Robust Networks

Sanjeev Goyal [*] Adrien Vigier[†]

March 10, 2010
First draft: July 2008

Abstract

Connections between individuals (firms, cities and countries) facilitate the exchange of goods, resources and information. Intelligent adversaries expend resources with to bring down these networks. Links then create value but also enable the spread of attacks. How does this tension in the role of links shape the architecture of the network?

A network is said to be *robust* if it performs well in the face of attacks.

We start with a study of a game in which study a game in which a designer chooses a network and an adversary then chooses an attack strategy. A robust network consists of equal size components whose number grows (and size falls) with the attack budget of the adversary.

We then consider the general problem of design and defence of a network which is facing an intelligent adversary. If defence and attack resources are small, relative to the number of nodes, the star network is robust and both players allocate all resources to the central node. If these resources are large then denser networks with dispersed defence strategies are robust.

[*] Faculty of Economics & Christ's College, University of Cambridge. Email: sanjeev.goyal@econ.cam.ac.uk

[†] Faculty of Economics & Queens College, University of Cambridge. Email: av301@cam.ac.uk

We are grateful to Murali Agastya, Mariagiovanna Baccara, Parimal Bag, Heski Bar-Isaac, Yann Bramoulle, Indranil Chakravarty, Matthew Elliot, Andrea Galeotti, Aditya Goenka, Matthew Jackson, Meg Meyer, Rufus Pollock, Bryony Reich, Rakesh Vohra, and seminar participants at Cambridge, Cornell, Laval, Madison, Northwestern, NUS, NYU, Paris, and Oxford for comments.

1 Introduction

Connections between individuals (firms, cities and countries) facilitate the exchange of goods, resources and information. Intelligent adversaries, who aim to impede the functioning of the network of connections, expend resources with a view to bringing down these networks. Links then become as much a liability as they are a source of value, since they enable the spread of attacks through the network. The following examples illustrate this tension in the functioning of links:

1. Transport networks: road, rail, and air connections facilitate exchange between locations, but help potential enemies transport attack from one location to another.

2. Criminal/terrorist organizations: Individuals with specialized skills communicate with each other to coordinate their actions. Communication requires connections – knowledge of identity and whereabouts, telephone numbers, e-mail addresses. These connections also expose an individual to possible threats: the detection and arrest of a well connected person may trigger a number of further arrests.

3. Computer networks: more connections enhance the efficiency of traffic management in the network, but also render the entire network more vulnerable to hackers and others.

How does this tension in the role of links shape the architecture of the network?

This paper studies the design and defence of networks which face an intelligent adversary. There are two players: a network designer and an adversary. The designer chooses a network among a set of given nodes and then allocates his defence resources to protect the nodes. The adversary observes the choices of the designer and then allocates his resources across nodes to attack the network. Successful attack on a node can potentially spread to neighboring nodes and the probability of this happening depends on the allocation of defence resources to the neighboring nodes.

Starting from a network, the interaction of defence and attack yields a set of surviving nodes and a corresponding residual network. This (residual) network defines the payoffs of the two players. The payoffs from a network reflect trade in goods and services, the exchange of information, and tasks which the nodes (or individuals located at nodes), carry out jointly. We assume that payoffs to the designer from a network are equal to the sum of the returns from the different components and that the returns from a component are increasing and convex in its size.[1] A network is said to be *robust* if it constitutes a (sub-game perfect) equilibrium outcome of the game between designer and the adversary.

We start with an analysis of the situation in which the designer has zero defence resources; this yields a game of network design and attack. Since there are no defence resources successful attack spreads across the network easily. The only way in which the designer can protect the network is by separating the network into distinct components. Since the adversary can observe the network prior to his choice of attack, he will always attack larger components in preference to smaller ones. Anticipating this, the designer chooses components to be of equal size. Moreover, the number of components grows and size falls as the attack budget of he adversary increases (Theorem 1). A fall in size of groups which communicate means that fewer and less complex tasks are performed by the network.

These findings echo discussions in the popular press. For instance, the editor of Newsweek magazine, Mr. Zakaria (2008) writes, "..the world's governments have effectively put them on the run... the Jihadists have had to scatter, work in small local cells....The terrorists have not been able to hit big, symbolic targets, especially ones involving Americans. So they blow up bombs in cafes, marketplaces, and subway stations .... They used to do terrorism, now they make video tapes".[2]

We then turn to the study of the general problem of design and defence of networks which face an adversary.[3] Given a network, an allocation of defence and attack resources gives rise to a game of conflict on the network. We suppose that the probability of successful direct attack on a node is increasing in the attack resources and decreasing in the defence resources allocated to the node.[4] The other key element is the spread of attack from one node to a neighboring node: we suppose that this probability is decreasing in the defence resources allocated to the neighboring node.

Thus the design and defence operate in tandem. On the one hand, given a network it is more attractive to protect central nodes. On the other hand, given limited defence resources

[1] A component in a network is a (maximal) set of interconnected nodes; for formal definitions see section 2. Observe that if returns from group size are concave then a collection of isolated nodes would maximize payoffs of the designer, irrespective of whether there is an adversary or not. So convex returns to component size is the interesting case for our purposes.

[2] For a study of the internal constraints on the growth of terrorist and criminal organizations, see Eilstrup-Sangiovanni and Jones (2008).

[3] Defence of networks is an important concern in many contexts. For example, countries make great efforts to protect major commercial and transport hubs; criminal organizations devote resources to protect their member's identities; anti-viruses and firewalls are installed on computers to protect them from viruses and malwares.

[4] We use the Tullock (1980) contest function to model the outcome of the conflict on a node.

it is attractive to make the protected nodes more central in the network. Our main finding is that that, if defence resources are small relative to the number of nodes then, the best course of action for the designer is to create a star network and allocate all defence resources to protect the center. Moreover, the adversary targets all his attack resources at the central node (Theorem 2).

Let us sketch the arguments underlying this result. Consider a star and suppose the designer allocates all his defence resources to the central node. For large enough number of nodes, the marginal value of eliminating a periphery node is very small compared to the value from a marginal increase in the chances of eliminating the central node (as eliminating this node then disrupts the network completely). So, in a star network the adversary is obliged to concentrate all his resources on one node. The probability of the entire network surviving is then simply proportional to the relative budgets of defence and attack.

The defence of dispersed networks – e.g., a ring or a core-periphery structure[5] – necessitates a more spread out allocation of resources. In response, the adversary can allocate his resources on these defended nodes to mimic the proportions of defence and attack resources (as in the hub of the star network). The key theoretical observation is that this profile of attack and defence resources yields a distribution on surviving nodes which stochastically dominates the distribution of surviving networks in the star network. Since the payoffs of the designer are convex in the size of the component, the payoffs of the designer are lower in the dispersed network. A comparison of Figures 5 and 6 illustrates this point: the distribution of surviving networks from a center-protected star is a mean-preserving spread of the distribution from a core-periphery network.

Empirical work has highlighted the salience of highly connected hub nodes in social and economic networks; see e.g., Barabasi (1999) and Goyal (2007). Many of these networks – such as the internet, terrorist groups and infrastructure networks – face intelligent adversaries. Theorem 2 provides a theoretical account of why hubs (and the corresponding defence allocations) are salient in such networks.

The principal contribution of this paper is to propose a tractable model of defence and attack in networks. In doing so, we build on and contribute to two rich strands of economics

[5] A ring network is a cycle containing all nodes; see Figure 2 for an illustration. A *core-periphery* network structure has two groups of nodes, the core and the periphery. The core nodes are fully linked among themselves, while the periphery nodes have a single link with one of the core nodes. Figure 4 illustrates a core-periphery network with two core codes.

research: the theory of networks and the theory of conflict/contests.[6]

The research on networks has been concerned with the formation, structure and functioning of social and economic networks; for book length surveys of this work, see Goyal (2007), Jackson (2008), and Vega-Redondo (2007). We build on the canonical model in the network literature – the connections model – to study a problem of practical interest. There is also a long and distinguished tradition of research in communication networks, see e.g., Bolton and Dewatripont (1994), Radner (1992, 1993), van Zandt (1999), and Garicano (2000). To the best of our knowledge, the present paper is the first to study design and defence of networks in the face of an intelligent adversary.[7]

The theory of conflict and contests explores the ways in which economic agents engage in conflict; well known contributions include Tullock (1967, 1980), Sandler and Hartley (2007), Dixit (1987), Hirshleifer (1991), Skaperdas (1996), and Baye (1998) and Kovenock, Baye, and de Vries (1996). An extensive literature studies conflict between two players across multiple battle sites with fixed budgets (the so-called Colonel Blotto games), see Hart (2008), Bier, Oliveros and Samuelson (2006), Powell (2008)) and Roberson (2006). The interest is in understanding the equilibrium allocation of resources and the payoffs outcomes as conflict functions and budgets vary. Our paper extends the theoretical framework in this area along two dimensions: one, we locate individual battles within a network of interconnected battles and allow for successful resources to be moved from one battle to neighboring battles, and two, we study the design of optimal interconnections across the battles.

Baccara and Bar-Issac (2007) study networks of power relations which face an adversary. The elimination of one agent leads to the elimination of connected others. The principal difference between our paper and their's is that we study of design and defence of networks, while they focus on the pure design problem. Moreover, in our paper networks facilitate communication and exchange while networks facilitate cooperation in their work. Due to these differences, the methods of analysis and the results in the two papers are quite different.

[6] There is also a literature on network security spread across disciplines such as computer science, statistical physics, engineering and operations research (Barabasi (1999); Nagaraja and Anderson (2007); Smith (2008); Levine (1999)). These literatures are vast, but to the best of our knowledge, the strategic analysis of network design and defence in the face of an intelligent adversary is novel.

[7] Bala and Goyal (2000b) study network formation among nodes faced with an exogenously given uniform probability of link deletion. Hong (2008) investigates the strategic complementarities between linking and protection. By contrast, our focus is on design and defence of a network faced by an intelligent adversary.

2 A simple model

We study a two player game between a designer $\mathcal{D}$ and an adversary $\mathcal{A}$. The designer has a collection of nodes and a protection budget, while the adversary has an attack budget. The designer moves first and chooses links between his nodes to construct a network. He then allocates resources across the nodes to protect the network. The network and the protection choices of $\mathcal{D}$ are observed by $\mathcal{A}$, who then chooses an attack strategy. The initial network design and the subsequent attack together yield a set of surviving nodes and a corresponding network which determine the realization of payoffs of the two players. The assumption that designer moves first is appropriate in the context of networks which require large physical investments or arise out of formal procedures. The leading example of the former is infrastructure networks – internet backbone, roads, railways, world wide web of links. Examples of the latter include formal lines of command and reporting relationships in an organization, e.g., an army, firm, gang, terrorist outfit.

We now set out the notation and the concepts which formally describe this game.

The designer: The designer $\mathcal{D}$ has a collection $N = \{1, ..., n\}$ of n nodes; for expositional simplicity, we will assume that n is an even number. The designer $\mathcal{D}$ chooses links between the nodes and allocates a defence budget $d \in \mathbf{N}$ across the nodes to protect the network. $\mathcal{D}$ chooses an allocation of his budget $\mathbf{d} = (d_1, d_2, ..., d_n)$, across the n nodes. This allocation must satisfy $\sum_{i \in N} d_i \leq d$. The set of such allocations is denoted by $\mathbf{D}$.

The payoffs to $\mathcal{D}$ arise out of exchange of information and goods or other tasks which the nodes, or individuals located at nodes, jointly carry out. Performing these tasks requires communication and interaction, which takes place via connections in a network chosen by $\mathcal{D}$.[8]

A link between two nodes i and j is represented by g_{ij}: we set $g_{ij} = 1$ if there is a link between i and j, and $g_{ij} = 0$ otherwise. Links are also assumed to be undirected, i.e. $g_{ij} = g_{ji}$. The links between the different pairs of individuals define a network g. The network such that $g_{ij} = 1$ for all i and j is called complete, and denoted g^c. A star network is one in which there exists a node, c say, such that $g_{ij} = 1$ if and only if $c = i$ or $c = j$. The node c is called the center of the star, while other nodes are referred to as periphery nodes.

We say that there is a path between two nodes i and j in the network g if there exists a sequence of nodes $i_{1,}.., i_k$ such that $g_{ii_1} = g_{i_1 i_2} = ... = g_{i_{k-1} i_k} = g_{i_k j} = 1$. Two nodes are said to be connected if and only if there is a path between them. A component of the network g

[8] A link between two nodes can be a physical connection (such as a road or a cable) or it may be a social link between individuals (reflecting mutual knowledge about identity, contact details, etc).

is a maximal connected subset. $\mathcal{C}(g)$ is the set of components of g; observe that $\mathcal{C}(g)$ defines a partition of N. We use $|C_k|$ to refer to the cardinality (or size) of a component $C_k \in \mathcal{C}(g)$.

A *core-periphery* network structure has two groups of nodes, $\mathcal{N}_1(g)$ and $\mathcal{N}_2(g)$. Nodes in $\mathcal{N}_1(g)$ constitute the periphery and have a single link each and this link is with a node in $\mathcal{N}_2(g)$; nodes in $\mathcal{N}_2(g)$ constitute the core and are fully linked with each other and with a subset of nodes in $\mathcal{N}_1(g)$. The star network is a special case of such an architecture in which the core contains a single node, i.e., $|N_2(g)| = 1$.

Following Myerson (1977), we assume that two nodes in the network can 'communicate' if and only if there is a path between them and that payoffs are additive across components. Let $f(m)$ denote the payoff of the designer from a component of size m. If $f(.)$ is decreasing, or concave, then a network of isolated nodes is always optimal for $\mathcal{D}$, independently from attack. Thus we restrict attention to the more interesting case in which $f(.)$ is increasing and convex. We also normalize payoffs so that $f(n) = 1$.

Assumption A.1: *The payoff to $\mathcal{D}$ from network g is given by*

$$\sum_{C_k \in \mathcal{C}(g)} f(|C_k|). \tag{1}$$

where $f(.)$ is increasing, convex and $f(n) = 1$.

