \frac{\nu - \delta}{2} \rfloor + 1 \\ \sigma_{\#K} & \text{if } i = \lfloor \frac{\nu - \delta}{2} \rfloor + 1 \\ \lfloor \frac{\nu - \delta}{2} \rfloor + 1 & \text{if } i = \sigma_{\#K} \end{cases} \quad \text{and } \forall i, j, \mathcal{C}_{i,j}^2 = \mathcal{X}_{\tilde{\sigma}_i, \tilde{\sigma}_j}$$

CS-2 Recall that $\nu - \delta = n - \nu$ and $2\nu - n = \delta$. From the choice of $\bar{\gamma}$, the constraint (CS-2) translates into: $\forall i, j \in \llbracket 1, 2\nu - n \rrbracket$, $\mathcal{A}_{i,j} \leq \gamma_i \gamma_j = 1$, $\forall i \in \llbracket 1, 2\nu - n \rrbracket, j \in \llbracket 2\nu - n + 1, \nu \rrbracket$, $\mathcal{B}_{i,j} \leq \gamma_i \gamma_j \leq 2$ and $\forall i, j \in \llbracket 2\nu - n + 1, \nu \rrbracket$, $\mathcal{C}_{i,j} \leq \gamma_i \gamma_j \leq 4$. The construction of \mathcal{M} induces that $\forall i, j$ $\mathcal{A}_{i,j} \leq 1$, $\mathcal{B}_{i,j} \leq 2$ and $\mathcal{C}_{i,j} \leq f_1 + 1 + 1$. Then,

$$\begin{aligned}
f_1 &\leq \frac{(k+1) - \frac{|\bar{\mathcal{Z}}|}{n-\nu} - (2\nu - n)}{n - \nu - 1} \\
&\leq \frac{(k+1)(n-\nu) - (2\nu - n)((k+1) - (\nu - 1)) - (2\nu - n)(n - \nu)}{(n - \nu - 1)(n - \nu)} \\
&= \frac{(k+1)(2n - 3\nu) + (2\nu - n)(2\nu - n - 1)}{(n - \nu - 1)(n - \nu)} \\
&\leq \frac{(n-2)(2n - 3\nu) + (2\nu - n)(2\nu - n - 1)}{(n - \nu - 1)(n - \nu)} = 4 - \frac{n}{n - \nu} < 3.
\end{aligned} \tag{17}$$

Since f_1 is an integer, then $f_1 \leq 2$ and thus $\mathcal{C}_{i,j} \leq 4$, hence, satisfying constraint (CS-2).

CS-3 \mathcal{A} and \mathcal{C}^2 have a zero diagonal as they are sums of matrices with zero diagonals. Since $\sigma_1 < \sigma_2 < \dots < \sigma_{\#K}$ then all \mathcal{J} matrices involved in the construction of \mathcal{C}^1 have zero diagonals. Therefore $\forall i, \mathcal{M}_{i,i} = 0$ and constraint (CS-3) is satisfied.

CS-4 Finally, let $I \subseteq \llbracket 1, \nu \rrbracket$.

- If I is a singleton, then by construction $\sum_{j \neq i} \mathcal{M}_{i,j} = k + 1$.
- Otherwise, note that from eq. 17 $\forall i, j, \mathcal{M}_{i,j} \leq 2$. Therefore, the proof is similar to that of Lemma 8: $\sum_{i \in I} \sum_{j \notin I} \mathcal{M}_{i,j} \geq \#I(k + 1 - (\#I - 1)(\max_{i,j} \mathcal{M}_{i,j})) \geq 1((k + 1) - 0) = k + 1$.

Hence, constraint (CS-4) is satisfied.

We now compute the number of links induced by \mathcal{M} . By construction, all nodes in $\llbracket 1, \delta \rrbracket$ have a degree equal to $k + 1$. Similarly, all nodes in $\llbracket \delta + 1, \nu \rrbracket$, except possibly one, say i , have a degree equal to $k + 1$. We now examine situations where node i has a degree strictly higher than $k + 1$. If $\#K \bmod 2 = 1$ and $(\nu - \delta)f \bmod 2 = 0$, then node i is identified with node σ_1 defined in \mathcal{C}^1 and has a degree equal to $k + 2$. Similarly, if $\#K \bmod 2 = 0$, i.e. all nodes in K have exactly one degree due to the construction of \mathcal{C}^1 , and if $f_2(\nu - \delta) \bmod 2 = 1$, then node i is identified with node $\lfloor \frac{\nu - \delta}{2} \rfloor + 1$. In that case, node i has a degree equal to $k + 2$. It remains to establish that the degree of i is equal to $k + 2$ if and only if $\nu(k + 1) \bmod 2 = 1$. We have to examine the different possibilities associated with $\nu(k + 1) \bmod 2 = 0$ or $\nu(k + 1) \bmod 2 = 1$, and $\delta \bmod 2 = 0$ or $\delta \bmod 2 = 1$. Suppose that $\nu(k + 1)\delta \bmod 2 = 1$. By construction of set K , we have $|\bar{\mathcal{Z}}| = x(\nu - \delta) + ((\nu - \delta) - \#K)$, with $x \in \mathbb{N}$, that is $\delta((k + 1) - (\delta - 1) - (\nu - \delta)) = x(\nu - \delta) + ((\nu - \delta) - \#K)$, with $x \in \mathbb{N}$. We have $\delta((k + 1) - (\delta - 1) - (\nu - \delta)) \bmod 2 = 1$, so $((\nu - \delta) - \#K) \bmod 2 = 1$ since $(\nu - \delta) \bmod 2 = 0$. Moreover, since $(\nu - \delta) \bmod 2 = 0$, we have $\#K \bmod 2 = 1$. Therefore, the degree of node i is equal to $k + 2$ since $\#K \bmod 2 = 1$ and $f(\nu - \delta) \bmod 2 = 0$. The examination of the different other possibilities associated with $\nu(k + 1) \bmod 2 = 0$ or $\nu(k + 1) \bmod 2 = 1$, and $\delta \bmod 2 = 0$ or $\delta \bmod 2 = 1$ is done by using the same arguments as in the case where $\nu(k + 1)\delta \bmod 2 = 1$. This fastidious examination allows us to conclude that the degree of node i is $k + 2$ if and only if $\nu(k + 1) \bmod 2 = 1$.

