

Network Design and Defence

Marcin Dziubiński^{*†}

Sanjeev Goyal[‡]

July 6, 2012

Abstract

Infrastructure networks – in communication and transport – are a key feature of an economy. The functionality of these networks depends on the connectivity and sizes of different components. However, these networks face a variety of threats ranging from natural disasters to intelligent attacks (carried out by human agents). How should networks be defended and designed to ensure connectivity?

We develop a strategic model to study this question. There are two ‘players’, a designer and an adversary. The designer forms links among a given set of n nodes. These links are costly. In addition, the designer may also choose to protect some nodes at a cost. The adversary then allocates his resources to attack nodes. Successful attack on a node leads to its elimination (along with all its connections). We study the sub-game perfect equilibrium of this game.

Our main finding is that if defence is affordable and reliable, then the network is sparse and heterogeneous, and either centrally or fully protected. On the other hand, if defence is relatively costly compared to linking then dense and homogeneous networks arise in equilibrium.

^{*}Faculty of Economics, University of Cambridge. Email: mkd29@cam.ac.uk

[†]Institute of Informatics, Faculty of Mathematics, Informatics and Mechanics, Warsaw University

[‡]Faculty of Economics & Christ’s College, University of Cambridge. Email: sg472@cam.ac.uk

We thank Oliver Baetz, Diego Cerdeiro, Andrea Galeotti, Michael Koenig, Dan Kovenock, Asuman Ozdaglar, and seminar participants at Cambridge, MIT, Microsoft (Cambridge and New England) and WEHIA 2012 (Paris) for useful comments.

1 Introduction

Infrastructure networks – in communication and transport – are a key feature of an economy. The functionality of these networks depends on the connectivity and sizes of different components. However, these networks face a variety of threats ranging from natural disasters to intelligent attacks. How should networks be defended to ensure connectivity? The study of optimal defence in turn raises the question of how should networks be designed to economize on defence?

We develop a model with the following features: there are two ‘players’, a designer and an adversary. The designer forms links among a given set of n nodes. These links are costly. In addition, the designer also chooses defence levels for the nodes. The adversary then attacks the network by removing up to k nodes. The goal of the designer is to maximize a payoff which depends on the value of the network less the costs of linking and defence, while the goal of the adversary is to minimize this payoff. We study the sub-game perfect equilibrium of the game.

We start with an analysis of a benchmark model based on the connectivity problem: here the designer only gains a positive payoff if the surviving nodes constitute a connected network.¹ As links are costly, the designer would like to form the minimum number of links (which is $n - 1$) to connect the n nodes. However, in a minimal connected network there will always exist nodes with a single link. It then follows that elimination of a single non-terminal node will disconnect such a network. One way forward for the designer is to add links thereby making individual nodes less critical for connectivity of the network. Alternatively, the designer can protect specific nodes, and keep the linking costs down. In the connectivity game, this trade-off between linking and protection finds a sharp resolution: in equilibrium the designer chooses either an empty network with no defence, a $(k + 1)$ -connected network with no defence or a star network with protected center (Proposition ??).² This result shows that if defence is affordable then protection is centralized, linking is heterogeneous and the network is sparse. If, on the other hand, defence is relatively costly compared to linking then dense and homogeneous networks arise in equilibrium.

In the second part of the paper we consider general payoff functions in which the value to the designer of a network is component additive, and the payoff from each component

¹A network is said to be connected if there is a path between every pair of nodes in it.

²An empty network contains no links, a $(k + 1)$ -connected network remains connected upon the deletion of k or fewer nodes but there exist a set of $k + 1$ nodes whose removal disconnects the network. A star network has $n - 1$ nodes connected to a single node and no other links; for formal definitions, see Section ??.

is increasing and convex in the size of the component. This smoother value function leads to richer phenomena. Firstly, it gives positive value to individual nodes in the network. Because of that, protecting all the nodes becomes attractive, if the defence is sufficiently cheap. Secondly, it gives positive value to disconnected networks. Therefore it may be profitable for the Designer to sacrifice small parts of the network by saving on the cost of linking. The main result of this paper illustrates how the insights of Proposition ?? generalize to this setting. Designer chooses either 0, 1 or n units of defence. In case of 1 unit of defence, the star network remains optimal. In case of n units of defence, the equilibrium network is either empty or a tree (Theorem ??). In case of 0 defence the network ranges from the empty network to $(k + 1)$ -connected and its structure depends on the shape of the component payoff function. If this function is sufficiently close to the connectivity function, then only empty or $(k + 1)$ -connected networks can be optimal. If it less convex, however, then other types of optimal networks are possible. Interestingly, if the number of units of attack is small, $k = 1$, or large, $k = n - 2$, then again only empty network or $(k + 1)$ -connected networks can be optimal, independent of the shape of the payoff function (Proposition ??).

There are two assumptions which are maintained through the above analysis. One is, that defence of a node is perfect. We consider different ways of relaxing this assumption: our main finding is that, when defence is imperfect, the Designer may find it optimal to protect multiple nodes. In particular, we identify a class of networks and protection profiles – complete bipartite networks with the smaller side protected – which arise in equilibrium. The second assumption is that the budget of the adversary is exogenous. We show that the attractiveness of centrally protected star and $(k + 1)$ -connected networks are robust to making the adversary budget endogenous.

Our paper contributes to two rich strands of research: the theory of networks in economics and the theory of reliable networks.

The research on networks has been concerned with the formation, structure and functioning of social and economic networks; for a survey of this work see ?, ? and ?.³ A key issue is the robustness of networks to different types of shocks – natural and man made. This subject is relatively little studied. The contribution of our paper is to build on a classical problem in graph theory: how to design networks to retain connectivity in the face of node and link removal. Our paper generalizes the class of payoffs beyond connectivity and supplements the design problem with the possibility of defence.

³There is also a distinguished tradition of research in communication networks, see e.g., ?, ?, ?, and ?.

In a recent paper, ? study the problem of design and defence of a network which faces an intelligent adversary. There is one key difference between the present paper and ?. In the present paper we assume that attack is restricted to individual nodes; by contrast, in ?, the focus is on the case where attack spreads smoothly through links of a network. Thus the two papers consider the two polar cases of zero and perfect contagion. The results in the two papers bring out the implications of this contagion: in the present paper, when defence units are 0, the designer defends the network by adding more links and so the optimal network is $(k + 1)$ -connected. By contrast, in ?, when defence units are 0, the designer defends the network by separating it into distinct components. In equilibrium with defence units present, in the present paper the equilibrium will typically involve protection of multiple nodes. By contrast, ? show that the designer will assign all resources to the central node of a star.⁴

The research on reliable networks dates at least fifty years back and spanned over several disciplines, including graph theory, operations research and computer science (for surveys see ???). The basic theoretical model is the classical problem of connectivity of graphs; for a statement see ?. What is the number of links needed to ensure that a network remains connected after k nodes (out of n) are removed? The answer to this problem was provided by ? who showed that exactly $\lceil n(k + 1)/2 \rceil$ links are needed.⁵ Our paper generalizes the classical connectivity problem in two dimensions: one, we allow for general increasing and convex function of networks and two, we allow for defence of nodes. We build on Harary's result to characterize equilibrium of a game of conflict between a designer and an intelligent adversary.

The rest of the paper is organized as follows. Section ?? presents the model of design, defence and attack in networks. Section ?? provides a characterization of outcomes when the payoff takes the form of connectivity based value function. Section ?? and Section ?? explore the scope of the results in the basic model. Section ?? concludes.