The following examples illustrate the scope of the convexity and component additivity assumption. We start with the connections model.

Example 1 *The connections model*[9]

Suppose that there are n individuals who all have one piece of information which is of equal value (say) 1, to everyone. A communication link between 1 and 2 allows 1 to access 2's information as well as information which 2 may have accessed via his links with others. Thus in a communication network g, 1 has access to all others in his component C_k. The payoff to 1 is then given by $|C_k|/n^2$ (where the denominator reflects a normalization to account for number of individuals). The total payoff to all individuals in a component is $|C_k|^2/n^2$. This payoff is increasing and convex in component size. The total value of communication in network g is:

$$\sum_{C_k \in \mathcal{C}(g)} \frac{|C_k|^2}{n^2}. \tag{2}$$

[9] This model draws on Bala and Goyal (2000a), Goyal (1993), and Jackson and Wolinsky (1996).

Thus, the returns from a network are component additive. △

The next example illustrates an application relating to trading in networks.

Example 2 *Specialization and trade*

Suppose there is a group with n individuals each of whom possesses a distinct skill which enables them to produce and supply one of n distinct goods. Over time, individuals need different goods: sometimes they need a good which they can provide personally, while on other occasions they require a good which only the others can provide. If Mr. X needs a good which only Mr. Y can provide then he has to communicate and arrange for the transfer of the good and the payment for it.

In a period, one person is chosen uniformly at random as a point of demand. We also suppose that all goods are equally likely to be needed. A pair of individuals can carry out exchange if either of them is picked and demands a good corresponding to their own skill or if one of them is picked and demands the good corresponding to the skill of the other person. So the surplus which a linked pair of nodes generates is $4/n^2$. Generalizing for a subset of m nodes, we get:

$$f(m) = \left[\frac{m}{n}\right]^2 \tag{3}$$

We will occasionally work with a slightly more general version of these payoffs:

$$f(m) = \left[\frac{m}{n}\right]^\alpha \tag{4}$$

where $\alpha > 1$.

The parameter α then reflects the returns from exchange in the network. For fixed n and α, this expression is increasing and convex in m. The payoff to the designer from a network g is simply the surplus generated across the different components, and is written as:

$$\sum_{C_k \in \mathcal{C}(g)} \left[\frac{|C_k|}{n}\right]^\alpha \tag{5}$$

△

Example 3 *Complementary skills and coordination*

Consider, as before, a group with n individuals each of whom possesses a distinct skill. The objective of the group is to carry out tasks which require varying number of skills. A task may be simple and require one skill only, possessed by individual X. In this case, Mr. X does not need to coordinate his activities with anyone else and simply carries out the task on his own. A task may also be of greater complexity and require a combination of skills possessed by, say, Mr. Y, W and Z. Carrying out this task requires coordination – for instance, different actions may have to be performed in a fixed sequence – and therefore can only be carried out by individuals who communicate with each other.

A connected set of m individuals can carry out m tasks each involving a single individual, $m(m-1)/2$ tasks each involving pairs of individuals, and so forth. The total number of tasks which this group can carry out is $2^m - 1$. Moreover, there are $2^n - 1$ tasks in all.

Suppose every task is equally likely to arise. The probability that a connected set of m individuals is able to carry out a task chosen at random is

$$f(m) = \frac{2^m - 1}{2^n - 1} \tag{6}$$

This expression is increasing and convex in size m. The probability that a network g carries out the task is simply the sum of the probability across the probability across different components, and is given by:

$$\sum_{C_k \in \mathcal{C}(g)} \frac{2^{|C_k|} - 1}{2^n - 1} \tag{7}$$

△

The adversary: The adversary, $\mathcal{A}$ has $a \in \mathbf{N}$ units of resource to attack the network created by $\mathcal{D}$. In particular, we assume that the adversary observes the network g and the protection choices $\mathbf{d} = (d_1, d_2, ...d_n)$ of the designer and then makes his attack decisions. The adversary chooses an allocation of his budget $\mathbf{a} = (a_1, a_2, ..., a_n)$, across the n nodes. This allocation must satisfy $\sum_{i \in N} a_i \leq a$. The set of such allocations is denoted by $\mathbf{A}$. We now develop a model of attack and defence on networks.

In line with the economics literature on contests (e.g., Tullock (1980), Skaperdas (1996)) we assume that the probabilities of 'winning' satisfy a set of standard properties: it is increasing in own effort, decreasing in other's effort, and symmetric across players. We also assume that these probabilities are homogeneous with respect to the efforts of the two players. This leads us to say that if $d_i > 0$ and $a_i \geq 0$, then the probability of successful attack is given by:[10]

$$\frac{a_i}{a_i + d_i}. \tag{8}$$

Observe that in this formulation, if $d_i = 0$ then the probability of successful attack is 1, for every $a_i > 0$. We would like to rule out such extreme effectiveness of small attack resources. This motivates our formulation that the probability of successful attack for $d_i \geq 0$, $a_i \geq 0$, is:

$$\min\{a_i, \frac{a_i}{a_i + d_i}\}. \tag{9}$$

Figure 2 illustrates the behavior of our contest function as we vary the attack and defence allocation on a node. We summarize our assumptions on the outcomes of direct attack and defence as follows:
Assumption A.2: *Attack on node i is successful with probability* $\min\{a_i, \frac{a_i}{a_i+d_i}\}$. *Success of attack is independent across nodes.*

A key element of our model is the possibility that attacks spread through the network. It is natural to suppose that the likelihood of spread is related to level of defence resources allocated to different nodes. A simple formulation is that successful indirect attack obeys a threshold property: a successful attack on node i moves to a directly linked node j, if and only if $d_j < 1$. More generally, we shall say that the path between two nodes i and j is *weak* if and only if $d_k < 1$ for all nodes $k \neq i$ on this path.[11]

Assumption A.3: *Successful attack on node i spreads to node j if and only if there exists a weak path between i and j and $d_j < 1$.*

Defence plays the role of a protective 'firewall'. To see this, consider a dispersed network like the ring and suppose that the designer and the adversary both have a budget $1 < a = d = k < n/2$. Suppose that he allocates 1 unit each to protect k nodes. Then, under our assumptions (**A.2**) and (**A.3**), the adversary can eliminate $n - k$ nodes by simply allocating one unit of attack resource to one node in each interval between the protected nodes. The

[10] The key result in Skaperdas (1996) states that any 'winning' probabilities satisfying the properties of being increasing in own effort, decreasing in other's effort, symmetric across players, and homogeneous with respect to the efforts of the two players can be written in the form $\frac{a_i^\beta}{a_i^\beta+d_i^\beta}$, and $\frac{d_i^\beta}{a_i^\beta+d_i^\beta}$, where $\beta > 0$. The Tullock contest function is obtained by setting $\beta = 1$.

[11] Observe that if $d = 0$ then any path between two nodes is weak.

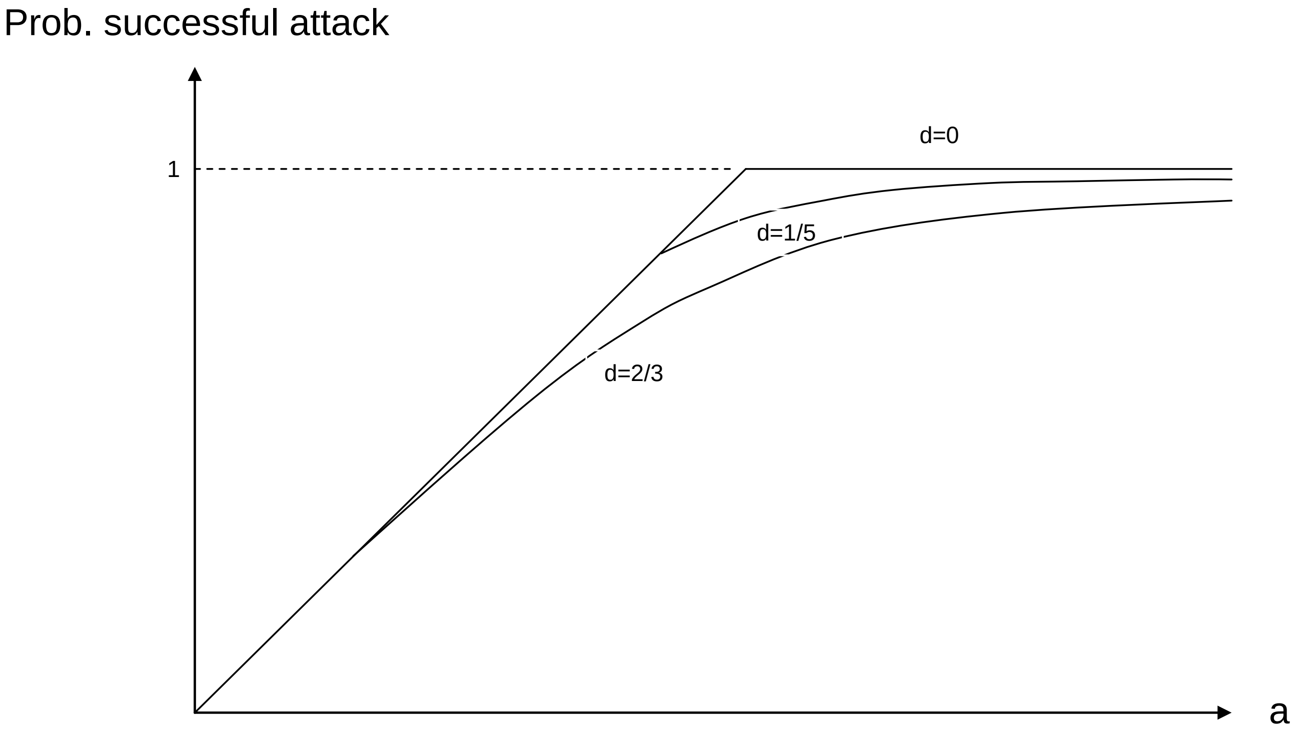


Figure 1: Probability of successful attack on a node

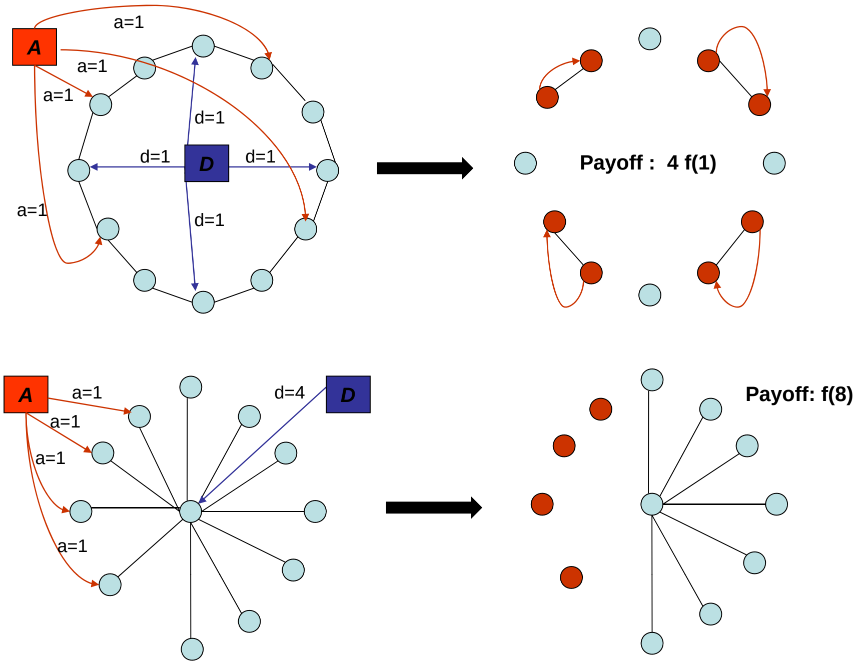


Figure 2: Attack & defence on ring and star; $n = 12, a = d = 4$.

maximum payoff the designer can earn is $kf(1)$. Figure 2 illustrates this network, the attack and defence strategies, the spread of attack and the residual surviving network. By contrast, consider the star network and let the designer allocate all resources to the central node. If adversary targets k unprotected nodes then the surviving network is a star with $n-k$ nodes. The payoff of the designer is $f(n-k)$. Figure 2 illustrates this. Given that $f(.)$ is increasing and convex, it follows that $f(n-k) > kf(1)$. Protecting the central node thus effectively blocks any spread of attacks.

The threshold structure of indirect attack means that spread of attacks exhibit a discontinuity at $d = 1$: for d_k smaller than but close to 1, the probability of successful indirect attack is 1, while at $d_k = 1$ there is probability 0 of successful indirect attack. In Appendix C we present and explore an alternative formulation of attack and contests which addresses this discontinuity. The analysis there clarifies the scope of our main results, Theorems 1 and 2.

We are now ready to define payoffs. First, $\mathcal{D}$ chooses strategy $(g, \mathbf{d})$. This strategy is observed by $\mathcal{A}$ who then chooses his strategy $\mathbf{a}$. For any given network g, a defence profile $\mathbf{d}$ and an attack profile $\mathbf{a}$, assumptions (**A.2**) and (**A.3**) define a probability distribution on the set of surviving sub-graphs $g' \subset g$. Denote this distribution by $P : \mathcal{D} \times \mathcal{A} \times \mathcal{G} \rightarrow \mathcal{G}$. So $P(g|\mathbf{a}, \mathbf{d}, g)$ is the probability of network g' surviving from the game of conflict with defence

d and attack **a** played out on the network g. We require

$$\begin{aligned} P(g|\mathbf{a}, \mathbf{d}, g) &\geq 0 \\ \sum_{g \in \mathcal{G}} P(g|\mathbf{a}, \mathbf{d}, g) &= 1 \end{aligned} \tag{10}$$

Let $C(g)$ be the partition into components in network g. The payoffs to the designer $\mathcal{D}$ when he chooses network g, and defence allocation **d** and adversary chooses allocation **a** are:

$$\sum_{g' \in \mathcal{G}} P(g'|\mathbf{a}, \mathbf{d}, g) \left[\sum_{C \in \mathcal{C}(g')} f(|C_k(g')|) \right] \tag{11}$$

Payoffs of the two players are assumed to sum to zero. We refer to the game just defined as the design-defence-attack ($\mathcal{DDA}$) game.