It follows that $\frac{|\mathcal{M}|}{2} = \left\lceil \frac{\nu(k+1)}{2} \right\rceil$, i.e. $\frac{|\mathcal{M}|}{2} = \underline{n}_1$. □

We now conclude the proof of Proposition 1. First, let us observe the conditions of Lemma 8. These are:

$$\left(\left\lfloor \frac{n}{\nu} \right\rfloor \right)^2 < \left\lfloor \frac{k+1}{\nu-1} \right\rfloor + 1 \tag{18a}$$

$$(2\nu - n)(2\nu - n - 1) < (3\nu - 2n)(k + 1) \tag{18b}$$

Suppose that eq. (18b) is satisfied and $n > \nu$. Then:

1. Consider the quadratic $x \mapsto (2\nu - x)(2\nu - x - 1)$. It is always non negative except in the (open) interval $(2\nu - 1, 2\nu)$. Since n is an integer, for any n and ν , we have $(2\nu - n)(2\nu - n - 1) \geq 0$.

Therefore, $3\nu - 2n > 0$. Thus $2\nu + \nu > 2n$, and hence $\nu > 2(n - \nu) > n - \nu$ (since $\nu < n$). Thus $2\nu > n > \nu$, or in other words $\lfloor \frac{n}{\nu} \rfloor = 1$.

2. Since $\nu > n - \nu$ and $n > \nu$, we have $\nu > 1$.

3. Now, suppose that $\nu - 1 > k + 1$. Then, $(2\nu - n)(2\nu - n - 1) < (3\nu - 2n)(k + 1) \Rightarrow (2\nu - n)(2\nu - n - 1) < (3\nu - 2n)(\nu - 1) \Rightarrow 4\nu^2 + n^2 - 4\nu n + n - 2\nu < 3\nu^2 - 2n\nu - 3\nu + 2n \Rightarrow \nu^2 + n^2 - 2\nu n + \nu < n \Rightarrow (n - \nu)^2 < n - \nu \Rightarrow n - \nu = 0$, which is impossible. Thus $\nu - 1 \leq k + 1$. Since $\nu > 1$, we have $\lfloor \frac{k+1}{\nu-1} \rfloor \geq 1$.

Now if $n = \nu$, then eq. (18b) implies that $n(n - 1) < n(k + 1)$. Since $n > 0$, this implies that $\nu - 1 < k + 1$, that is $\lfloor \frac{k+1}{\nu-1} \rfloor \geq 1$.

We have shown that for all $n \geq \nu$, any solution of eq. (18b) satisfies $(\lfloor \frac{n}{\nu} \rfloor)^2 = 1 < 1 + 1 \leq 1 + \lfloor \frac{k+1}{\nu-1} \rfloor$. Therefore, constraint (18a) is implied by constraint (18b) and can thus be omitted.

Second, we have $\sum_{\ell=1}^{\nu} \gamma_{\ell} = n$ and by Lemma 1, $\sum_{\ell=1}^{\nu} (\gamma_{\ell} - 1) = p$ since (N, E_P, \emptyset) is acyclic. It follows that $\nu = n - p$. Since $\nu = n - p$, $\lfloor \frac{\nu(k+1)}{2} \rfloor = \lfloor \frac{(n-p)(k+1)}{2} \rfloor$, so $n_1 = n_1(p, k)$. Similarly, $(2\nu - n)((k + 1) - \frac{2\nu - n - 1}{2})$ is equal to $(n - 2p)(k + 1 - \frac{n - 2p - 1}{2})$, and so $n_2 = n_2(p, k)$. Finally, note that equation $(n - 2p)((k + 1) - (n - 2p - 1)/2) = (k + 1)p$ is quadratic in p . Let x_1 and x_2 be the two real roots of the polynomial when they exist (which occurs when $n \leq \frac{(3k+5)^2}{8(k+1)}$). Then, $p_1(k, n) = \lfloor x_1 \rfloor + 1$ and $p_2(k, n) = \lceil x_2 \rceil - 1$.