⁴ In an early paper, ? study non-cooperative network formation among nodes faced with an exogenously given uniform probability of link deletion. By contrast, our focus is on defence and design of a network facing an intelligent adversary. For other papers on optimal networks which face attacks with contagion risk, see ? and ?. In all these papers, there is no technology for defence of nodes/links. For a recent study of conflict in exogenous networks, see ?

⁵ Given $x \in \mathbb{R}$, $\lceil x \rceil$ denotes the minimal integer i such that $i \geq x$. Similarly, $\lfloor x \rfloor$ denotes the maximal integer i such that $i \leq x$.

2 Model

We consider a 2-player game with a Designer (\mathcal{D}) and an Adversary (\mathcal{A}). There are two stages in the game. In the first stage, the Designer chooses a network and a subset of nodes to defend. In the second stage, the Adversary observes the network and defence and then chooses a subset of nodes to attack.

Let $N = \{1, \dots, n\}$, with $n \geq 3$ be a given finite set of nodes. The Designer \mathcal{D} chooses links between pairs of nodes and creates a network $g \subseteq g^N = \{ij : i, j \in N, i \neq j\}$, where ij is an abbreviation for $\{i, j\}$.⁶ He also chooses a subset of nodes to protect, $\delta \subseteq N$.

In the second stage, the Adversary assigns his attack budget $k > 0$ to a subset of nodes $X \subseteq N$, with $|X| \leq k$. This strategy is called a *cut*. Given the network g the removal of X nodes creates a *residual network* $g - X = \{ij \in g : i, j \in N \setminus X\}$.

We assume that defence is perfect: a protected node cannot be removed by an attack, while any attacked unprotected node is removed with certainty. Given a defence δ and a cut X , a set $Y = X \setminus \delta$ will be removed from the network.

Both links and defence resources are costly: each link in a network has a positive cost $c_l > 0$ and each unit of defence costs $c_d > 0$. The cost of a defended network (g, δ) is $c_l|g| + c_d|\delta|$. The payoff to the Designer from choosing (g, δ) when the Adversary chooses a cut X is

$$\Pi^D(\delta, X; g) = \Phi(g - (X \setminus \delta)) - c_l|g| - c_d|\delta|. \quad (1)$$

The payoff to the adversary is

$$\Pi^A(\delta, X; g) = -\Phi(g - (X \setminus \delta)). \quad (2)$$

Two nodes i and j are *connected* if they have a link between them, or if there exists a sequence of nodes i_0, \dots, i_l , called a *path* from i to j , such that $i = i_0$, $j = i_l$ and for all $1 \leq r < l$, $i_r i_{r+1} \in g$. Two paths i_0, \dots, i_l and $j_1, \dots, j_{l'}$ are *node independent* if the sets of their internal nodes are disjoint, $\{i_1, \dots, i_{l-1}\} \cap \{j_1, \dots, j_{l'-1}\} = \emptyset$. A network g is *connected* if every two nodes $i, j \in N$ are connected. A non-empty set of nodes such that there is a path between each pair of them is called a *component*. The set of all components of g is denoted by $\mathcal{C}(g)$.

The objective of the Designer is to maximize the value of the residual network at minimum

⁶ Notice that the network is undirected, with links from a node to itself forbidden.

cost; the objective of the Adversary is to minimize the value of the residual network. We are interested in the sub-game perfect equilibria of the game (which will be referred to as ‘equilibrium’, for short). Note that in every sub-game, facing a defended network (g, δ) , the Adversary will choose a set of nodes to eliminate such that:

$$\Phi(g - (X \setminus \delta)) = \min_{\substack{Y \subseteq N \\ |Y| \leq k}} \Phi(g - (Y \setminus \delta)). \quad (3)$$

Therefore, the Designer selects a network-defence combination (g, δ) such that for all cuts X with $|X| \leq k$

$$\pi^D(\delta, X; g) \geq \max_{\substack{g' \subseteq g^N \\ \delta' \subseteq N}} \left(\left(\min_{Y \subseteq N, |Y| \leq k} \Phi(g - (Y \setminus \delta')) \right) - (|\delta'|c_d + |g'|c_1) \right). \quad (4)$$

We study this problem starting with a benchmark model where the the value function is based on whether the residual network is connected or not. Despite its simplicity, the model provides us with insights that can be used to study a more general value function which is component additive and the value of each component is a strictly increasing and convex function of its size.

3 Benchmark: The connectivity problem

We start with an analysis of a benchmark model based on the connectivity problem: here the Designer only gains a positive payoff if the surviving nodes constitute a connected network. Our main result is that if defence is affordable then protection is centralized, linking is heterogeneous and the network is sparse. If, on the other hand, defence is relatively costly compared to linking then dense and homogeneous networks arise in equilibrium. This result is obtained under the assumption that the Adversary budget is exogenous and that defence is perfect. We consider different ways of relaxing these assumptions. We show that the attractiveness of centrally protected star and $(k + 1)$ -connected networks is robust to making the adversary budget endogenous. We show that, when defence is imperfect, the Designer may find it optimal to protect multiple nodes. In particular, we identify a class of networks and protection profiles – complete bipartite networks with the smaller side protected – which arise in equilibrium.

Formally, in the benchmark model, we assume that the value function Φ of residual network is:

$$\Phi(g) = \begin{cases} 1, & \text{if the network is connected,} \\ 0, & \text{otherwise.} \end{cases} \quad (5)$$

Therefore, the Designer either chooses the empty network with no defence or selects a lowest cost network-defence combination (g, δ) such that for all $X \subseteq N$ with $|X| \leq k$, $g - (X \setminus \delta)$ is connected.

Before we state our result on network-defence combinations chosen in equilibrium we need to introduce a notion of minimal k -connected networks. A cut $X \subseteq N$ is a *separator* for network g if $g - X$ is not connected. Given $k \geq 2$, a network g is *k -connected* if either $|N| = k + 1$, or $|N| \geq k + 2$ and there exists a separator $X \subseteq N$ with $|X| = k$ for g , and any cut $X \subseteq N$ with $|X| < k$ is not a separator for g . A network of connectivity level k which minimizes the number of links is called *minimal k -connected*. The set of minimal k -connected networks is denoted by $\mathcal{M}(k, n)$.

It is easy to see that every node of a k -connected network has degree at least k , as otherwise it could be separated by removing all its neighbors. Thus the minimal number of links in such a network is $\lceil nk/2 \rceil$. As was shown by ?, this number is also sufficient. The proof of this result is constructive – Harary describes how to obtain the desired family of graphs, which are now called *Harary graphs*.⁷ Notice that Harary’s result implies that all networks in $\mathcal{M}(k, n)$ are almost regular (possibly with one node having degree $k + 1$, when nk is odd). To give the reader some idea of how Harary networks look like, we provide some examples in Figure ???. For full description of the construction the interested reader is referred to ?.

The main result of this section provides a characterization of equilibrium network architecture, defence and attack.

Proposition 1. *Consider the Designer-Adversary game and suppose that $k \leq n - 2$. In equilibrium*

1. *The Designer chooses protected network (g, δ) which is as follows:*

- *If $c_1 < 1 / \lceil \frac{n(k+1)}{2} \rceil$ and $c_d > c_1 \left(\lceil \frac{n(k-1)}{2} \rceil + 1 \right)$, then $g \in \mathcal{M}(k, n)$ and $\delta = \emptyset$.*

⁷ Throughout the paper we will call them *Harary networks*. We would also like to note that obtaining these networks is computationally easy, and they can be constructed in time linear with respect to the number of links required.