We will say that a network $g \in \mathcal{G}$ is *robust* if it maximizes the expected payoff of the designer faced with an intelligent adversary. More formally, a network g is *robust* if it is a sub-game perfect equilibrium outcome of the game.

3 The design and attack game

This section studies the case where the designer has no defence resources, i.e., $d = 0$. The only way in which the designer can protect the network is by separating nodes into distinct components. We show that it is optimal for the adversary to target at most one node in each component. A robust network consists of equal size components whose number grows and size falls as the attack budget of $\mathcal{A}$ increases. We then explore network robustness in the face of random attacks: networks consist of fewer components which are typically of unequal size. Thus understanding the nature of the adversary is critical for the design of networks.

We first observe that if $d = 0$ then a component $C \in C(g)$ survives if and only if *none* of its nodes is successfully directly attacked. The probability of this event is simply $\prod_{i \in C}(1 - a_i)$. Then, component additivity of payoffs implies that the expected payoff from network g to the designer $\mathcal{D}$ facing attack **a** is:

$$\sum_{C \in \mathcal{C}(g)} f(|C_k(g)|) \prod_{i \in C_k(g)} (1 - a_i) \tag{12}$$

Our first result characterizes optimal attack strategies and robust networks in this game. We show that it is optimal for $\mathcal{A}$ to target at most one node in each component. Moreover, since the adversary can observe the network, he will attack larger components first. So a robust network consists of equal size components. Their number grows and size falls as the attack budget of $\mathcal{A}$ increases. The following result summarizes these assertions.

Theorem 1 *Suppose (**A.1**)-(**A.3**) hold and $d = 0$. In equilibrium, the adversary targets at most one node in any component. If $a < n/2$ then a robust network contains at least $a + 1$ components of equal and maximal size, and at most one component which is smaller. If $a \geq n/2$, then the empty network is robust.*[12]

Proof: First, we establish that at most one node is attacked in a component. When $\mathcal{A}$ attacks two nodes with positive resources, there arises a state in which both nodes are eliminated: this is wasteful as elimination of one node is sufficient to remove the entire component.[13]

Second, there must be at least $a + 1$ components: if the number of components is fewer than $a + 1$, then $\mathcal{A}$ can set $a_i = 1$ for one node in each component and thereby ensure that $\mathcal{D}$ earns zero payoff. A network with $a + 1$ components on the other hand, guarantees $\mathcal{D}$ strictly positive payoff as at least one component survives any attack of $\mathcal{A}$ with some probability.

Third, we show that there are at least $a+1$ maximal size components. Suppose this is not the case and let component C_1 have maximal size. As part of his response, $\mathcal{A}$ must eliminate C_1. Next, form a new network g' from g in which C_1' is obtained from C_1 by isolating a single node, leaving the rest of the network unchanged. In g', either C_1' has maximal size, or at most $a - 1$ components have size strictly greater than it. Hence, without loss of generality, we may assume that C_1' is eliminated as part of the best response by $\mathcal{A}$. But then $\mathcal{D}$ does strictly better with g' than g since by doing so she saves the node she isolated. This contradicts the hypothesis that g is optimal.

Fourth, we show that at most one component has size strictly smaller than the maximal size $\bar{s}$. Suppose we can find two such components. $\mathcal{D}$ can then take a node from the smaller of the two components and place it in the larger component. The larger component still remains

[12] The empty network is uniquely robust for $n/2 \leq a < n$. At $a = n$, all networks yield 0 payoff to $\mathcal{D}$ and so all of them are robust.

[13] Observe that under (A.2), our formulation equates resources and probability of successful attack; this formulation is analytically convenient and is borrowed from Baccara and Bar-Isaac (2008). To see how this formulation may be motivated, suppose that the adversary allocates resources a_i to node i and that these resources translate into a probability of successful attack on node i via a technology summarized in the function $e(a_i) \in [0, 1]$. Our model then refers to the case in which $e(.)$ is linear. It is easy to see that the targeted attack result remains valid if the function $e(.)$ is convex, but may not hold if $e(.)$ is 'very' concave.

(weakly) smaller than the maximal size components, and it now follows from convexity of $f(.)$ that payoffs to $\mathcal{D}$ are strictly increased.

Finally, observe that if $a \geq n/2$ then $\mathcal{A}$ can always eliminate every component with 2 or more nodes. Hence it is optimal to only have components with single nodes, i.e., a network made of isolated nodes is robust.

■

By Theorem 1, a robust network consists of at least $a+1$ components. Loosely speaking then the number of components is weakly increasing in the budget of $\mathcal{A}$. The precise number of components in a robust network and how they change with the budget of $\mathcal{A}$ however hinges upon the convexity of $f(.)$.

Smaller components allow more nodes to survive in the face of attack, but reduce the payoff from any surviving node. When the returns from exchange increase, the latter effect strengthens. Intuitively, larger returns from exchange increase the reluctance of the designer to break up the network. So greater convexity of $f(.)$ induces robust networks which consist of fewer and larger components. The following example illustrates these effects.

Example 4 *Robust networks in a model of specialization and trade.*

Suppose payoffs are given by

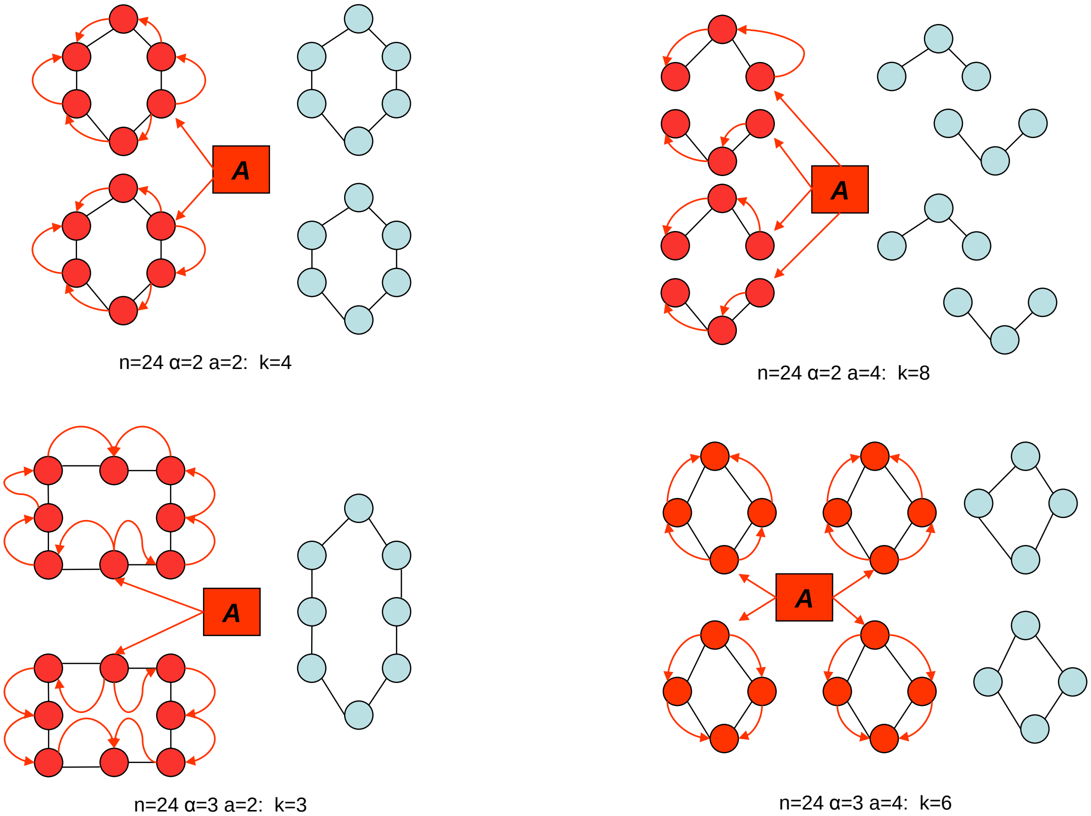
$$f(m) = \left[\frac{m}{n}\right]^{\alpha} \tag{13}$$

And, assume $a > 0$ and $d = 0$. What is the optimal number of components in a robust network and how does this vary with α and a? Our computations establish that: *if $a < n/2$ and $\frac{\alpha a}{\alpha - 1} \in \{a+1, .., n\}$ divides n, then the unique robust network consists of $\frac{\alpha a}{\alpha-1}$ equal size components. So, for $\alpha = 2$ a robust network has $2a$ equal size components whereas for $\alpha = 3$ this number falls to $3a/2$ components.*

Figure 3 illustrates robust networks for $n = 24$, $\alpha = 2, 3$ and for adversary budgets $a = 2, 4$.

Let $a < n/2$. Suppose $\frac{\alpha a}{\alpha-1} \in \{a+1, .., n\}$; observe that if $\alpha = 2$ then this is true for all a. We first show that a network with $\frac{\alpha a}{\alpha-1}$ equal size components is best among all networks with equal size components. We then show that any network in which one component has less than maximal size is dominated by some network with equal size components. The claim then follows from Theorem 1.

Consider a network with equal size components, and let s denote this size. Using arguments from Theorem 1 we know that, in any sub-game perfect equilibrium, $\mathcal{A}$ targets one node in a



n=24 α=2 a=2: k=4

n=24 α=2 a=4: k=8

n=24 α=3 a=2: k=3

n=24 α=3 a=4: k=6

Figure 3: Robust networks: $n = 24$, α=2,3 and $a = 2, 4$

components. So the payoff of $\mathcal{D}$ from choosing such a network is

$$f(s)(\frac{n}{s} - a) \tag{14}$$

It is easily checked that for $f(.)$ given by (13), this is maximized at $s = \frac{n(\alpha-1)}{a\alpha}$.

Next, consider a network with all but one component having maximal size $\overline{s}$, and one component of size $0 < s < \overline{s}$. Let $\beta = (n - s)/\overline{s}$ denote the number of maximal size components. Using arguments from Theorem 1, in any sub-game perfect equilibrium, the payoff of $\mathcal{D}$ from choosing this network is

$$f(\overline{s})(\beta - a) + f(s) \tag{15}$$

By convexity of $f(.)$, this payoff is less than

$$f(\overline{s})(\beta - a) + \frac{s}{\overline{s}} f(\overline{s}) \tag{16}$$

which, substituting for β and simplifying, can be written as

$$f(\overline{s})(\frac{n}{\overline{s}} - a) \tag{17}$$

But by the first step, for $f(.)$ given by (13) this last expression is less than payoffs attained with a network of $\frac{\alpha a}{\alpha - 1}$ equal size components. So a network with $\frac{\alpha a}{\alpha - 1}$ equal size components dominates any network in which one component has less than maximal size. By Theorem 1, it follows that a network with $\frac{\alpha a}{\alpha - 1}$ equal size components is robust.

$\triangle$

Theorem 1 and Example 4 suggest that as the adversary budget grows, the designer responds by splitting the network into smaller components. This is consistent with recent discussions in the media, with regard to the effects of larger government budgets in the fight against terrorism.

3.1 Strategic *vs.* random attack

We next examine robust networks in the face of uniform random attack. Uniform random attack refers to the case where the adversary assigns $q = a/n$ to every node in the network. This is a natural model for biological or physical attacks. It also serves as a benchmark which helps us understand the role of adversarial intelligence. We develop two general points: one, robustness in the face of random uniform attacks typically entails networks with components of unequal size, and two, the number of components in a robust network will be very different depending on whether the attack is random or strategic.

Example 5 *Random attack and unequal components*

Suppose that $n = 4$, $f(1) = 0.01$, $f(2) = 0.05$, $f(3) = 0.50$, $f(4) = 1.00$. Let us write down the payoffs as uniform random attack $q = a/4$ varies across $a \in \{0, 1, 2, 3, 4\}$. The payoffs to designer from the empty network g^e are $4[1 - q][0.01]$, the payoffs from a network with two equal components is $(1 - q)^2[0.10]$, the payoffs from a network with two unequal components is $(1 - q)^3[0.50] + (1 - q)[0.01]$ and the payoff from the connected network is $(1 - q)^4[1]$. It is now straightforward to check that the robust network at $a = 0, 1$ is connected, at $a = 2$ it contains two unequal components, at $a = 3$ it contains either two unequal components or four components. All networks yield value zero at $a = 4$ of course.

On the other hand, faced with strategic attack, the robust network at $a = 0$ is connected, at $a = 1$ is contains two equal components, at $a = 2, 3$, it contains four components. All networks yield value zero at $a = 4$.

Thus the nature of the adversary – strategic OR random – plays a crucial role in determining whether the components in a robust network are equal to unequal.

△

We now turn to the implication of random attack for the number of components in the robust network.

Example 6 *Number of components: random vs. strategic attack*

Suppose component payoffs are given by 3. We let $V(g, q)$ indicate the payoff of the designer from network g under uniform random attack $a_i = a/n$. Let g^k be a network with k equal size components. It is easily checked that

$$V(g^k, q) = \frac{1}{k}\,[1-q]^{\frac{n}{k}} \tag{18}$$

At $a_i = 0$, the optimal network is clearly connected. Since payoffs are continuous in a_i the connected network is also optimal for attack probability close to 0.

Now fix $a = 1$ and let n get large. For large n, the robust network is connected. The probability of task completion in a connected network is given by $(1-\frac{1}{n})^n \sim e^{-1} \sim \frac{1}{2.72} = .38$. Our computations in example 4, on the other hand, tell us that under strategic attack robust networks have two components and the probability of task completion is $1/4 = .25$, *irrespective of the number of nodes.* Thus the number of components in the robust networks differ – 1 *vs* 2 – and the probability of task completion is also very different – 0.38 *vs* 0.25 – when we compare random attack with strategic attack.