Appendix II: A construction for Matrix $\overline{\mathcal{Z}}$

A possible construction for matrix $\overline{\mathcal{Z}}$ is to proceed according to the following process:

Input: Number of rows δ , columns $\nu - \delta$ and value of row sum $(k + 1) - (\nu - 1)$

Output: A possible matrix $\overline{\mathcal{Z}}$

1 Initialize matrix $\overline{\mathcal{Z}}$ to the zero matrix

2 Set $\overline{\mathcal{Z}}_{1,j} = 1$ for all $j \leq (k + 1) - (\nu - 1)$

3 **for** each row r from 2 to δ **do**

4 **for** each column q from 1 to $\nu - \delta$ **do**

5 Compute the partial sum $w_q = \sum_{i \in [1, r]} \overline{\mathcal{Z}}_{i, q}$

6 Select exactly $(k + 1) - (\nu - 1)$ columns among the $\nu - \delta$ columns having the lowest sum of w_q , i.e.

7 finds $W_r \in \arg \min_{C \subseteq [1, \nu - \delta]} \left\{ \sum_{q \in C} w_q, \#C = (k + 1) - (\nu - 1) \right\}$

8 Set these elements to one: $\overline{\mathcal{Z}}_{r, j} = 1$ with $j \in W_r$

References

- [1] D. Acemoglu, A. Malekian and A. Ozdaglar. Network security and contagion. *National Bureau of Economic Research*, 2013.
- [2] M. Baccara and H. Bar-Isaac. How to organize crime. *The Review of Economic Studies*, 75(4):1039-1067, 2008.
- [3] V. Bala and S. Goyal. A noncooperative model of network formation. *Econometrica*, 68(5):1181-1229, 2000.

- [4] V. Bala and S. Goyal. A strategic analysis of network reliability. *Review of Economic Design*, 5:205-228, 2000.
- [5] G. Brown, M. Carlyle, J. Salmerón, and K. Wood. Defending critical infrastructure. *Interfaces*, 36(6):530-544, 2006.
- [6] A. Cabrales, P. Gottardi, and F. Vega-Redondo. Risk-sharing and contagion in networks. *Working paper*, 2014, <https://ideas.repec.org/p/cte/werepe/we1301.html>.
- [7] D. Cerdeiro, M. Dziubiński, and S. Goyal. Contagion risk and network design. *Working paper*, 2015, <https://ideas.repec.org/p/fem/femwpa/2015.56.html>.
- [8] W.H. Cunningham. Optimal attack and reinforcement of a network. *Journal of the ACM (JACM)*, 32(3):549-561, 1985.
- [9] M. Dziubiński and S. Goyal. Network design and defense. *Games and Economic Behavior*, 79:30-43, 2013.
- [10] S. Goyal. *Connections: an introduction to the economics of networks*. Princeton University Press, 2012.
- [11] S. Goyal and A. Vigier. Attack, defense, and contagion in networks. *The Review of Economic Studies*, 81(4):1518-1542, 2014.
- [12] M. Groetschel, C.L. Monma and M. Stoer. Design of survivable networks. *Handbooks in Operations Research and Management Science*, 7:617-672, 1995.
- [13] H. Haller. Network vulnerability: a Designer-Disruptor game. *Working paper*, 2015, <https://ideas.repec.org/p/vpi/wpaper/e07-50.html>.
- [14] H. Haller and S. Sarangi. Nash networks with heterogeneous links. *Mathematical Social Sciences*, 50(2):181-201, 2005.
- [15] F. Harary. The maximum connectivity of a graph. *Proceedings of the National Academy of Sciences of the United States of America*, 48(7):1142, 1962.
- [16] S. Hong. Enhancing transportation security against terrorist attacks. *Working paper*, Vanderbilt University, 2009, <http://gtcenter.org/Archive/Conf09/Conf/Hong778.pdf>.
- [17] B. Hoyer and K. De Jaegher. Strategic network disruption and defense. *Journal of Public Economic Theory*, 2016.
- [18] M.O. Jackson. *Social and economic networks*, volume 3. Princeton University Press, 2008.
- [19] M.O. Jackson and A. Wolinsky. A strategic model of social and economic networks. *Journal of economic theory*, 71(1):44-74, 1996.
- [20] H. Kerivin and A. R. Mahjoub. Design of survivable networks: A survey. *Networks*, 46(1):1-21, 2005.
- [21] J. Landwehr. Network design and imperfect defense. 2015. Center for Mathematical Economics Working Paper No. 537, <https://ideas.repec.org/p/bie/wpaper/537.html>.
- [22] M. McBride and D. Hewitt. The enemy you can't see: An investigation of the disruption of dark networks. *Journal of Economic Behavior & Organization*, 93:32-50, 2013.
- [23] M. Tambe. *Security and game theory: Algorithms, deployed systems, lessons learned*. Cambridge University Press, 2011.
- [24] F. Vega-Redondo. *Complex social networks*. Number 44. Cambridge University Press, 2007.