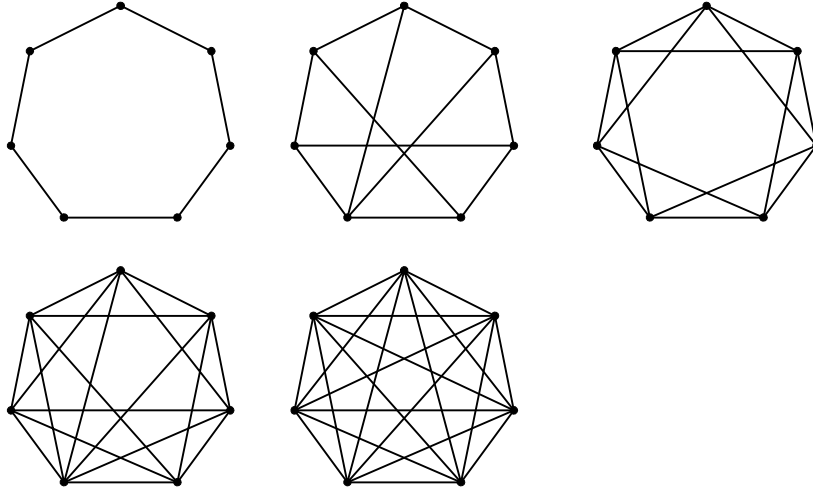


Figure 1: Harary networks over $n = 7$ nodes with connectivity $k = 2, \dots, 6$.

- If $c_1(n - 1) + c_d < 1$ and $c_d < c_1 \left(\left\lceil \frac{n(k-1)}{2} \right\rceil + 1 \right)$, then g is a star and the central node is protected.
- Otherwise g is empty and $\delta = \emptyset$.

2. The Adversary chooses: a separating cut for (g, δ) , if it exists, if it does not exist then all cuts yield the same payoff.

It follows that payoff to \mathcal{D} is $1 - \lceil n(k+1)/2 \rceil c_1$ in the case of $(k+1)$ -connected network, $1 - c_d - (n-1)c_1$ in the case of center protected star and 0 in the case of empty network. The payoff to \mathcal{A} is -1 when faced with a $(k+1)$ -connected network or a center protected star, and it is 0 faced with the empty network. Possible optimal networks are presented in Figure ??.

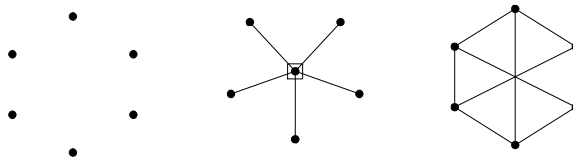


Figure 2: Equilibrium in connectivity problem: $n = 6$ $k = 2$

Proof. In the first step we clarify the nature of networks under zero or positive defence. In case of no defence the network is either empty or connected. If it is connected, then the residual

network must be connected as well, irrespective of any cuts. Thus in case of no defence \mathcal{D} must choose either empty or minimal $(k + 1)$ -connected network. In case of positive defence, the residual network must again be connected. So the initial network set up by \mathcal{D} must be connected. Observe that connectedness of the residual network can be guaranteed by a star network with 1 unit of defence assigned to the centre. This protected network has minimal number of links across all connected networks and minimal defence across all networks with positive defence. Any other network with the same number of links must be a tree, and any tree requires protecting all internal nodes (i.e. non-leaves) to stay connected after an attack of the Adversary. Thus the center-protected star is the only candidate for an optimizing \mathcal{D} .

The payoff to the Designer from the empty network is 0. Observe that in a $(k + 1)$ -connected network every node must have degree at least $k + 1$; otherwise it can be isolated by the deletion of its neighbors. ? proved that degree $k + 1$ for every node – so $\lceil n(k + 1)/2 \rceil$ links in all – are also sufficient for $k + 1$ connectedness. Applying this theorem, we can write the payoff from the minimal $(k + 1)$ -connected network as

$$1 - c_1 \left\lceil \frac{n(k + 1)}{2} \right\rceil. \quad (6)$$

Finally, the payoff from the centrally protected hub-spoke network is

$$1 - c_1(n - 1) - c_d \quad (7)$$

A comparison of the payoffs from the empty network, a minimal $(k + 1)$ -connected network and a centrally protected star network yields the desired parameter restrictions. \square

The characterization of optimal defended networks, for different values of costs, is summarized in Figure ???. Here we define $R(n, k) = n(k + 1) \bmod 2$, so $R(n, k) = 1$ if n is odd and k is even, and $R(n, k) = 0$, otherwise. We can see that if cost of linking, c_1 , is higher than $1/\lceil n(k + 1)/2 \rceil$, then only a centrally protected star or the empty network can be optimal. Raising the cost of defence makes the empty network more attractive. Similarly, if cost of defence, c_d , is higher than $1 - \frac{2(n-1) - n(k+1) \bmod 2}{n(k+1) - n(k+1) \bmod 2}$, then only a minimal $(k + 1)$ -connected network or the empty network can be optimal (and raising the cost of linking makes the latter more attractive). On the other hand, if the costs are sufficiently low, $c_1 < 1/\lceil n(k + 1)/2 \rceil$ and $c_d < 1 - \frac{2(n-1) - n(k+1) \bmod 2}{n(k+1) + n(k+1) \bmod 2}$, then either a minimal $(k + 1)$ -connected network or a centrally protected star is optimal, depending on the relation between the costs, c_d/c_1 . If c_d/c_1 is sufficiently low, i.e. lower than $\lceil n(k - 1)/2 \rceil + 1$, then centrally protected star is optimal, and if

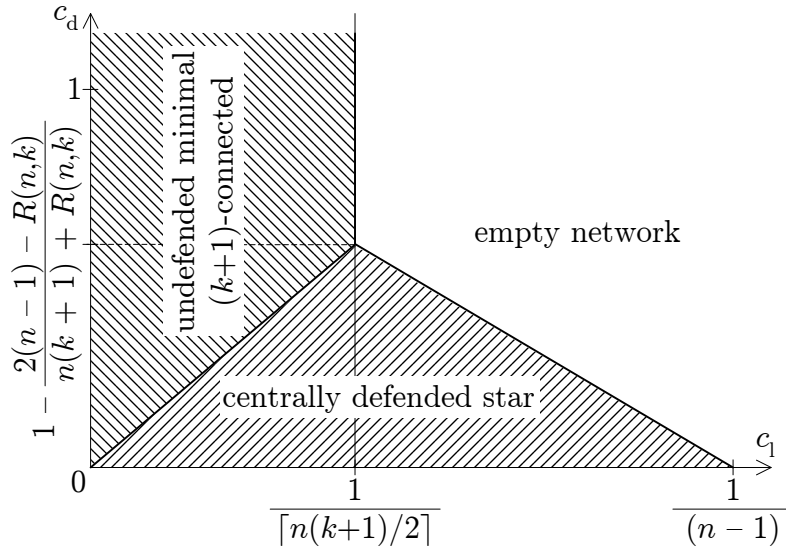


Figure 3: Equilibrium in connectivity problem: parametric characterization

it is higher, then minimal $(k + 1)$ -connected network is optimal.

3.1 Discussion

Proposition ?? provides a sharp characterization of equilibrium in the design and defence game. It also illustrates the trade off between two possible ways of defending a network. The Designer could either use defence and connect the network minimally, or construct a network with no defence and larger number of links, which guarantee sufficient redundancy in the number of paths between any two nodes.⁸

Below we explore the scope of the result by discussing the role of assumptions on the nature of the attack and the defence technology.