△

Thus, from a practical point of view, understanding the nature of the adversary is important. If the number of nodes n is large, a designer anticipating an attack budget of 1 and uniform random attack will design a connected network. In the face of an intelligent adversary such a network would yield a payoff of 0; by contrast, the designer could obtain a payoff of 1/4 in a network with two components of equal size!

4 Network design, defence and attack

We now examine the general problem of designing and defending a network which faces an intelligent adversary. Our main result says that, if the number of nodes is large relative to

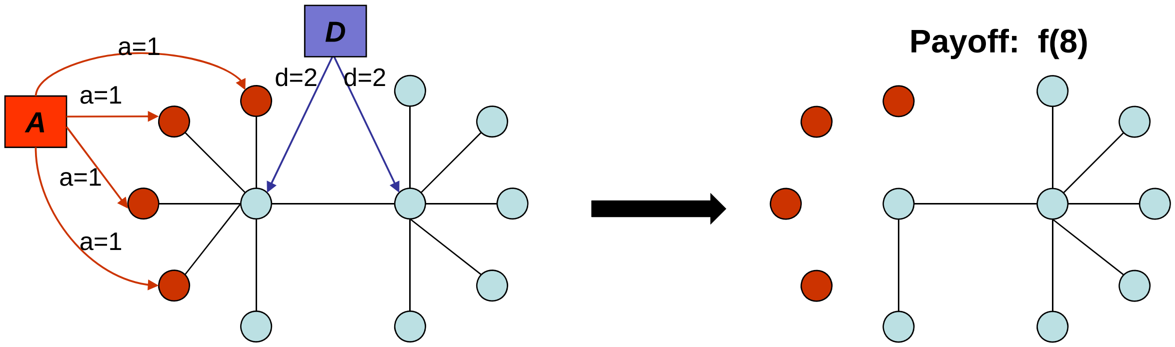


Figure 4: Core-periphery network: attacking the periphery nodes

the budget of the designer and adversary, the star network is robust. The designer allocates all his resources to defend, and the adversary allocates all his resources to attack, the center.

The shape of the returns function, $f(.)$ plays an important role in our analysis of defence and attack. Consider a (2 node) core-periphery network and let the designer allocate $d/2$ units to protect each core node. There are two simple strategies of attack: one, the adversary allocates $a/2$ units each to the two core nodes and two, the adversary allocates resources to a peripheral nodes. Given assumptions (**A.2**) and (**A.3**), the payoff to the designer under the former attack strategy is $1/4 + f(n/2)/2$, and under the latter attack strategy it is $f(n-a)$. The attractiveness of different attack strategies depends on the criticality of a nodes. Figures 4 and 4 illustrate. In Examples 1 and 2, any fixed number of nodes, k, has negligible impact on payoffs as the number of nodes gets large. So, for large n, $1/4+f(n/2)/2 < f(n-a)$ and it is optimal for the adversary to target the two core nodes. If, on the other hand, the function $f(.)$ has a threshold feature and $f(k) = 0$, for all $k < X$ (where $X > n - a$), and $f(k) = 1$ for $k \geq X$, then adversary clearly prefers to target a peripheral nodes and eliminate them for sure. These considerations motivate the following stronger assumption on payoffs.

Assumption A.1′: *The payoff to $\mathcal{D}$ from network g is given by (1), where $f(.)$ is increasing, convex, $f(n) = 1$ and $\lim_{n\to\infty} f(n-a)/f(n) = 1$ for all $a \in \mathbf{N}$.*

We can now state the main result of this section.

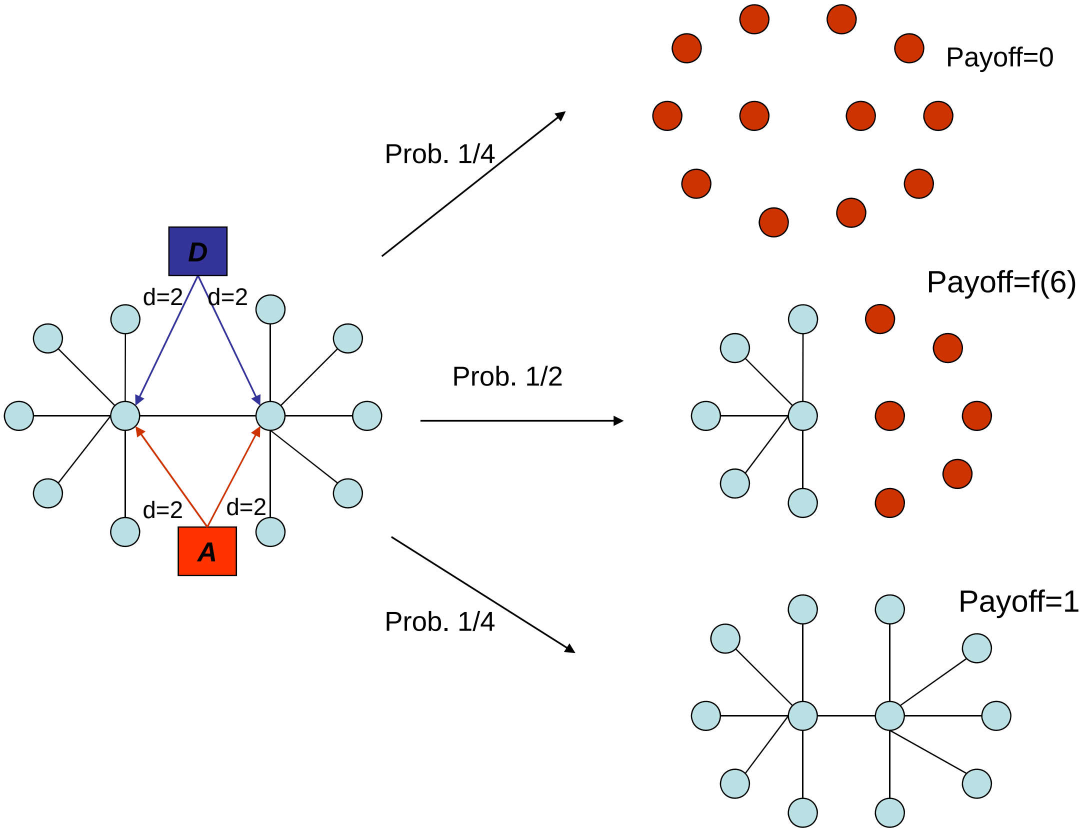


Figure 5: Core-periphery network: attacking the core nodes

Theorem 2 *Suppose* (***A.1***$'$)*-*(***A.3***) *hold. Fix budgets $a \geq 1$ and $d \geq 1$.*[14] *For n sufficiently large, the star network is robust in the class of connected networks. In equilibrium, each player allocates all his resources to the central node.*

The proof of this result is given in Appendix A. There are three steps in the proof of the theorem. First we show that given a star network in which all defence resources are allocated to the central node, $\mathcal{A}$ chooses to allocate c units of resource to the center, and concentrates all remaining resources on $a - c$ nodes in the periphery. Spreading resources across more nodes implies a spread in the distribution of surviving nodes. But observe that by (**A.3**) a center-protected star ensures connectedness of all surviving nodes. Thus, by convexity of $f(.)$, spreading resources across many nodes reduces the expected payoff of $\mathcal{A}$.

Second, we show that given a star network $\mathcal{D}$ can ensure himself a payoff of $d/(d + a)$. Suppose $\mathcal{D}$ allocates all his resources to protect the central node. Consider the incentives of $\mathcal{A}$. If $\mathcal{A}$ allocates k units of resource to peripheral nodes, then he eliminates k of them for sure. But this allocation of k units away from the central node lowers the probability of

[14] The case of $a = 0$ is uninteresting as (**A.1**) implies that the robust network is connected and nothing more can be said in the absence of attack about the network architecture. The case of $d = 0$ has already been covered in Theorem 1 above.

eliminating it correspondingly. Under assumption ($\mathbf{A.1'}$), for large enough n, the marginal value of eliminating a single node is very small compared to the value from a marginal increase in the chances of eliminating the central node. Thus $\mathcal{A}$ will find it optimal to allocate all resources to the central node. This also shows that, with a star network, $\mathcal{D}$ can ensure himself a payoff of $d/(d+a)$.

Third, we show that, given any network other than a star and any allocation of defence resources, there exists an attack by $\mathcal{A}$ which yields $\mathcal{D}$ payoff less than $d/(d+a)$. To see why this must be true, consider a core-periphery network with 2 core nodes and suppose $a = d = 4$. Suppose $\mathcal{D}$ defends the 2 core nodes in this network with equal resources. $\mathcal{A}$ can always allocate 2 units to each defended node. Now it can be checked that the distribution of surviving nodes in the star network with protected center is a mean-preserving spread of the distribution of surviving nodes with 2 nodes defended in a (2 node) core-periphery network. Figures 4 and 4 illustrate these outcomes. Moreover, the surviving nodes in a center-protected star constitute a component. Since $f(.)$ is increasing and convex, the expected payoff of $\mathcal{D}$ with 2 nodes defended is strictly smaller than the expected payoff under the star network in which $\mathcal{D}$ allocates all resources to defend the central node. The arguments in our proof show that this intuition can be generalized to cover arbitrary allocations and all possible connected networks.

The empirical work highlights the importance of hub nodes in real world networks (see e.g., Barabasi (1999)). In an influential paper, using simulations, Albert, Jeong and Barabasi (2000) highlight the vulnerability of hub-spoke architectures to intelligent adversaries: successful attack on only a few nodes is enough for the disintegration of the network at large. Our result highlights the other side of hub-spoke architectures: successful defence of only a few nodes contains the spread of attacks through the network. This pressure prevails and our analysis demonstrates that in an interesting class of economic environments – where payoffs from exchange are increasing and convex – networks with hubs are robust. Thus our work provides a theoretical account for the salience of hubs in these networks.

We have so far focused on the case where budgets are small relative to the number of nodes. While we believe this is the natural scenario in large networked systems, for the sake of completeness we now turn to the case where budgets are large. We first study the case where only one of the players has large budget and then examine the case where both players have large budgets. In the rest of this section we restrict attention to the connections model; for expositional simplicity, we also suppose that resource allocations across nodes only take integer values.

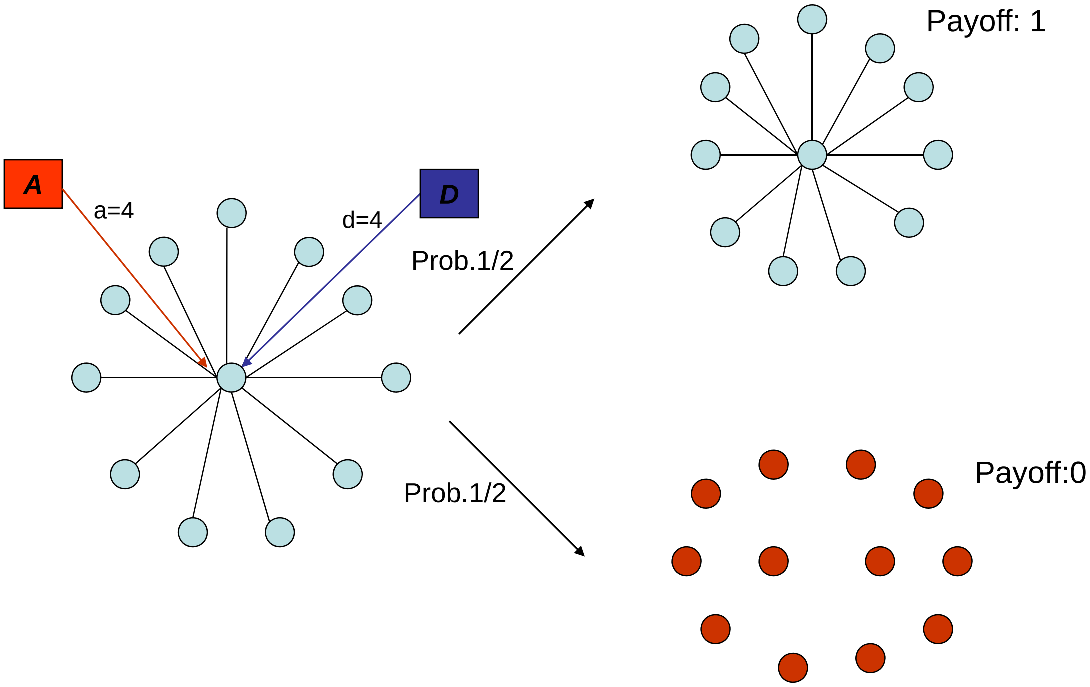


Figure 6: Attack & defence on a star

Proposition 1 *Consider the payoffs in the connections model. Assume that (****A.2****)and(****A.3****) hold and that allocations of resources on nodes taken integer values only.* (i). *If $d = 1$ then, for all a, the star network is robust within the class of connected networks.* (ii) *If $a = 1$ then, for $d < n$, the star network is robust within the class of connected networks. For $d = n$ the complete and ring networks are robust, and payoff dominate the star network.*

The proof of this proposition is given in Appendix A. The argument for the first part is straightforward: in a star network suppose $\mathcal{D}$ protects the central node. If adversary optimally allocates t units to the central node the payoff to $\mathcal{D}$ is $f(n-a+t)/1+t$. Next take any connected network $g^{'}$ and suppose node, $\alpha^{'}$ is defended. If adversary allocates t to node α', and 1 unit each to $a-t$ other nodes, the payoff of designer is at most $f(n-a+t)/1+t$ (as attacks on some of the nodes will spread to unprotected neighboring nodes). Since the network g was arbitrary and the adversary's strategy is feasible, the star is robust.