3.1.1 Assumptions on attack

In the benchmark model, we make three assumptions on the attack: (i) the budget is exogenous, (ii) it is perfectly known to the Designer and (iii) the Adversary observes the network and chooses the attack optimally. We will argue that the the insights of Proposition ?? – with

⁸ As as shown by ?, a network is k -connected if there are at least k node independent paths between any two nodes. Notice that this implies, in particular, that for each pair of nodes that are at least $k(k - 1)/2$ cycles containing both of them.

respect to the units of defence and the architecture of networks – are robust with respect to these assumptions.

Let us start with the assumption that Adversary budget is exogenous. Suppose attacking each node has a cost $c_a > 0$. Given the connectivity objective of the designer, the adversary will choose a maximum of k units of attack where $k \cdot c_a = 1$. So the Designer will play as in the game studied above with $k = 1/c_a$ units of attack. This implies that as c_a increases, the maximum number of units of attack falls, and this makes the centrally defended star less attractive relative to the $(k + 1)$ -connected network. An increase in cost of linking, c_l , makes the center-protected star more attractive, while increases in cost of defence, c_d , result in a decrease of attractiveness of centrally-protected star.

Next consider the issue of perfect knowledge of adversary budget. Suppose, for example, that the set of possible types (i.e. the number of possessed units of attack) is finite and is $K = \{1, \dots, k^*\}$. Let beliefs of the Designer about the type of the Adversary be represented by a discrete CDF $F : K \rightarrow [0, 1]$. Then we can apply the arguments in Proposition ?? and conclude that designer will choose an empty network with no defence, or a centrally defended star or a $(k + 1)$ -connected network with minimal number of links. In the last case the level of connectedness k^* would solve:

$$\max_{k \in K} F(k) - c_l \left\lceil \frac{n(k+1)}{2} \right\rceil. \quad (8)$$

Finally, consider the situation when the adversary is not strategic but removes k nodes from g at random. Both the centrally protected star and minimal $(k + 1)$ -connected networks remain connected after any random attack of size k . But less connected and undefended nodes may dominate. Observe that the maximum possible payoff for \mathcal{D} from a connected network is $1 - (n - 1)c_l$. So the maximum payoff for the designer is $\max(0, 1 - (n - 1)c_l)$. A network-defence pair (g, δ) is said to be ε -optimal if

$$p(g, \delta, k) - |g|c_l - |\delta|c_d > \max(0, 1 - (n - 1)c_l) - \varepsilon, \quad (9)$$

where $p(g, \delta, k)$ is the probability of (g, δ) staying connected under a random removal of k nodes. In the case of a star network

$$p(g, \delta, k) = \frac{n - k}{n} \rightarrow 1, \text{ when } n \rightarrow +\infty. \quad (10)$$

Thus under random attack, if the number of nodes is sufficiently large, the star network is ε -optimal.

3.1.2 Imperfect Defence

In the benchmark model we assume that defence is perfect. This is clearly a strong assumption and it is important to understand how our result depends on it.

There are different ways in which imperfect defence can be formulated. Here we will suppose that there is a probability $\pi \in [0, 1]$ that a defended node can be eliminated by attack. Observe that the basic model above is a special case when $\pi = 0$, while $\pi = 1$ is the case where defence is useless, and so we are back in the pure design problem, with $|\delta| = 0$. The interest here is in understanding how networks and defence chosen in equilibrium depend on the value of π .

We have been unable to provide a general characterization of this problem. The following example brings out three features of the problem which we believe are quite general: One, the outcomes obtained in the perfect defence case – empty network, the center protected star, and the $k + 1$ connected network – remain equilibria (depending on parameter values). Two, as defence becomes less reliable the designer may find it optimal to protect multiple nodes and create bi-partite networks, with one protected side. This generalizes the center protected star network obtained in the basic model. Three, we find that richer defence and network patterns – all nodes protected in a cycle network – which trade off the costs of defence against the cost of linking, may arise in equilibrium.

Example 1 (Imperfect defence). Let $n = 6$ and $k = 2$. The optimal defended networks in this case are either the empty network, a centrally protected star, a complete bipartite network with one part of size 2 fully protected, and the other side of size 4, a fully protected 2-connected network with minimal number of links (a cycle), and a 3-connected network with minimal number of links and no defence. The networks are summarized in Figure ??.

To see that no other defended network can be optimal, observe first that any such network must be 2-connected, as the list above contains optimal 1-connected defended networks (centrally protected stars) and optimal defended network which are more than 2-connected (3-connected networks with minimal number of links). Secondly, observe that any such network must have 2, 3, 4 or 5 units of defence. This is because the optimal networks with 1 and 0 units of defence are already included in the list above. In the case of 6 units of defence, the cycle (which is the unique 2-connected network with minimal number of links) is also included.

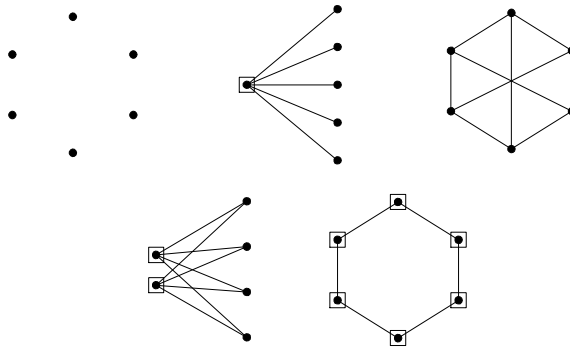


Figure 4: Equilibrium with Imperfect defence: $n = 6$ and $k = 2$.

Thirdly, observe that any such network must have 6 or 7 links. It must have at least 6 links, as it must be 2-connected. It cannot have more than 7 links, as the highest probability of staying connected after an optimal attack it can guarantee is $1 - \pi^2$ (if every separating 2-cut contains 2 protected nodes), and this probability is guaranteed by a complete bipartite network with ones side of two nodes fully protected, which has 8 links. Lastly, observe that any node in such a network either has two protected neighbors or has degree at least three. This is because with no protected neighbor and degree at most two the node could be separated by a 2-cut removing all its neighbors. With one protected neighbor and degree at most two the probability of disconnecting the node (and consequently the whole network) is π , like in the case of centrally protected star, while the cost of centrally protected star is lower.

To see that optimal defended network with 2, 3, 4 or 5 protected nodes is not possible, suppose that it has 6 links first. In this case the network must be a cycle (the unique 2-connected network with minimal number of links). Notice that with less than 6 protected nodes, such a network always has a node of degree 2 neighboring an unprotected node. This contradicts the fourth observation above and implies that the network cannot have 6 links. Suppose that it has 7 links. It is easy to check that any 2-connected network with 7 links must contain a cycle over all nodes. Now, with less than 5 protected nodes, there are at least three nodes in this cycle with an unprotected neighbor and since the seventh link can connect only two of them, at least one node is of degree two and has one unprotected neighbor. This contradicts the fourth observation above and the remaining case is 5 protected nodes (and the two neighbors of an unprotected node connected with a link). For such a network to be better than a cycle with all nodes protected and better than a complete bipartite network with one

side of two nodes fully protected, the following inequalities must be satisfied:

$$1 - \pi^2 - 5c_d - 7c_1 > 1 - \pi^2 - 6c_1 - 6c_d$$

$$1 - \pi^2 - 5c_d - 7c_1 > 1 - \pi^2 - 8c_1 - 2c_d,$$

which simplifies to $c_d > c_1$ and $c_1 > 3c_d$, a contradiction.