The argument for the second part, $a = 1$, $0 \leq d < n$, builds on two observations. Fix some network g and some (possibly dispersed) defence allocation $\mathbf{d} = (d_1, d_2, .., d_n)$. The first observation is that the probability of successful attack on a protected node i, $1/(d_i + 1)$, is larger than the probability of successful attack on the central node in the star network, $1/(d + 1)$. On the other hand the set of nodes indirectly attacked will be smaller than in the case of the case of the center-protected star (where all n nodes are eliminated). The second observation is about the number of indirectly affected nodes in the dispersed defence case. We show that there always exists a node i such that successful attack on this node exposes n_i other unprotected nodes to indirect attack where $n_i \geq n(d_i/d)$. So the maximum payoff to the designer from such a network and dispersed defence is

$$\frac{d_j}{d_j+1}f(n) + \frac{1}{d_j+1}f(n - \frac{d_j}{d}n) \tag{19}$$

Indeed, we can write the difference in payoff between the dispersed protected network and the center protected star as:

$$\frac{d_j}{d_j+1}f(n) + \frac{1}{d_j+1}f(n - \frac{d_j}{d}n) - \frac{d}{d+1}f(n) \tag{20}$$

So the attractiveness of network g and a possibly dispersed attack $\mathbf{d}$ relative to a center-protected star will depend on the payoff function $f(.)$. We show that in the connections model this difference is always non-positive. Hence the star with protected center is robust.

Next we explain why the star is not robust for $d = n$. If the designer protects every node and the adversary attacks the central node, the payoff to the designer is

$$\frac{n-1}{2}f(1) + \frac{1}{2}f(n) \tag{21}$$

By contrast, in the ring network, with all nodes protected, the payoff to the the designer is

$$\frac{1}{2}f(n-1) + \frac{1}{2}f(n) \tag{22}$$

By convexity of $f(.)$, this is clearly larger. If alternatively $\mathcal{D}$ leaves one node unprotected, and $\mathcal{A}$ attacks this node, then payoff to the designer is at most $f(n-1)$ which is clearly smaller than the payoff from the ring.

We turn finally to the case where he budgets of *both* players are large relative to the number of nodes. We have been unable to characterize robust networks for general a, d and n. The following example illustrates the complexity of the problem in a model with four nodes.

Example 7 *Attack and defence with large budgets*

Consider the payoffs in the connections model. Suppose $n = 4$, assumptions (**A.2**) and (**A.3**) hold and allocations take on integer values only. Let g^{s} indicate a star network, g^{c}

denote the complete network, g^{rd} denote the ring network with one diagonal link, g^{r} the ring network, g^{l} the line network, and g^{sp} the star network with a link between two of the periphery nodes.

The matrix shown below summarizes our computations (the details of the computations are presented in the Appendix A.)

	a=1	a=2	a=3	a=4
d=1	g^{s}	g^{s}	g^{s}	g^{s}
d=2	$\{g^{\mathrm{s}}, g^{\mathrm{l}}\}$	$\{g^{\mathrm{s}}, g^{\mathrm{sp}}, g^{\mathrm{rd}}, g^{\mathrm{l}}\}$	$\{g^{\mathrm{s}}, g^{\mathrm{sp}}, g^{\mathrm{rd}}, g^{\mathrm{l}}\}$	$\{g^{\mathrm{s}}, g^{\mathrm{sp}}, g^{\mathrm{rd}}, g^{\mathrm{l}}\}$
d=3	$\{g^{\mathrm{s}}, g^{\mathrm{sp}}, g^{\mathrm{r}}, g^{\mathrm{l}}\}$	$\{g^{\mathrm{sp}}, g^{\mathrm{rd}}, g^{c}\}$	$\{g^{\mathrm{sp}}, g^{\mathrm{rd}}, g^{c}\}$	$\{g^{\mathrm{sp}}, g^{\mathrm{rd}}, g^{c}\}$
d=4	$\{g^{\mathrm{r}}, g^{\mathrm{rd}}, g^{c}\}$	g^{c}	g^{c}	g^{c}

To understand the effects of larger budgets relative to the number of nodes consider $d = 4$ and $a = 1$, and suppose $\mathcal{D}$ chooses a star network. If $\mathcal{D}$ protects less than four nodes, $\mathcal{A}$ can eliminate one node for sure. So the maximal payoff of $\mathcal{D}$ with less than four nodes protected is $f(3) = 9/16$. Next, suppose $\mathcal{D}$ allocates one unit of resource on every node in the network. Clearly, the optimal response of $\mathcal{A}$ is to attack the central node. The resulting payoff of $\mathcal{D}$ is $3f(1)/2 + f(4)/2 = 19/32$. So in a star network spreading resources is optimal for $\mathcal{D}$. But observe that in this case additional links between periphery nodes augment the payoffs that $\mathcal{D}$ can guarantee himself, by ensuring connectedness of these nodes in the event that attack on the center is successful. If $\mathcal{D}$ protects all nodes in the complete network for example, then his minimum payoff is $f(3)/2 + f(4)/2 = 25/32$. This shows that for n=4, d=4, and a=1, the complete network strictly dominates a star. Figures 4 and 4 illustrate this example.

$\triangle$

We conclude this section with two remarks.

One, we observe that Theorem 2 and the other results in this section all maintain the assumption that the network is connected. Example 4, in the previous section, clarifies the key role of convexity of the returns function $f(.)$ in shaping the number of components in a robust network. If $f(.)$ is sufficiently convex then the designer will choose to have a connected network and Theorem 2 and the subsequent examples defines the architecture of such a connected network and predicts the nature of attack and defence strategy. If, on the other hand, $f(.)$ is close to being linear then multiple components will be better. Proposition 2 in Appendix B develops a set of sufficient conditions on the returns function $f(.)$ which ensure connectedness of the robust network.

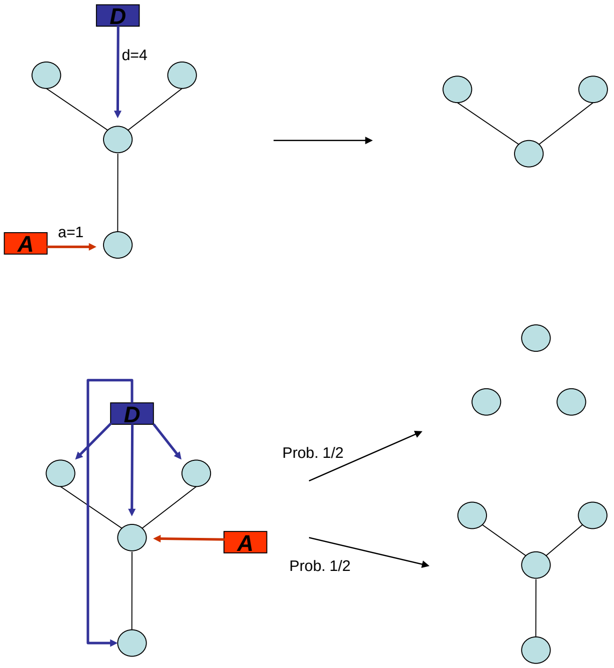


Figure 7: Star network, n=4 d=4 a=1

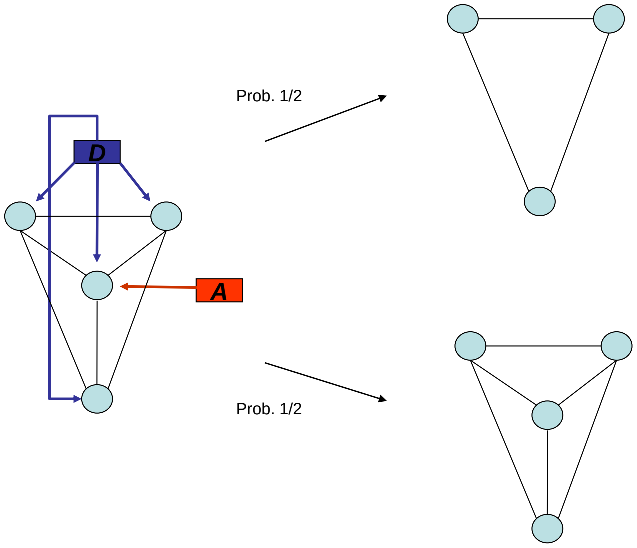


Figure 8: Complete network, n=4 d=4 a=1

Two, we have assumed that the adversary observes the allocation of defence resources made by the designer prior to making his own choices. In some applications, the designer and adversary may have an interest as well as an ability to conceal their attack and defence strategies. It is possible to show that if budgets are small relative to the number of nodes, the star is robust. Proposition 3 in Appendix B presents this result in a setting with different orders of move and allows for multiple components.

5 Conclusion

This paper explores the design and defence of networks which face an intelligent adversary.

We first study a game in which a designer constructs a network while an adversary attacks nodes in this network. Optimal attack involves targeting only a few nodes and ignoring the rest. In response, robust networks consist of equal size components. Their number grows and size falls as the attack budget of the adversary increases. Fewer and less complex tasks are being performed by the network, and this effect strengthens the smaller the returns from exchange in the network. For fixed budget of attack, the number of components is higher and performance poorer as compared to the case of uniform random attack.

We then extend the strategic options of the designer: he can now choose a network and allocate resources to defend nodes. This defines a game of design, defence and attack. Our second result is that if the defence and attack resources are small relative to the number of nodes, then a star network is robust. In equilibrium, both players allocate all resources to the central node. On the other hand, if the budgets of both the designer and adversary are large then denser networks with dispersed defence allocations are robust.

This paper restricts attention to the case where a designer, or social planner, chooses the network and the protection strategy. The related problem, in which nodes choose their own pattern of interaction as well as protection, is an important avenue for research. The companion papers, Goyal and Vigier (2009a, 2009b) study decentralized linking and protection.

6 Appendix A

The following Lemma is useful in the proof of Theorem 2.

Lemma 1 *Let $\{I_1, .., I_k\}$ denote a set of i.i.d. Bernoulli random variables with $P(I_i = 1) = \delta$*

and $P(I_i = 0) = 1 - \delta$, for all i. Then $(n_1 + .. + n_k)I_1$ is a mean-preserving spread of $n_1I_1 + .. + n_kI_k$.

Proof. Suppose, without loss generality, that $n_1 \leq .. \leq n_k$. We prove the result by induction on k.

Suppose $k = 2$. Let F and G denote the cumulative distribution functions of $(n_1 + n_2)I_1$ and $n_1I_1 + n_2I_2$, respectively. Define $1 - \delta = \alpha$. Then

$$F(x) = \begin{cases} \alpha & \text{if } 0 \leq x < n \\ 1 & \text{if } x = n \end{cases} \tag{23}$$

and

$$G(x) = \begin{cases} \alpha^2 & \text{if } 0 \leq x < n_1 \\ \alpha & \text{if } n_1 \leq x < n_2 \\ 1 - \delta^2 & \text{if } n_2 \leq x < n \\ 1 & \text{if } x = n \end{cases} \tag{24}$$

So, using Theorem 1 in Rothschild and Stiglitz (1970), $(n_1 + n_2)I_1$ is a mean-preserving spread (MPS) of $n_1I_1 + n_2I_2$ if and only if

$$\alpha - \alpha^2 = 1 - \delta^2 - \alpha \tag{25}$$

or, substituting for δ

$$\alpha - \alpha^2 = 2\alpha - \alpha^2 - \alpha \tag{26}$$

So the result holds for $k = 2$.

Next, suppose the result holds up to $k \geq 2$.

We first show that if Y is a MPS of X then, for any random variable Q independent of X and Y, $Y + Q$ is a MPS of $X + Q$. This is equivalent to show that there exists a random variable Z such that $Y + Q =_d X + Q + Z$ and $E[Z|X + Q] = 0$ (where $=_d$ means 'have same distribution'. See Rothschild and Stiglitz (1970), Theorem 2).

Since Y is a MPS of X, we can find a r.v. Z such that $Y =_d X + Z$ and $E[Z|X] = 0$. Given that Q is independent of X and Y, we then have $Y + Q =_d X + Z + Q$. In addition

$$\begin{aligned} E[Z|X+Q] &= E[E[Z|X,Q]|X+Q] \\ &= E[E[Z|X]|X+Q] \\ &= E[0|X+Q] \\ &= 0 \end{aligned} \tag{27}$$

where the first equality uses the tower property of conditional expectations, the second equality is immediate by independence of Z and Q, and the third equality is by definition of Z. Thus, as claimed, $Y + Q$ is a MPS of $X + Q$.

But then setting $X = n_1 I_1 + n_2 I_2$, $Y = (n_1 + n_2) I_1$, $Q = n_3 I_3$, and using the result for $k = 2$ shows that $(n_1 + n_2) I_1 + n_3 I_3 + .. + n_{k+1} I_{k+1}$ is a MPS of $n_1 I_1 + n_2 I_2 + n_3 I_3 .. + n_{k+1} I_{k+1}$. Hence, by induction, $(n_1 + n_2 + .. + n_{k+1}) I_1$ is a MPS of $n_1 I_1 + n_2 I_2 + n_3 I_3 .. + n_{k+1} I_{k+1}$ and the proof is complete.

■

Proof of Theorem 2: First we show that with a center-protected star there exists $c \geq 0$ such that the optimal response of $\mathcal{A}$ consists in allocating c units of resource to the central node and exactly 1 unit of resource to $a - c$ periphery nodes. Two, we show that with a center-protected star $\mathcal{D}$ guarantees himself payoff of $d/(a + d)$. Last we show that, for any network with defence, there is an attack by $\mathcal{A}$ which yields $\mathcal{D}$ payoff less than $d/(a + d)$.

Step 1: Consider a star network with $n > a + 1$, and suppose $\mathcal{D}$ allocates all his resources to protecting the central node. We will show that if $\mathcal{A}$ allocates c units of resource to the central node, then his best allocation of remaining resources consists in targeting $a - c$ nodes with exactly 1 unit of resource each.

Suppose we can find two periphery nodes, i_1 and i_2 say, such that $0 < a_{i_1} \leq a_{i_2} < 1$. We will show that $\mathcal{A}$ obtains strictly higher payoff in this sub-game if he transfers a small amount of resources from i_1 to i_2.

Let M denote the set of nodes other than i_1 and i_2 on which attack is unsuccessful (note, this is a random variable), and let $m = |M|$. In the event that the central node belongs to $N \backslash M$, all nodes are removed and the payoff of $\mathcal{D}$ is trivially zero. If on the other the central node belongs to M then, by (**A.3**), no attack spreads through the network. In addition, observe that the structure of the network ensures connectedness of M. So the payoff of $\mathcal{D}$ is

$f(m)$ at least.