It is elementary to verify that for any of the networks listed in Figure ?? there are ranges of values of parameters $\pi \in (0, 1)$, $c_1 > 0$ and $c_d > 0$ for which the network is the unique optimal. Notice that these optimal networks can be divided with respect to their expected value (i.e. one minus the probability of getting disconnected by an intelligent adversary):

- 0: The empty network.
- $1 - \pi$: Centrally protected star.
- $1 - \pi^2$: Complete bipartite network with protected side of two nodes and the fully protected cycle.
- 1: Unprotected 3-connected network.

At level $1 - \pi^2$ there are two possible networks that can be optimal. If the cost of linking is small and cost of defence is high, then complete bipartite network with two nodes protected is better, while with low cost of defence and higher cost of linking, fully protected cycle is better.

We are unable to obtain a complete characterization of equilibrium outcomes in the imperfect defence model. But we can make a number of observations on general properties of equilibrium, which are worth reporting. Firstly, notice that for any n , k and π there is a unique family of networks that yield the expected value $1 - \pi$: centrally protected stars. This is because they have the lowest cost among connected networks with protection. Similarly with the family of networks that yield the expected payoff 1: these are the $(k + 1)$ -connected networks with minimal number of links.

Secondly, an increase in the number of units of attack, k , leads to increase of the cost of networks that yield value higher than $1 - \pi$. If the costs of linking and defence are sufficiently low, so that the complete network dominates all the networks with protection, for all k , then increase in k makes subsequent $(k + 1)$ -connected networks with minimal number of links optimal. If the cost of linking is high but cost of defence is sufficiently low, so that centrally

protected star dominates the empty network, then increase in k increases the attractiveness of centrally protected star. If the costs are sufficiently high, then increase in k increases the attractiveness of the empty network.

Increase in cost of defence, c_d , increases the attractiveness of networks without protection: the empty network and $(k + 1)$ -connected networks. Increase in cost of linking, c_1 , increases the attractiveness of the empty network and the centrally protected star. Increase in the probability of protection failure, π , decreases the attractiveness of networks with protection, which makes the empty network or $(k + 1)$ -protected networks more attractive.

4 General Value Function

In the benchmark we assumed that the designer derives a payoff of 0 from any (residual) network which is not connected. We now relax this assumption and study a value function which is component additive and in which the value from each component, f , is an increasing and convex function of the size of component. This formulation covers a variety of models which have been studied in the economic literature⁹ The connectivity problem analyzed above is its limiting case. Our analysis yields two insights: firstly, when designer uses defence the nature of equilibrium networks is similar to what we found in the connectivity problem, secondly, when he chooses 0 defence, a variety of networks which lie between the empty network and the $(k + 1)$ -connected network may arise in equilibrium.

Given a (residual) network g , the payoff

$$\Phi(g) = \sum_{C \in \mathcal{C}(g)} f(|C|), \tag{11}$$

where $\mathcal{C}(g)$ is the set of components of g and $f : \mathbb{R}_+ \rightarrow \mathbb{R}_+$. We will assume that f is (strictly) increasing and (strictly) convex. This assumption will be referred to as **(A.1)**.

Our analysis of the general model is summarized in the following result.

Theorem 1. *Consider the Designer-Adversary game. Suppose $1 \leq k \leq n - 2$, and **(A.1)** holds. In equilibrium*

- *The Designer chooses defence $|\delta| = 0, 1$ or n . In case $|\delta| = 0$, a variety of networks – including the empty network and a $(k + 1)$ -connected network – can arise. In case*

⁹ For example this formulation covers the connections model ? and ?. For a discussion of other examples, see ?.

$|\delta| = 1$ the network is a star with a protected center. In case $|\delta| = n$ the network is either empty or minimally connected (a tree).

- The Adversary chooses a separating cut if (g, δ) permits such a cut. When (g, δ) permits no separating cuts, every cut yields the same payoff to the Adversary and is optimal.

The main difference with the connectivity problem is that disconnected networks yield different payoffs depending on the size of the components in the network. By contrast, in the connectivity problem all disconnected networks yield payoff 0 to the Designer. This potentially complicates the design and defence problem greatly. To make progress, we proceed in steps. The first step brings out an important implication of the convexity of payoffs: designer will choose defence units 0, 1 and n . The second step characterizes the networks which accompany the positive defence choice: in case of 1 unit of defence the network is a star and in case of n units of defence the network is either empty or a tree. Possible optimal networks are presented in Figure ???. Proposition ??? below provides a partial characterization of equilibrium networks when defence $\delta = 0$

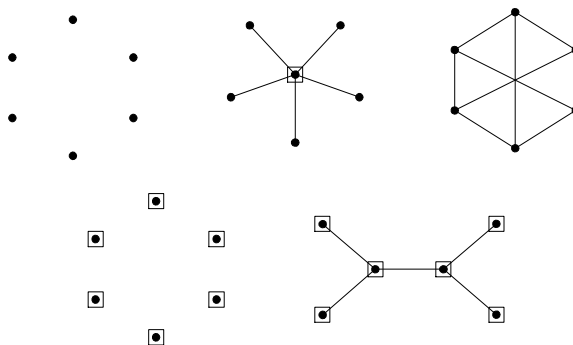


Figure 5: Equilibrium with General Payoffs: $n = 6$ and $k = 2$

Proof. Suppose that (g, δ) is an equilibrium strategy of the Designer and $|\delta| = d > 0$. We will show that either $d = 1$ and g is a star with protected center or $d = n$ and the network is either empty or a tree. The proof is constructive.

Let C_1, \dots, C_m be the components of g and g_1, \dots, g_m the subnetworks of g over these components. Suppose, without loss of generality, that component C_1 contains a protected node (at least one such component must exist, since $d > 0$). A component C_i contains at least

$|C_i| - 1$ links. Starting from g and keeping d constant, we construct a network g' with defence δ' as follows:

- Convert g_1 into a star network g'_1 with a protected node being the center.
- From each of the components C_2, \dots, C_m remove all but one node and connect each the removed nodes to the center of g'_1 .
- If feasible, move defence from protected nodes in one-node components to unprotected nodes in the newly created star g'_1 , thus obtaining a new defence, δ' .

Observe that g' has m components as does g , component C'_1 contains $n - (m - 1)$ nodes, while all the components C'_2, \dots, C'_m contain exactly one node. The number of links in g' is (weakly) smaller than the number of links in g (with equality only in the event that all components of g were minimally connected, i.e. trees). Any k -cut X applied to g' makes weakly less damage as compared to the cut on network g : it does not disconnect any paths between non-attacked nodes. Thus the pair (g', δ') yields a (weakly) higher payoff than (g, δ) .

There are two cases to be considered next: $d = 1$ and $d > 1$.

Case $d = 1$: If $d = 1$, then only the center of the component C'_1 is protected. As f is convex, it is optimal for the Adversary to attack component C'_1 first. So (g', δ') dominates (g', \emptyset) only if $|C'_1| > k + 1$. As this is an optimal outcome, it follows that the marginal value of the last periphery node, $n - m$, is greater than the cost of link c_1 . It now follows from convexity of f that linking an additional singleton node to the center of g'_1 is strictly profitable. Iterating, we conclude that if $d = 1$ then the optimal network is a center-protected star.