Given M, we next examine the impact of attack on i_1 and i_2. Since success of attack is independent across nodes, the payoff of $\mathcal{D}$ can be written as

$$(1-a_{i_1})(1-a_{i_2})f(m+2)+[a_{i_1}(1-a_{i_2})+a_{i_2}(1-a_{i_1})]f(m+1)+a_{i_1}a_{i_2}f(m) \tag{28}$$

Letting $\overline{a}=a_{i_1}+a_{i_2}$, this becomes

$$(1-\overline{a})f(m+2)+\overline{a}f(m+1)+a_{i_1}(1-a_{i_1})[f(m+2)-2f(m+1)+f(m)] \tag{29}$$

By convexity of $f(.)$, $f(m+2)-2f(m+1)+f(m)>0$. So the former expression is increasing in a_{i_1} (recall, $a_{i_1}\leq\overline{a}/2$), and $\mathcal{A}$ obtains strictly higher payoff if he transfers a small amount of resources from i_1 to i_2.

Since the argument above was for arbitrary realization of the set M, transferring resources also ensures $\mathcal{A}$ strictly higher payoff in the overall sub-game. It is then immediate that if $\mathcal{A}$ allocates c units of resource to the central node, his best allocation of remaining resources consists in targeting $a-c$ nodes with exactly 1 unit of resource each (this is feasible since we assumed at the outset $n>a+1$).

Step 2: We show that, with a center-protected star, $\mathcal{D}$ guarantees himself payoff $d/(a+d)$.

First, notice that since $f(.)$ is increasing and convex, for all $c<a$

$$f(n-a+c)\leq f(n-a+c)+(a-c)f'(n-a+c)\leq 1 \tag{30}$$

So, by $(\mathbf{A.1'})$:

$$\lim_{n\to\infty} f'(n-a+c)=0,\ \forall\, c<a \tag{31}$$

Next, consider a center-protected star. By step 1, we can restrict attention to attacks which consist in allocating $c\geq 0$ units of resource to the central node and exactly 1 unit of resource to $a-c$ periphery nodes. Given $d\geq 1$, we have $\min\{c,\frac{c}{c+d}\}=\frac{c}{c+d}$, for all $c\geq 0$. So the payoff of $\mathcal{D}$ under such an attack is

$$\frac{d}{c+d}f(n-a+c) \tag{32}$$

Differentiating with respect to c yields

$$-\frac{d}{(c+d)^2}f(n-a+c)+\frac{d}{(c+d)}f'(n-a+c) \tag{33}$$

By (**A.1′**), and (31), this expression tends to $-\frac{d}{(c+d)^2}f(n)<0$. So, for n sufficiently large, the best response of $\mathcal{A}$ to a center-protected star consists in allocating all resources to the central node. The resulting payoff of $\mathcal{D}$ is $d/(a+d)$.

Step 3: We show that, for any network with defence, there is an attack by $\mathcal{A}$ which yields $\mathcal{D}$ payoff less than $d/(a+d)$.

Consider an arbitrary (connected) network g, and arbitrary allocation of defence resources across nodes in this network. Let k denote the number of nodes for which $d_i \geq 1$ in this allocation.

If $k=0$, choose a node at random, j say, and suppose $\mathcal{A}$ sets $a_j = a$. Given that $d \geq 1$, it follows that $d_j < d$. And since $\min\{a, \frac{a}{a+d}\}$ is non-increasing in d, we have

$$\min\{a_j, \frac{a_j}{a_j+d_j}\} \geq \min\{a_j, \frac{a_j}{a_j+d}\} = \frac{a}{a+d} \tag{34}$$

By (**A.3**), the resulting payoff of $\mathcal{D}$ is therefore at most $d/(a+d)$.

Next, suppose $k \geq 1$ and label the nodes for which $d_i \geq 1$ from 1 to k. Let O denote the set of remaining nodes. So a node j belongs to O if and only if $d_j < 1$. Given this network with defence, suppose $\mathcal{A}$ sets $a_i = \frac{a}{d}d_i$, $i \in \{1,..,k\}$, and burns any remaining resources he may be left with. Notice that, for all i in $\{1,..,k\}$:

$$\min\{a_i, \frac{a_i}{a_i+d_i}\} = \frac{a_i}{a_i+d_i} = \frac{a}{a+d} \tag{35}$$

Define the sequence of sets N_i, $i \in \{1,..,k\}$ recursively. Let N_1 denote the set of nodes comprising node 1 and all nodes in O that can be reached from 1 through a path itself in O. N_2 the set of nodes comprising node 2 and all nodes in $O \backslash N_1$ that can be reached from 2 through a path itself in $O \backslash N_1$, and so on until N_k. Let $n_i = |N_i|$, $i \in \{1,..,k\}$.

Since g is connected, each node in O can be reached through a path in O from at least one node $i \in \{1,..,k\}$ (and so, in particular, $n_1+..+n_k = n$). If each node in O can be reached through a path in O from *exactly* one node $i \in \{1,..,k\}$, call this case 1. Otherwise, call this case 2.

In what follows, we let $\{I_1, .., I_k\}$ denote a set of independent Bernoulli random variables such that $P(I_i = 1) = \frac{d}{a+d}$, for all $i \in \{1, .., k\}$.

Consider case 1 first. Observe that, given the profile of attack, it follows from (**A.3**) that nodes in N_i survive if and only if i survives, for all $i \in \{1, .., k\}$. So the total number of surviving nodes has same distribution as $n_1 I_1 + .. + n_k I_k$. Given increasing and convex $f(.)$, the expected payoff of $\mathcal{D}$ is thus at most $E[f(n_1 I_1 + .. + n_k I_k)]$. Convexity of $f(.)$ and Lemma 1 then show (see e.g., Rothschild and Stiglitz (1970))

$$E[f(n_1 I_1 + .. + n_k I_k)] \leq E[f((n_1 + .. + n_k) I_1)] = \frac{d}{a+d} f(n) \tag{36}$$

Hence, by step 2, the resulting payoff of $\mathcal{D}$ is less than he can guarantee himself with a star.

Next, consider case 2. For all $i \in \{1, .., k\}$, successful attack on node i spreads to all other nodes in N_i as well as possibly some nodes in N_j, $j \neq i$. So the distribution of the total number of surviving nodes is first order stochastically dominated by that of $n_1 I_1 + .. + n_k I_k$. The expected payoff of $\mathcal{D}$ is thus at most $E[f(n_1 I_1 + .. + n_k I_k)]$, since $f(.)$ is increasing and convex. Again, by (36), this shows that the resulting payoff of $\mathcal{D}$ is less than he can guarantee himself with a star.

We have thus found an attack by $\mathcal{A}$ which, given any network with defence, yields $\mathcal{D}$ lower payoff than that he can guarantee himself with a star. Hence, the star is robust.

■

Proof of Proposition 1: (i). Consider a star network, and suppose $\mathcal{D}$ allocates his unit of resource to protect the central node. Let α denote the central node, and t denote the optimal units of resource which $\mathcal{A}$ allocates on α. The resulting payoff of $\mathcal{D}$ is thus

$$\frac{1}{1+t} f(n - a + t) \tag{37}$$

Next, consider an arbitrary connected network $g^{'}$, in which a single node, $\alpha^{'}$ say, is defended. Consider the following attack by $\mathcal{A}:$ allocate t units of resource to attack node $\alpha^{'}$, and 1 unit of resource to $a - t$ other nodes. Clearly, the resulting payoff of $\mathcal{D}$ is at most (37). We have thus found an attack by $\mathcal{A}$ which, given any network with allocation of a unit resource for defence, yields $\mathcal{D}$ lower payoff than that he can guarantee himself with a star. Hence, the star is robust.

(ii) If $d = 0$, the result is trivial. So suppose $1 \leq d < n$. Consider a star network, and suppose $\mathcal{D}$ allocates all his resources to protect the central node. Either the optimal response of $\mathcal{A}$ is to target a node in the periphery (call this case 1), or it is to allocate his unit resource to attack the central node (case 2).

Case 1: In this case the resulting payoff of $\mathcal{D}$ is $f(n-1)$. Since $d < n$, for any connected network and allocation of defence resources in it, there exists an undefended node. If $\mathcal{A}$ then allocates his unit resource on this node, the resulting payoff of $\mathcal{D}$ is at most $f(n-1)$. So the star is robust in this case.

Case 2: In this case the resulting payoff of $\mathcal{D}$ is $\frac{d}{d+1}f(n)$. For comparison, consider an arbitrary connected network, g say, and allocation of defence resources $\mathbf{d}$ on it. As in Theorem 2, label the nodes for which $d_i \geq 1$ from 1 to k. Let O denote the set of remaining nodes, and define the sequence of sets N_i recursively in the following way, $i \in \{1, .., k\}$. Let N_1 denote the set of nodes comprising node 1 and all nodes in O that can be reached from 1 through a path consisting of nodes which are in O. N_2 the set of nodes comprising node 2 and all nodes in $O \backslash N_1$ that can be reached from 2 through a path itself in $O \backslash N_1$, and so on until N_k. Let $n_i = |N_i|$, $i \in \{1, .., k\}$.

Since g is connected, each node in O can be reached through a path in O from at least one node $i \in \{1, .., k\}$. Hence $n_1 + .. + n_k = n$, and it follows that $n_i \geq \frac{d_i}{d}n$ for at least one $i \in \{1, .., k\}$. Let j denote one such node, i.e. $n_j \geq \frac{d_j}{d}n$.

Next, suppose $\mathcal{A}$ sets $a_j = 1$. The resulting payoff of $\mathcal{D}$ is at most

$$\frac{d_j}{d_j+1}f(n) + \frac{1}{d_j+1}f(n-n_j) \tag{38}$$

And, given $n_j \geq \frac{d_j}{d}n$, this in turn is at most

$$\frac{d_j}{d_j+1}f(n) + \frac{1}{d_j+1}f(n-\frac{d_j}{d}n) \tag{39}$$

Subtracting from this expression the payoff obtained by $\mathcal{D}$ in a star with protected center yields

$$\frac{d_j}{d_j+1}f(n) + \frac{1}{d_j+1}f(n-\frac{d_j}{d}n) - \frac{d}{d+1}f(n) \tag{40}$$

Substituting for $f(.)$ using payoffs given by (3) yields:

$$\left(\frac{d_j}{d_j+1} - \frac{d}{d+1}\right) + \frac{1}{d_j+1}\left(1 - \frac{d_j}{d}\right)^2 \tag{41}$$

And, multiplying by $(d_j+1)(d+1)d^2$

$$\left[d_j(d+1)d^2 - (d_j+1)d^3\right] + (d+1)(d-d_j)^2 \tag{42}$$

This simplifies to

$$(d-d_j)^2 + dd_j(d_j - d) \tag{43}$$

And further to:

$$(d-d_j)(d-d_j-dd_j) \tag{44}$$

Since $d_j \geq 0$, this expression is in turn less than

$$d(d-d_j)(1-d_j) \tag{45}$$

Either $d_j = 1$ and this expression equals 0, or $d_j > 1$ and it is negative. We have thus found an attack by $\mathcal{A}$ which yields $\mathcal{D}$ lower payoff than the star with protected center. Since the network g and the allocation $\mathbf{d}$ were arbitrary, the star network with protected center is robust.

Finally, consider $d = n$. In the complete and ring networks, if $\mathcal{D}$ allocates one unit of resource on every node his resulting payoff is

$$\frac{1}{2}f(n-1) + \frac{1}{2}f(n) \tag{46}$$

Consider next an arbitrary network, (say) g. If all nodes nodes are protected the maximum payoff to $\mathcal{D}$ is clearly given by (46). Alternatively, at least one node is not protected. In this case, $\mathcal{A}$ may choose to attack this node. This leaves $\mathcal{D}$ with payoff at most $f(n-1)$, which is less than (46). So both the complete and ring networks are robust.

$\blacksquare$

Example 7: The cases in which $d = 1$ or $a = 1$ follow from Proposition 1. Next, notice that it is a weakly dominant strategy for $\mathcal{D}$ to: (i) have a link between any two protected nodes; (ii) not have links between unprotected nodes. With these observations in mind, we

investigate below all remaining cases. Throughout, we refer to the resulting payoffs of $\mathcal{D}$.
d=2, a=2: Note that in this case $\mathcal{A}$ can always induce payoff of 1/4 by targeting two unprotected nodes. Moreover the optimal strategy of $\mathcal{A}$ with a center-protected star (CP star) consists in targeting two peripheral nodes. This induces payoff of 1/4 and so the CP star is optimal.
d=2, a=3: The optimal strategy of $\mathcal{A}$ with a CP star consists in targeting three peripheral nodes. The payoff induced is 1/16. With two defended nodes, the optimal strategy of $\mathcal{A}$ consists in targeting two unprotected nodes and one defended node. The payoff induced is $5/32 > 1/16$. Two defended nodes is thus optimal.
d=2, a=4: The optimal strategy of $\mathcal{A}$ with a CP star consists in targeting all nodes. The payoff induced is 1/32. With two defended nodes, the optimal strategy of $\mathcal{A}$ again consists in targeting all nodes. The payoff induced however is $3/32 > 1/32$. So two defended nodes is optimal.
d=3, a=2: The optimal strategy of $\mathcal{A}$ with a CP star consists in targeting peripheral nodes. The payoff induced is 1/4. With two defended nodes $\mathcal{A}$ targets two unprotected nodes and payoff is again 1/4. With three protected nodes, the optimal strategy of $\mathcal{A}$ consists in targeting the unprotected node and one other node. The payoff induced is $13/32 > 1/4$. So three defended nodes is optimal.
d=3, a=3: The optimal strategy of $\mathcal{A}$ with a CP star consists in targeting peripheral nodes. The payoff induced is 1/16. With two protected nodes, $\mathcal{A}$ targets two unprotected nodes and the least protected node. The payoff induced is 5/32. With three protected nodes, the optimal strategy of $\mathcal{A}$ consists in targeting the unprotected node and two other nodes. The payoff induced is 9/32. Three defended nodes is therefore optimal.
d=3, a=4: The optimal strategy of $\mathcal{A}$ with a CP star consists in targeting all nodes. The payoff induced is 3/64. With two protected nodes, the optimal strategy of $\mathcal{A}$ consists in targeting all nodes. The payoff induced is 11/96. With three protected nodes, $\mathcal{A}$ again targets all nodes. The payoff induced is 3/16. Thus three protected nodes is optimal.
d=4, a=1: If $\mathcal{D}$ protects less than four nodes, his maximum payoff in a sub-game perfect equilibrium is $f(n-1)$. If he protects four nodes on the other hand, $\mathcal{D}$ can guarantee himself payoff $f(n-1)/2 + f(n)/2$. This is greater than $f(n-1)$. So four protected nodes is optimal.
d=4, a=2: The optimal strategy of $\mathcal{A}$ with a CP star consists in targeting peripheral nodes. The payoff induced is 1/4. The same payoff results with two protected nodes. With three protected nodes, $\mathcal{A}$ targets the unprotected node and one of the least defended node. The payoff induced is 13/32. With four protected nodes, $\mathcal{A}$ targets two nodes and induces payoff

19/32. Thus four protected nodes is optimal.
d=4, a=3: The optimal strategy of $\mathcal{A}$ with a CP star consists in targeting peripheral nodes. The payoff induced is 1/16. With two equally protected nodes, $\mathcal{A}$ targets the unprotected nodes and one of the protected nodes. The resulting payoff is 3/16. With three protected nodes, $\mathcal{A}$ targets the unprotected node and one of the least protected nodes. The resulting payoff is 9/32. With four protected nodes, $\mathcal{A}$ targets three nodes, with payoff 7/16. Thus, four protected nodes is optimal.
d=4, a=4: The optimal strategy of $\mathcal{A}$ with a CP star consists in targeting all nodes. The payoff induced is 1/32. With two equally protected nodes, $\mathcal{A}$ targets all nodes. The resulting payoff is 5/36. With two unequally protected nodes, $\mathcal{A}$ again targets all nodes. The resulting payoff is 1/8. With three protected nodes $\mathcal{A}$ targets all nodes and the payoff is 3/16. Similarly, with four protected nodes $\mathcal{A}$ targets all nodes and the payoff is 5/16. So four protected nodes is optimal.

■

7 Appendix B

This appendix considers and networks containing multiple components and also considers alternative orders of moves between the designer and adversary, In some applications, the designer and adversary may have an interest as well as an ability to conceal their attack and defence strategies. We inquire whether the robustness of the star network is valid in such settings. So far we have focused on the case where, after choosing the network, $\mathcal{D}$ allocates resources first, followed by $\mathcal{A}$. Let us refer to this as the *DDA* game. The reverse scenario in which (after $\mathcal{D}$ chooses the network) $\mathcal{A}$ allocates his resources first, followed by $\mathcal{D}$, will be referred to as the $\mathcal{DAD}$ game. Simultaneous attack and defence will be referred to as the $\mathcal{D}*$ game.

The possibility of defence naturally alters the incentives of the designer to split the network into separate components. To study connectedness of robust networks in the case where $d > 0$, we strengthen our assumptions regarding convexity:

Assumption A.4: Suppose **A.1′** holds. In addition, for all $m \in \{2, 3, .., n\}$:
(a). $f(m-1) \geq \frac{1}{2}f(m)$
(b). $f(\frac{m}{2}) \leq \frac{1}{3}f(m)$

Observe that payoffs given in (3) satisfy (a) and (b) for $m \geq 4$. To make progress we will restrict attention on $a = d = 1$, and integer allocation of attack and defence resources. The following result allows for all the different orders of move and also permits an arbitrary number of components.

Proposition 2 *Consider the $\mathcal{DDA}$ game. Let $a = d = 1$, and suppose that allocation of resources can only take on integer values. Suppose (**A.1′**)-(**A.3**) hold. Then, if (**A.4a**) holds the star is robust within the class of connected networks. If in addition (**A.4b**) holds, then the star is robust among all networks.*

Proof: First consider a connected network with one node defended, j say. If $\mathcal{A}$ targets this node, the resulting payoff of $\mathcal{D}$ is $f(n)/2$. Now consider a star network, and suppose $\mathcal{D}$ protects the central node. By (**A.4a**) the resulting payoff of $\mathcal{D}$ is $f(n)/2$ at least. The argument is completed by showing that the maximum payoff of $\mathcal{D}$ in networks with two or more components is no more than $f(n)/2$. *$\mathcal{DDA}$ game:* Consider a star network, and suppose $\mathcal{D}$ protects the central node. By (**A.4a**) the optimal response of $\mathcal{A}$ is also to target the central node. This shows that with a star $\mathcal{D}$ guarantees himself payoff $f(n)/2$.

Consider next an arbitrary connected network g, and arbitrary defence by the designer. Suppose $\mathcal{A}$ attacks the defended node in the network. The resulting payoff of $\mathcal{D}$ is $f(n)/2$. Since $\mathcal{D}$ guarantees himself this payoff with a star, it follows that the star is robust within the class of connected networks.

Now consider networks with multiple components. We first establish the following useful result: under (**A.4b**), for all $\overline{m} \geq \underline{m}$

$$\frac{1}{2} f(\overline{m} + \underline{m}) \geq 1/2 f(\overline{m}) + f(\underline{m}) \tag{47}$$

Let $m = \overline{m} + \underline{m}$, and consider the following expression over $x \in [0, \frac{m}{2}]$:

$$\frac{1}{2} f(\frac{m}{2} + x) + f(\frac{m}{2} - x) \tag{48}$$

By convexity of $f(.)$, this expression is maximized at a corner, $x = 0$ or $x = m/2$. At $x = 0$ this is $3f(\frac{m}{2})/2$. It is $f(m)/2$ at $x = m/2$. By (**A.4b**), the maximum is $f(m)/2$. So (47) holds, as claimed.

Now consider an arbitrary network g with $k \geq 1$ components, labeled 1 to k. By the previous step, we can replace network g by another one, g' say, in which each component is a star and that yields the designer, in any sub-game perfect equilibrium of the $\mathcal{DDA}$ game, payoff at least as large as the payoff he obtains with g. Hence, in what follows, we restrict attention to networks in which each component is a star.

Let $n_1 \leq .. \leq n_k$ the size of the components in g. In particular, $n_1 + .. + n_k = n$. Suppose $\mathcal{D}$ defends the central node in the largest component, and let $\mathcal{A}$ attack the same node. The resulting payoff of $\mathcal{D}$ is

$$\frac{1}{2}f(n_k) + f(n_{k-1}) + .. + f(n_1) \tag{49}$$

By (47), this in turn is, at most

$$\frac{1}{2}f(n_k + n_{k-1}) + f(n_{k-2}) + .. + f(n_1) \tag{50}$$

and, repeating the argument, at most

$$\frac{1}{2}f(n_k + n_{k-1} + .. + n_1) = \frac{1}{2}f(n) \tag{51}$$

This also shows that whichever node $\mathcal{D}$ chooses to protect in g then, if $\mathcal{A}$ attacks the central node in the largest component, the resulting payoff of $\mathcal{D}$ is at most $f(n)/2$. Since we showed above that $\mathcal{D}$ could guarantee himself payoff $f(n)/2$ with a star, it follows that the star is robust among all networks.

■

Proposition 3 *Consider the $\mathcal{DAD}$ or $\mathcal{D}*$ game. Let $a = d = 1$, and suppose that allocation of resources can only take on integer values. Suppose **(A.1′)**-(**A.3**) hold. Then, if (**A.4a**) holds the star is robust within the class of connected networks. If in addition (**A.4b**) holds, then the star is robust among all networks.*

Proof: We will start by establishing the following property of any connected network. *There exists a node, i say, with the property that for any node $j \neq i$, there exist j-independent paths between i and at least half the nodes in the network.* This property allows us to show that, independently of the order of moves, the payoff of $\mathcal{D}$ is at most $f(n)/2$ in any connected network. Since he can guarantee himself exactly this payoff in a star network by defending the central node, this shows that the star is robust within the class of connected networks. Again, we have to demonstrate that the maximum payoff of $\mathcal{D}$ in networks with two or more components is no more than $f(n)/2$.

The following Lemma is useful in the proof of Proposition 1.

Lemma 2 *For any connected network g there exists a node i with the property that, for any $j \in g$ fixed, $j \neq i$, there exist j-independent paths connecting i and $\frac{n}{2}$ nodes in g at least.*

Proof. We provide a proof for minimally connected networks. If g is not minimally connected then there exists a minimally connected network g' obtained from g by deletion of links. Then a node i satisfying the property in g' also satisfies it in g.

The proof is by induction on n, the total number of nodes. For $n = 2$ the property is obviously satisfied. Let $n > 2$ and assume the property holds for any network with $n - 1$ or less nodes. Let g be minimally connected with n nodes. Consider g' obtained from g by removing a leaf l in g (i.e. a node with degree 1). Using the induction hypothesis on g' we can find i' satisfying the property for g'. Next, let i denote the neighbor of i' on the unique path between i' and l in g. We show that one of i or i' must satisfy the property for g. If i' satisfies the property for g then we are done. Suppose i' fails to satisfy the property for g. Let $Y_{s\backslash t}(g)$ ($y_{s\backslash t}(g)$) denote the set (cardinality) of nodes which can be connected to s in g through some t-independent path. Note that $Y_{i\backslash i'}(g) = N\backslash Y_{i'\backslash i}(g)$, and so $y_{i\backslash i'}(g) = n - y_{i'\backslash i}(g)$. Since $y_{i'\backslash i}(g) < \frac{n}{2}$ by hypothesis, it follows that $y_{i\backslash i'}(g) > \frac{n}{2}$. Next, let j denote a neighbor of i in g other than i'. Note that $y_{i\backslash j}(g) \geq y_{i'\backslash i}(g') + 1$. By definition of i' we have $y_{i'\backslash i}(g') \geq \frac{n-1}{2}$. Hence $y_{i\backslash j}(g) \geq \frac{n-1}{2} + 1 > \frac{n}{2}$. Thus we have shown that for any neighbor t of i in g, $y_{i\backslash t}(g) \geq \frac{n}{2}$. Since for any node non-neighbor t' of i there exists a neighbor t with $y_{i\backslash t'}(g) > y_{i\backslash t}(g)$, the proof is complete.

■

Proof of Proposition 3: We take each order of moves in turn.

$\mathcal{DAD}$ and $\mathcal{D}*$ games: Consider a star network, and suppose $\mathcal{D}$ protects the central node. By (**A.4a**) the optimal response of $\mathcal{A}$ is also to target the central node. This shows that with a star $\mathcal{D}$ guarantees himself payoff $f(n)/2$.

Consider next an arbitrary connected network g, and arbitrary defence (possibly using a mixed strategy). Suppose $\mathcal{A}$ targets a node i, identified in Lemma 2. By definition of i, and convexity of $f(.)$, notice that the resulting payoff of $\mathcal{D}$ is at most $f(n)/2$. So, by the first step, the star is robust within the class of connected networks.

The proof for the second part on general components is analogous to the proof given for $\mathcal{DDA}$ game and is omitted. ■

8 Appendix C: A richer model of attack

This section considers a richer model of the spread of attack which goes some way toward addressing the discontinuity of indirect attack. We now suppose that if attack resources a_i prevail over the defence at site i then they are available for further attacks on neighboring nodes. If they fail in their attack then they are neutralized and removed from the network. Similarly, for defence resources. If attack is defeated then the defence resources remain intact; if attack prevails then the defence resources are removed from the network.

Once attack resources a_i on node i prevail over defence resources the adversary can move them to a neighboring node j and engage in a contest with the defence resources d_j at that node. The outcome of the contest on node j is in turn defined by assumption (**A.2**).

The attack and defence conflict is now dynamic: we start with a network. The defence and attack allocation defines, via Assumption (A.2), a set of surviving nodes, and surviving defence and attack resources. The adversary moves the surviving attack resources to neighboring nodes with a view to maximizing damage on the network. The only restriction we impose is that resources not be left idle at a captured node and that movement across unprotected or already captured nodes are instantaneous. This attack resource re-allocation leads to a new round of contests on different nodes in the network. The outcomes of this conflict is defined by the contest function given in equation 9. The process proceeds in this way until there is no attack resource left or all nodes have been captured.

Since attack resources at a captured node are obliged to move to un-captured nodes in every period, in every period at least one unit of defence or attack resource is removed from the network. So there is an upper bound on the number of periods for the game, given by $a + d$.

The payoffs of the players are defined with respect to the network which survives at the end of the game. At every stage, given a network and attack and defence resource profile profile, the adversary relocates resources with a view to maximizing its eventual payoffs (and minimizing the payoffs of the designer).

Assumption A.3′: *Consider a network g.*

(i). Successful attack on node i means that the attack resources a_i remain intact and the defence resources d_i are removed from the network. Similarly, if defence prevails then the attack resources a_i are removed from the network and the defence resources d_i remain intact.

(ii). The designer is obliged to relocate (surviving) attack resources $(a_1, ...a_k)$ to defended nodes in the neighborhood of the successfully attacked nodes. If there are no such nodes, then

it must move to the neighbors of neighboring nodes and so forth. The attack resources move across nodes i with $d_i = 0$, instantaneously. (iii). The game ends when either all the defence or all the attack resources are removed from the network.

We first observe that when $d = 0$, attack spreads instantaneously under assumption ($\mathbf{A.3'}$), and so the dynamic model of attack leaves Theorem 1 unaffected. Let us turn to Theorem 2. The following result obtains a partial generalization when resources of attack and defence are equal, i.e., $d = a > 0$.