Case $d > 1$: We show that convexity of f implies that all the nodes must be defended and the network must be connected. First, observe that adversary will start by attacking unprotected periphery nodes (if any) in component C'_1 . So it is worthwhile to protect a periphery node in component C'_1 , as it will be eliminated otherwise. This implies that all unprotected periphery nodes will be eliminated in C'_1 . Now, convexity of f implies that marginal returns from protecting additional periphery nodes is increasing while the cost of protection is linear. Thus payoff to the Designer is strictly increased by protecting all periphery nodes in C'_1 . Arguments analogous to those used in the case of $d = 1$ imply that if there are singleton nodes with protection, then attaching them to the center of g'_1 will strictly increase payoff to the Designer as well. Finally, consider the situation in which all nodes from g'_1 are

protected and all singleton isolated nodes are not protected. If this is optimal, then the cost of linking and protection, $c_1 + c_d$, is smaller than the marginal value $-f(|C'_1|) - f(|C'_1| - 1) -$ of doing so. Convexity of f implies that the marginal value of adding an extra periphery node and protecting it is strictly larger, while the cost is still $c_1 + c_d$. It follows that the star network with all protected nodes strictly payoff dominates configuration (g', δ') . Observe finally that the payoff in any tree with all nodes protected is equal. This completes the proof. \square

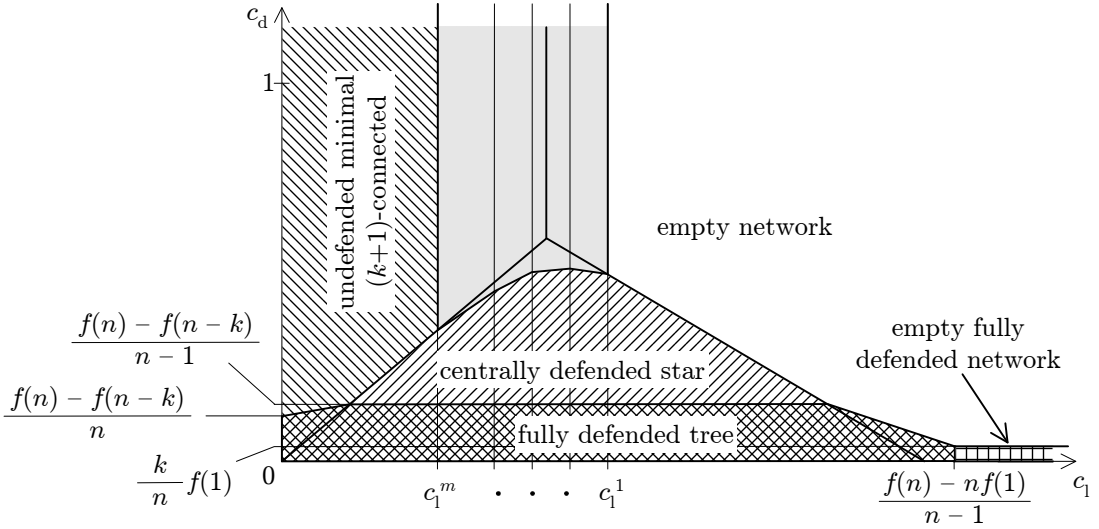


Figure 6: Equilibrium with general payoffs: parametric characterization.

The characterization of optimal defended networks for the case of general value function and for different values of costs is illustrated in Figure ???. The 'grey area' covers the networks without protection, which are not fully characterized.

We now discuss the architecture of equilibrium networks when defence $|\delta| = 0$. The following result summarizes our analysis.

Proposition 2. *Consider the Designer-Adversary game. Suppose $1 \leq k \leq n - 2$, and (A.1) holds. Consider equilibrium in which designer chooses defence $\delta = \emptyset$.*

1. *If $k = 1$ then the network is either empty or a cycle containing all nodes.*
2. *If $k = n - 2$ the the network is either empty or the complete network.*
3. *If $1 < k < n - 2$ then there exist numbers of links $0 = l_0 < \dots < l_m = l^* = \left\lceil \frac{n(k+1)}{2} \right\rceil$ and levels of cost of linking $0 = c_1^{m+1} < c_1^m < \dots < c_1^0 = +\infty$ such that for $c_1 \in (c_1^{i+1}, c_1^i)$*

- *There exists an optimal defended network with l_i links and no defence.*
- *Every optimal defended network with no defence has l_i links.*

When $k = 1$, the equilibrium network is either empty or a cycle containing all nodes. It is not profitable to have more than n links since a cycle guarantees the maximal payoff $f(n - 1)$ in the face of attack $k = 1$. A network with positive number of links less than n is not optimal due to the convexity of f . Details of the argument are given in the Appendix. When adversary budget $k = n - 2$, the equilibrium network is either the empty or the complete network: an (non-empty) incomplete network can be disconnected by an adversary with budget $k = n - 2$, so depending on the costs of linking c_1 the designer will choose either the empty or the complete network.

Proposition ?? holds out the possibility that equilibrium networks with no protection are $(k + 1)$ -connected, for general k . The following example shows that Proposition ?? cannot be extended to cover general values of k .

Example 2 (Optimal networks between empty and $(k + 1)$ -connected.). Consider f of the form

$$f(x) = \left(\frac{x}{n - k} \right)^2. \quad (12)$$

Suppose that there are $n = 28$ nodes and that the Adversary has $k = 2$ units of attack. Let g be a network constructed as follows:

- Nodes $1, \dots, 14$ are connected to form a 3-connected network with minimal number of links, which will be referred to as the core.
- Each of the nodes $i \in \{15, \dots, 28\}$ is connected to the unique node $i - 14$ and will be referred to as a periphery node.

The network has $14(3/2) + 14 = 35$ links. The Adversary cannot disconnect the core and his best strategy is to remove two nodes of the core, thus disconnecting two periphery nodes. Payoff to the Designer from g is $289/338 - 35c_1$ Payoff to the Designer from the empty network is $1/26$ and from a 3-connected network with minimal number of links is $1 - 42c_1$. From these inequalities we can deduce that there is a range of cost values ($0.0208 < c_1 < 0.0233$) for which g payoff dominates the empty network and a 3-connected network.

The example above suggests the following general method of constructing networks which would dominate both the empty network and network from $\mathcal{M}(k + 1, n)$:

- Select a fraction of nodes and connect them to form a $(k + 1)$ -connected core.
- The remaining nodes are periphery nodes, each connected to one core node and distributed evenly across the core nodes.

Observe that this strategy shares some similarity with the idea of centrally protected star: the core nodes are “protected” from getting disconnected, and the adversary can only disconnect periphery nodes from the core.

To describe this idea more precisely, take $m \in \{2, \dots, \lfloor \frac{n}{k+2} \rfloor\}$ and let h_m be a network where

- $\lceil \frac{n}{m} \rceil$ nodes form a $(k + 1)$ -connected core with minimal number of links and
- Each node from the core is connected either to $m - 1$ or $m - 2$ periphery nodes and each of the periphery nodes is connected to exactly one node of the core (that is $n - \lceil \frac{n}{m} \rceil(m - 1)$ nodes of the core are connected to $m - 1$ periphery nodes and the rest are connected to $m - 2$ periphery nodes).

Example of such networks are illustrated in Figure ???. Notice that the network in the numerical example above has $m = 2$, i.e. half of the nodes are in the core.

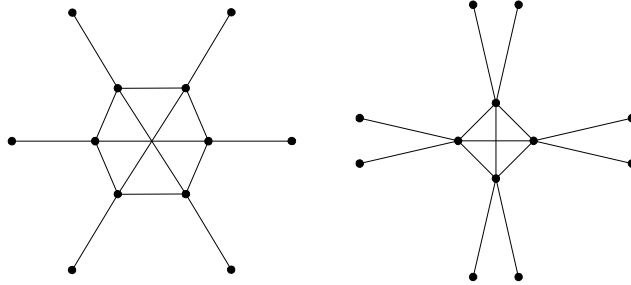


Figure 7: Example of protected hubs networks: $n = 12$, $k = 3$ and $m = 2, 3$.