Theorem 3 *Suppose (**A.1**′), (**A.2**), and (**A.3**′) hold. Fix $a = d > 0$. For n sufficiently large, the star network is robust in the class of connected networks.*

Sketch of Proof. Consider first a star network. If $\mathcal{D}$ allocates all resources to the center, $\mathcal{A}$ optimally allocates one unit of resource to a distinct nodes in the periphery (notice, this is always feasible if $n \geq a + 1$). The resulting payoff of $\mathcal{D}$ is

$$\underline{B} = \mathbb{E}[f((n - d)I_0] \tag{52}$$

where I_0 denotes a Bernoulli random variable with mean 1/2. Hence $\underline{B}$ is a lower bound for the payoff $\mathcal{D}$ can guarantee himself with a star network.

Consider a connected network g and allocation of defence resources $\mathbf{d}$. Let $X(\mathbf{d}) = \{i_1, ..., i_k\}$ denote the set of nodes with non-zero defence resources, and $O = N - X$. Let $O_{i_t} \subset \big(O \cup \{i_t\}\big)$, $t = 1, ..., k$ denote the subset of nodes which can be reached from i_t through a path lying in $O \cup \{i_t\}$. Construct the sequence of sets $(N_{i_t}(g, \mathbf{d}))_{1 \leq t \leq k}$ recursively as follows:[15]

$$N_{i_1} = O_{i_1} \ , \ N_{i_2} = O_{i_2} - N_{i_1} \ , \ ... \ , \ N_{i_k} = O_{i_k} - \bigcup_{t=1}^{k-1} N_{i_t}$$

Let $n_{i_t} = |N_{i_t}|$, $t = 1, ..., k$. Relabeling nodes if necessary, assume that $n_{i_1} \geq n_{i_2} \geq ... \geq n_{i_k}$. Note that by connectedness of g, $\sum_{1 \leq t \leq k} n_{i_t} = n$. In what follows we let $X' = \{i_t \in X : d_{i_t} \leq n_{i_t} - 1\}$.

Case 1: $X' = X$.

[15] Note that in some cases the sequence constructed will depend on the particular order assigned to elements in X.

Suppose $\mathcal{A}$ allocates one unit of resource on d_{i_t} distinct nodes from $N_{i_t} - \{i_t\}$, $t = 1, ..., k$. By ($\mathbf{A.3'}$) the resulting payoff of $\mathcal{D}$ is bounded above by

$$\mathbb{E}[f(\sum_{i_t \in X} (n_{i_t} - d_{i_t}) I_{i_t})]$$

where $(I_{i_t})_{1 \leq t \leq k}$ denotes a family of i.i.d. Bernoulli random variables with mean 1/2.

By Lemma

$$\mathbb{E}[f(\sum_{i_t \in X} (ni_t - d_{i_t}) I_{i_t})] \leq \mathbb{E}[f((\sum_{i_t \in X} n_{i_t} - d_{i_t}) I_0)] \tag{53}$$

Hence, by (52), we have found an attack strategy leaving $\mathcal{D}$ at most the payoff he can guarantee himself with a star. The adversary is assumed to maximize one period payoffs at every stage of the conflict dynamic. So the strategy which maximizes damage will do at least as well as the above strategy.

Case 2: $X' \neq X$.

Notice that $X' \neq \emptyset$ if $n \geq a + d$. In particular, given $n_{i_1} \geq n/d$, $i_1 \in X'$ if $n \geq d(d+1)$. So suppose $\mathcal{A}$ allocates d_{i_t} unit resources on i_t, $t = 2, ..., k$ and $d_{i_1} + \sum_{i_t \in X - X'} d_{i_t}$ unit resources on i_1. By ($\mathbf{A.3'}$) the resulting payoff of $\mathcal{D}$ is bounded above by

$$\overline{B} = \mathbb{E}[f(n_{i_1} I'_{i_1} + \sum_{i_t \in X' - \{i_1\}} n_{i_t} I_{i_t} + \sum_{i_t \in X - X'} n_{i_t})]$$

where the Bernoulli random variables are independent, with mean 1/2 except I'_{i_1} which mean is

$$\frac{d_{i_1}}{d_{i_1} + (d_{i_1} + \sum_{i_t \in X - X'} d_{i_t})}$$

In particular, since this is less than 1/2 and $f(.)$ is increasing, we have

$$\overline{B} < \mathbb{E}[f(\sum_{i_t \in X'} n_{i_t} I_{i_t} + \sum_{i_t \in X - X'} n_{i_t})] \tag{54}$$

where all Bernoulli random variables are independent with mean 1/2. In what follows we let

$$\overline{B}' = \mathbb{E}[f(\sum_{i_t \in X'} n_{i_t} I_{i_t} + \sum_{i_t \in X - X'} n_{i_t})]$$

Observe that

$$\begin{aligned}\overline{B}' - \overline{B} &\geq \big(f(n) - f(n-n_{i_1})\big)\big(\mathbb{P}(I_{i_1}=1) - \mathbb{P}(I'_{i_1}=1)\big)\mathbb{P}(I_{i_2} = ... = I_{i_k} = 1) \\ &\geq \frac{n_{i_1}}{n} f(n)\bigg(\frac{1}{2} - \frac{d_{i_1}}{d_{i_1} + (d_{i_1} + \sum_{i_t \in X - X'} d_{i_t})}\bigg)\big(\frac{1}{2}\big)^{k-1} \\ &\geq \frac{1}{d}\bigg(\frac{1}{2} - \frac{d}{2d+1}\bigg)\big(\frac{1}{2}\big)^{k-1}\end{aligned} \tag{55}$$

Let $\varepsilon = \bigg(\frac{1}{2} - \frac{d}{2d+1}\bigg)\big(\frac{1}{2}\big)^{k-1}$ in what follows.

Next, observe that $\sum_{i_t \in X-X'} n_{i_t} \leq \sum_{i_t \in X-X'} d_{i_t} < d$. By (**A.1**$'$), it follows that

$$\overline{B}' \to_{n\to\infty} \mathbb{E}[f(\sum_{i_t \in X'} n_{i_t} I_{i_t})] \tag{56}$$

Also, by Lemma

$$\mathbb{E}[f(\sum_{i_t \in X'} n_{i_t} I_{i_t})] \leq \mathbb{E}[f(\big(\sum_{i_t \in X'} n_{i_t}\big) I_0)] \tag{57}$$

where, by the previous remark and renewed application of (**A.1**$'$)

$$\mathbb{E}[f((\sum_{i_t \in X'} n_{i_t}) I_0)] \to_{n\to\infty} \underline{B} \tag{58}$$

Combining (55), (56), (57), and (58) gives

$$\begin{aligned}\overline{B} &= \overline{B}' - (\overline{B}' - \overline{B}) \\ &\leq \overline{B}' - \varepsilon \\ &\to_{n\to\infty} \mathbb{E}[f(\sum_{i_t \in X'} n_{i_t} I_{i_t})] - \varepsilon \\ &\leq \mathbb{E}[f((\sum_{i_t \in X'} n_{i_t}) I_0)] - \varepsilon \\ &\to_{n\to\infty} \underline{B} - \varepsilon \\ &< \underline{B}\end{aligned} \tag{59}$$

Hence, again, we have found an attack strategy which leaves $\mathcal{D}$ with less payoff than he can guarantee himself with a star.

Since any connected network and allocation of defence resources is included in one of cases 1 and 2 examined above, it ensues that the star network is robust.

■

We turn next to the case where $a \neq d$. The following example shows that the star network need not be robust for $a < d$.

Example 8

Suppose $a = 1$ and $d = 2$. Consider a star network to start with. Simple inspection shows that allocating both unit resources to the central node is optimal for $\mathcal{D}$. $\mathcal{A}$ allocates his unique resource to one of the periphery nodes with resulting payoff to $\mathcal{D}$ equal to $\frac{1}{2}f(n-1)$.

Next consider a network consisting of two stars, each containing half the total number of nodes, and one link between the two central nodes. Suppose $\mathcal{D}$ allocates a unit of resource on each central node. It is immediate that an optimal response for $\mathcal{A}$ consists in allocating his unique resource to one of the periphery nodes and so, the resulting payoff to $\mathcal{D}$ is $\frac{1}{2}f(n-1) + \frac{1}{4}f(\frac{n}{2}) > \frac{1}{2}f(n-1)$.

△

9 References

1. Albert R, Jeong H, Barabási, A-L (2000), Error and attack tolerance of complex networks, *Nature*, 406: 378-82.

2. Baccara, M. and H. Bar-Isaac (2008), How to organize crime? *Review of Economic Studies*, 75, 4, 1039-1067.

3. Barabasi, A-L (1999), *Linked.* Perseus Books.

4. Bala, V. and S. Goyal. (2000a), A non-cooperative model of network formation, *Econometrica*, 68, 5, 1181-1229.

5. Bala, V. and Goyal, S. (2000b), An analysis of strategic reliability, *Review of Economic Design*, 5, 205-28.

6. Baye, M. (1998), *Recent Developments in the Theory of Contests: Advances in Applied Microeconomics.* JAI Press.

7. Kovenock, D. M. R. Baye and C. G. de Vries (1996), The all-pay auction with complete information, *Economic Theory*, 8, 2, 291-305.

8. Bier, V., S. Oliveros and L. Samuelson (2006), Choosing what to Protect: Strategic Defensive Allocation against an Unknown Attacker, *Journal of Public Economic Theory*, 9, 1-25.

9. Bolton, P. and M. Dewatripont (1994), The firm as a communication network, *Quarterly Journal of Economics*, 109, 809-839.

10. Bolton, P. and M. Dewatripont (1994), The firm as a communication network, *Quarterly Journal of Economics*, 109, 809-839.

11. Dixit, A. (1987), Strategic behavior in contests, *American Economic Review*, 77, 891-898.

12. Eilstrup-Sangiovanni, M. and C. Jones (2008), Strengths and Weaknesses of Networks: Why al-Qaeda may be Less Dangerous than Most Think, *International Security*, 33, 2, 7-44.

13. Farley, J. D. (2003), Breaking Al Queda Cells: A mathematical analysis of counter-terrorism operations, *Studies in Conflict and Terrorism*, 26, 399-411.

14. Farley, J. D. (2006), Building the perfect terrorist cell, Conference Talk.

15. Garicano, L. (2000), Hierarchies and the Organization of Knowledge in Production, *Journal of Political Economy*, volume 108, pages 874-904.

16. Garicano, L. and R. Posner (2005), Intelligence failures: an organization theory perspective, *Journal of Economic Perspectives*, 19, 4, 151-179.

17. Garoupa, N. (2007), Optimal Law enforcement and criminal organization, *Journal of Economic Behavior and Organization*, 63, 461-474.

18. Goyal, S. (1993), Sustainable communication networks, *Tinbergen Institute Discussion Paper, TI 93-250*, Rotterdam-Amsterdam.

19. Goyal, S. (2007), *Connections: an introduction to the economics of networks.* Princeton University Press.

20. Goyal, S. and A. Vigier (2009a), Interaction, infection and control. *Mimeo*, Cambridge University.

21. Goyal, S. and A. Vigier (2009b), Vaccination, social networks and public policy. *Mimeo*, Cambridge University.

22. Grotschel, M., C.L. Monma and M. Stoer (1995), Design of survivable communication networks, in M.O. Ball, TL. Magnanti, C.L. Monma and G.L. Nemhauser (eds) *Handbooks of Operations Research and management science: Network Models.* North Holland. Amsterdam, 617-672.

23. Gutfraind, A. (2009), The complexity of Markovian Network Interdiction, *Mimeo*, Cornell University.

24. Hart, S. (2008), Discrete Colonel Blotto and General Lotto games, *International Journal of Game Theory*, 36, 3, 441-460.

25. Hirshleifer, J. (1991), The paradox of power, *Economics and Politics*, 3, 177-200.

26. Hong, S. (2008), Hacking-proofness and Stability in a Model of Information Security Networks, working paper.

27. Jackson, M. O. and A. Wolinsky (1996), A strategic model of social and economic networks, *Journal of Economic Theory*, 71, 44-74.

28. Krueger, A. (1974), The Political Economy of the Rent-Seeking Society, *American Economic Review* 64, 3, 291?303.

29. Levine, S. (1999), *Fragile Dominion: Complexity and the Commons* Perseus Books, Reading, MA.

30. Myerson, R. (1977), Graphs and cooperation in games, *Mathematics of Operations Research*, 2, 225-229.

31. Nagaraja, S., Anderson, R. (2007) The topology of covert conflict, *Cambridge Computer Laboratory Technical Report 637.*

32. Powell, (2008), Sequential non-zero sum Blotto: allocating defence resources prior to attack, *Games and Economic Behavior*, forthcoming.

33. Radner, R (1992), Hierarchy: The Economics of Managing, *Journal of Economic Perspectives*, 30, 3, 1382-1415.

34. Radner, R. (1993), The organization of decentralized information processing, *Econometrica*, 61, 5, 1109-1146.

35. Roberson, B. (2006), The Colonel Blotto Game, *Economic Theory*, 29, 1?24.

36. Rothschild, M. and J. E. Stiglitz (1970), Increasing risk: I. A definition, *Journal of Economic Theory*, 2, 3, 225-243.

37. Sandler, T. and K. Hartley (2007), *The Handbook of Defence Economics, Volume 2: Defence in a Globalized World*. Elsevier. Amsterdam.

38. Smith, C. J (2008), Preface to special issue on *Networks: Games, Interdiction, and human interaction problems on networks*, Volume 52, 3, 109-110.

39. Skaperdas, S. (1996), Contest success functions, *Economic Theory*, 7, 2, 283-290.

40. Tullock, G. (1967), The Welfare Costs of Tariffs, Monopolies, and Theft, *Western Economic Journal* 5, 3, 224-232.

41. Tullock, G. (1980), Efficient rent seeking, *Towards a theory of the rent-seeking society*, edited by Buchanan, J., Tollison, R., and Tullock, G., Texas A&M University Press.

42. Van Zandt, T. (1999), Decentralized information processing in the theory of organizations, *Contemporary Economic Issues Volume 4: economic design and behavior*, edited by Murat Sertel. MacMillan Press. London.

43. Zakaria, F. (2008), The Rise of the Rest, *Newsweek*, May 12.