Using basic algebra it can be shown that if

$$2 \leq k \leq \frac{n}{m - \sqrt{m} + 1} - \frac{4m - 2(m + 1)\sqrt{m} - 1}{2(m - \sqrt{m} + 1)} - 1 \quad (13)$$

then there exists a range of costs of linking such that the network h_m yields higher payoff to the Designer than the empty network and any $(k + 1)$ -connected network with minimal number of links. There are also ranges of values of k for which there exist ranges costs

of linking such that the network h_m yields the higher payoff than the network $h_{m'}$ for any $m' \in \{2, \dots, \lfloor \frac{n}{k+2} \rfloor\} \setminus \{m\}$. We omit them here, as it can be that none of these networks is optimal when the set of all networks over N is considered.

The example above suggests that the set of optimal networks in the case of general value function is richer than in the case of connectivity problem studied in Section ???. However with proper normalization, so that $f(n-k) = 1$, the connectivity value function defined in Equation (??) is the lower envelope, on the interval $[0, n-k]$, of the family of value functions studied in this section. This suggests that if f is sufficiently close to the connectivity value function, so that the marginal value of a single node in the residual component is sufficiently high, then we could expect that the optimal networks would be like those listed in Proposition ??. Indeed, this is the case. Consider the following property of convex and strictly increasing $f : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ normalized to satisfy $f(n-k) = 1$:

$$f(n-k) - (n-k)f(1) > \left\lceil \frac{n(k+1)}{2} \right\rceil (f(n-k-1) - (n-k-1)f(1)). \quad (14)$$

The property is illustrated in Figure ??. Notice that this property is satisfied by f if $f(n-k) >$

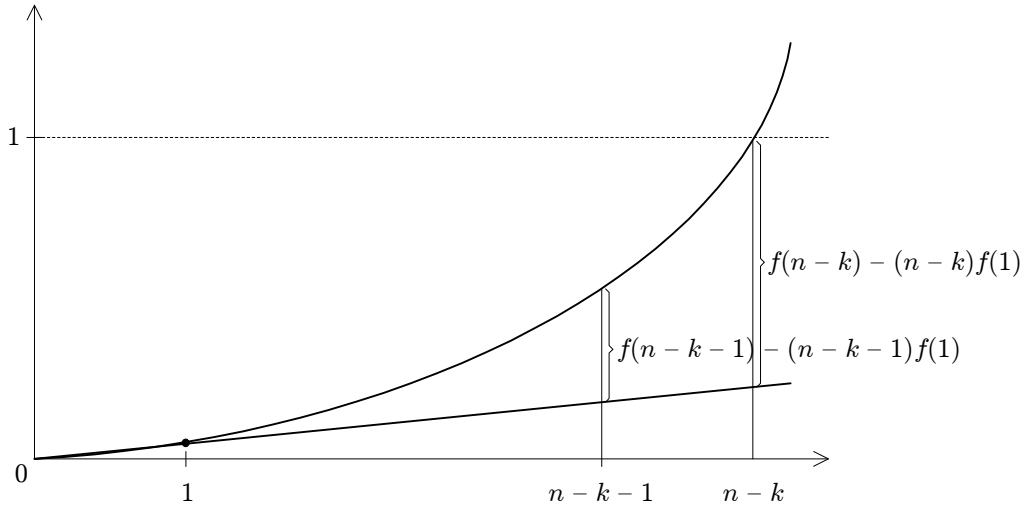


Figure 8: Sufficient condition on general payoffs.

$n(k+1)f(n-k-1)$ holds. With this observation it is easy to see that if f is sufficiently close

to the connectivity value function than it satisfies (??).¹⁰

Proposition 3. *Consider the Designer-Adversary game with $N = \{1, \dots, n\}$ and $1 \leq k \leq n - 2$. If f satisfies Condition (??) and defence $\delta = \emptyset$ then the equilibrium network is either empty or $(k + 1)$ -connected.*

Proof. Fix f , n and k . If there is $l \in \{1, \dots, l^* - 1\}$, where $l^* = \left\lceil \frac{n(k+1)}{2} \right\rceil$, and $c_1 > 0$ such that a network with l links is better than \emptyset and than any network in $\mathcal{M}(k + 1, n)$, then the following two inequalities are satisfied

$$\begin{aligned}\varphi(l) - lc_1 &\geq f(n - k) - l^*c_1, \\ \varphi(l) - lc_1 &\geq (n - k)f(1).\end{aligned}$$

The inequalities reduce to

$$\frac{f(n - k) - \varphi(l)}{l^* - l} \leq c_1 \leq \frac{\varphi(l) - (n - k)f(1)}{l}. \quad (15)$$

Inequality (??) cannot be satisfied if

$$\frac{l}{l^*} (f(n - k) - (n - k)f(1)) > \varphi(l) - (n - k)f(1). \quad (16)$$

Since $l \geq 1$ and $\varphi(l) \leq f(n - k - 1) + f(1)$, so Inequality (??) is satisfied if

$$f(n - k) - (n - k)f(1) > l^*(f(n - k - 1) - (n - k - 1)f(1)). \quad (17)$$

□

5 Conclusions

In this paper we studied the problem of network defence and design in face of intelligent adversary. We started with an analysis of a benchmark model based on the connectivity problem: here the designer only gains a positive payoff if the surviving nodes constitute a connected network. Our main result is that if defence is affordable then protection is centralized, linking is heterogeneous and the network is sparse. If, on the other hand, defence is relatively costly

¹⁰A functional form for which this property is satisfied is: $f(x) = e^{x^2} - 1$.

compared to linking, then dense and homogeneous networks arise in equilibrium. These insights carry over to a more general model, where value of the network is based on a component additive function with value of each component being convex and strictly increasing.

These results are obtained under the assumption that the budget of the Adversary is exogenous and that defence is perfect. We consider different ways of relaxing these assumptions. We show that the attractiveness of centrally protected star and $(k + 1)$ -connected networks are robust to making the Adversary budget endogenous. We show that, when defence is imperfect, the Designer may find it optimal to protect multiple nodes. In particular, we identify a class of networks and protection profiles – complete bipartite networks with the smaller side protected – which arise in equilibrium.

One interesting open problem is the characterization of optimal networks when the defence is not perfect. Preliminary work reported in this paper suggests that with defence being sufficiently reliable, the centralized strategy or protection is optimal. Lower reliability, however, may lead to a more dispersed defence patterns, including protecting all the nodes with reduced connectivity. We leave a complete analysis of this problem for future research.

Appendix

Proof of Proposition ??. The proof of parts 2 and 3 is immediate and omitted. We provide a proof of part 1. Suppose the equilibrium network is non-empty; we will show that it must be a cycle containing all nodes. Since $k = 1$, a cycle network ensures connectivity of the residual network. So a connected network contains either n links or $n - 1$ links. The tree network with $n - 1$ links will be disconnected by the adversary. It follows from the convexity of the f that the tree is either dominated by the empty network or the ring. So we move next to an equilibrium network g with multiple components.

Let the network g contain m components, C_1, \dots, C_m . Observe that every component has a maximum of $|C_i|$ links, as a cycle guarantees connectivity of the residual surviving component. Suppose that one component is a cycle and all other components are singletons. Let the cycle be component C_1 . Now it follows from convexity of f that maximum damage is caused by removal of a node in the cycle; so optimal adversary strategy entails removal of a node from the cycle. Note that marginal return to the last node in the cycle $f(|C_1| - 1) - f(|C_1| - 2) > c_1$ since g is an equilibrium network. But it follows from convexity that $f(|C_1|) - f(|C_1| - 1) > f(|C_1| - 1) - f(|C_1| - 2) > c_1$. Since f is convex, the payoff to the Designer can be raised by

expanding the cycle. A contradiction.

Next consider a network g which contains a cycle component C_1 and other components C_2, \dots, C_m , at least one of them being a non-singleton. Starting from g construct a new network g' by taking all but one node from each of the non-singleton components and adding it to the C_1 . So define $C'_1 = (\bigcup_{j=1}^m C_j) \setminus \{i_2, \dots, i_m\}$ and $C'_j = \{i_j\}$, where $i_j \in C_j$, for $2 \leq j \leq m$. Next, form a cycle over the nodes of C'_1 . Observe that the number of links in g' is not larger than the number of links in g . We will show below that g' delivers (weakly) greater value than g . Since a cycle over all nodes is strictly better than g' , g cannot be optimal. This is a contradiction.

Strict convexity of f yields the following property: Let $m \geq 2$, $n \geq m$, $n_1, \dots, n_m \geq 0$, $n_1 + \dots + n_m = n$ and $n_i < n - m$, for all $i \in \{1, \dots, m\}$. Then

$$f(n - m) + (m - 1)f(1) > f(n_1) + \dots + f(n_m), \quad (18)$$

This can be shown by simple induction, starting from the observation that $f(x - 1) + f(1) > f(x - y) + f(y)$, for any $y > 1$ and for $x \geq 3$.

Clearly, the adversary will target a node in the cycle component. Let this node be denoted by x and $X = \{x\}$ be the cut of the Adversary. The payoff from residual network $g' - X$ is

$$\Phi(g' - X) = f(n - m) + (m - 1)f(1) \quad (19)$$

Next consider the value of the residual network $g - X$. Suppose, without loss of generality, that $x \in C_m$ (we do not make assumptions on the sizes of components of g). The payoff from $g - X$ is

$$\Phi(g - X) = f(n_1) + \dots + f(n_{m-1}) + f(n_m^1) + f(n_m^2), \quad (20)$$

where $n_m^1 + n_m^2 = n_m - 1$, $n_m^1 \geq n_m^2$, are the sizes of the components which the network component over C_m falls into after removal of i_m (in particular we allow $n_m^2 = 0$). Since $n_i < n - m - 1$, for all $i \in \{1, \dots, m\}$ (because $|C_1| \geq 3$, as the subnetwork of g over C_1 is a cycle), it follows from (??) that $\Phi(g' - X) > \Phi(g - X)$.

If g does not contain any cycle components, it must be a collection of trees. Assume that sizes of the components are weakly decreasing, that is $|C_1| \geq |C_2| \geq \dots \geq |C_m|$. Let m' be the index of the last component which is not a single node; for all $m' < j \leq m$, $|C_j| = 1$ (since g is not empty, so $m' \geq 1$). We will consider two cases separately:

Case 1: $|C_1| = 2$. In this case the first m' components are two-node linear networks. The marginal benefit from adding m' links to the empty network to create g is

$$(m' - 1)f(2) + (m - m' + 1)f(1) - (m + m' - 1)f(1) = (m' - 1)f(2) - 2(m' - 1)f(1) \quad (21)$$

The marginal benefit from adding m' links to g to obtain network g' which contains a single cycle component of size $2m'$ is

$$\begin{aligned} f(2m' - 1) + (m - m')f(1) - ((m' - 1)f(2) + (m - m' + 1)f(1)) \\ = f(2m' - 1) - (m' - 1)f(2) - f(1). \end{aligned} \quad (22)$$

Marginal payoff in the latter case is strictly larger, as $f(2m' - 1) > 2(m' - 1)f(2) - (2m' - 3)f(1)$. The cost of m' links is the same, so we have shown that g cannot be optimal for the Designer.

Case 2: $|C_1| \geq 3$. We will show that g is strictly payoff dominated by other networks. Let g' be a network obtained from g by replacing the components $C_1, \dots, C_{m'}$ with a cycle and a collection of m' loose nodes. The number of links in g and g' are the same. Observe that faced by an intelligent adversary with $k = 1$ units of attack, a ‘linear’ network with z nodes performs better than any other tree with z nodes.¹¹ Assume then that $g_1, \dots, g_{m'}$ are linear networks. The intelligent adversary will remove a middle node of the largest component (it may be one of two central nodes if that component has even number of nodes). Assume, without loss of generality, that the node is removed from C_1 and let $n_0 = \lfloor (|C_1| - 1)/2 \rfloor$, $n_1 = \lceil (|C_1| - 1)/2 \rceil$, and $n_i = |C_i|$, for $i = 2, \dots, m'$. Also, let $n' = n - m + m'$ denote the number of nodes in the first m' components of g . Since $|C_1| \geq 3$ so $n_0 \geq 1$. Moreover, since $n_0 + \dots + n_{m'} = n' - 1$ and $n_i \geq 1$, for all $i \in \{0, \dots, m'\}$, $n_i \leq n' - m' - 1$ for all $i \in \{0, \dots, m'\}$. We take up the case of strict inequality and equality separately as they require different arguments.

Case 2.1: $n_i < n' - m' - 1$, for all $i \in \{0, \dots, m'\}$.

The payoff to the designer from the residual network after an optimal attack on g is

$$f(n_0) + \dots + f(n_{m'}) + (m - m')f(1), \quad (23)$$

¹¹It is always possible to find a node in a tree such that every subtree connected to that node has $\leq \lceil n/2 \rceil$ nodes and for at most one of the subtrees this holds with equality.

The value of the residual network after an optimal attack on g' is

$$f(n' - m' - 1) + m'f(1) + (m - m')f(1). \quad (24)$$

and the latter is strictly larger, by Equation (??).

Case 2.2: $n_i = n' - m' - 1$, for some $i \in \{0, \dots, m'\}$.

In this case it must be that $n_j = 1$ for all $j = \{1, \dots, m'\} \setminus \{i\}$. The equality $n_i = n' - m' - 1$ is possible only if the sizes of components $n_0, \dots, n_{m'}$ in the residual network are all equal to 1, apart from n_i . By definition of m' this is possible only if the size of the first component in g , $|C_1| = 3$, and $m' = 2$. This implies that $n_0 = n_1 = 1$ and $n_2 \in \{2, 3\}$. If $n_2 = 2$, then it is easy to see that g is strictly worse than a network g' where the first component is replaced by a single node and a two-node linear network (as it yields the same residual network but has lower cost). If $n_2 = 3$, then g consists of two linear component networks of size 3 and $m - 2$ single nodes. Compare this network with two networks: g' , consisting of two components of size two and m single nodes, and g'' , consisting of a cycle of size 6 and $m - 2$ single nodes. Marginal benefit from adding two links to g' to obtain g is $f(3) - f(2) - f(1)$, while marginal benefit from adding two links to g to obtain g'' is $f(5) - f(3) - 2f(1)$. The latter is strictly bigger, as $f(5) - f(3) > f(3) - f(1) - (f(2) - 2f(1))$. Since marginal costs are in both cases the same, so either g' or g'' is strictly better than g . This completes the proof. \square