Robust Networks

Sanjeev Goyal * Adrien Vigier†

July 2, 2010

Abstract

Connections between individuals (firms, cities and countries) facilitate the exchange of goods, resources and information. Adversaries exploit these links to spread their attacks and lower the effectiveness of the network. Links thus 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 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 adversary. If defence and attack budgets are small, relative to the number of nodes, the star network is robust and the designer allocates all his resources to protect the central node. If defence and attack budgets are large then denser networks are robust and dispersed defence allocation is better.

*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 Nizar Allouch, Murali Agastya, Mariagiovanna Baccara, Oliver Bätz, Parimal Bag, Heski Bar-Isaac, Yann Bramoulle, Indranil Chakravarty, Matthew Elliot, Andrea Galeotti, Edoardo Gallo, Aditya Goenka, Matthew Jackson, Gilat Levy, Meg Meyer, Francesco Nava, Volcker Nocke, Romans Pancs, Rufus Pollock, Rony Razin, Bryony Reich, Rakesh Vohra, and seminar participants at Cambridge, Cornell, Laval, LSE, Madison, Mannheim, Northwestern, NUS, NYU, Oxford, Paris, Tilburg and Warwick for comments.

1 Introduction

Connections between individuals (firms, cities and countries) facilitate the exchange of goods, resources and information. Adversaries exploit these links to spread their attacks and lower the effectiveness of the network. Links thus create value but also enable the spread of attacks. 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 flow of information and 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 on nodes. The likelihood of successful attack on a node depends on the defence and attack resources allocated to it. Successful attacks on a node can potentially spread to linked nodes (depending on the distribution of defence resources).[1]

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

[1]Our work is naturally viewed as a study of contexts with an actor who has the authority to design networks and defend them; the three examples mentioned above illustrate this possibility. But our paper may also be viewed as a *normative* analysis of how networks should be designed and defended against attacks. This latter interpretation suggests that our insights may also be useful in contexts – e.g., inter-bank networks and vaccination against disease – where linking and defence decisions are made by individual nodes. The companion papers, Goyal and Vigier (2009a, 2009b) study games of decentralized linking and protection.

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.[2] A network is said to be *robust* if it constitutes a (sub-game perfect) equilibrium outcome of the game between the designer and the adversary.

We begin with an analysis of the situation in which the designer has zero defence resources. As 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 attack nodes in a larger component in preference to nodes in smaller components. 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 the adversary increases (Theorem 1). A fall in the size of groups 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".[3]

We then turn to the study of the general problem of design and defence of networks which face an adversary.[4] Given a network, an allocation of defence and attack resources gives rise to a set of inter-connected conflicts. 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.[5] 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. Our main finding is that, if defence and attack resources are small relative

[2]A component in a network is a (maximal) set of interconnected nodes; for formal definitions see section 2. 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.

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

[4]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.

[5]We use the standard Tullock (1967, 1980) contest function to model the outcome of the conflict on a node.

to the number of nodes then, a star network in which the designer allocates all resources to protecting the central node is robust (Theorem 2).

Let us sketch the arguments underlying this result. Consider a star network: for large enough number of nodes, the marginal value of protecting a periphery node is very small compared to the value from a marginal increase in the chances of protecting the central node (as elimination of this node then disrupts the network completely). So, in a star network the designer will concentrate all his resources on the central node. Suppose to fix ideas that the adversary allocates all resources to attacking the central 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[6] – necessitates a more spread out allocation of resources. Faced with such a network, the adversary can allocate his resources on these defended nodes to mimic the proportion of defence and attack resources as in the hub of the star network. The key theoretical observation is that this dispersed profile yields a distribution on surviving networks which stochastically dominates the distribution of surviving networks in the star network. Since the payoffs of the designer are convex in the size of a component, the designer obtains lower payoffs (and the adversary higher payoffs) in the dispersed network. Thus a star network with protected central node is robust. A comparison of Figures 5 and 6 illustrates this point: the distribution of surviving nodes from a center-protected star is a mean-preserving spread of the distribution from a core-periphery network (with multiple core nodes).

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 an efficiency based justification for the salience of hubs in large 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 research: the theory of networks and the theory of conflict/contests.[7]

[6] 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 6 illustrates a core-periphery network with two core codes.

[7] 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.

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). 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). We build on the canonical model in the networks literature – the connections model – to study design and defence of networks in the face of an intelligent adversary.[8] To the best of our knowledge this is the first paper to do so.

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 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.

The theory of contests studies allocation of resources in situations of conflict; influential contributions include Tullock (1967, 1980), Sandler and Hartley (2007), Esteban and Ray (2010), Dixit (1987), Hirshleifer (1991), Skaperdas (1996), 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)) Szentes and Rosenthal (2003), 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 builds on the standard model of contest success function and extends the theoretical framework along two dimensions: one, we locate individual battles within a network of interconnections 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.

The rest of the paper is organized as follows. Section 2 presents our model of design, defence and attack in networks. Section 3 takes up the case where the defence budget is zero and analyzes the pure design and attack model; here we also contrast the effects of strategic and random attacks, respectively, on the design of robust networks. Section 4 presents our analysis of design, defence and attack in networks. Section 5 concludes. Appendix A contains the proofs of the longer results, appendices B-D explore the generality of our results.

[8]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 and an adversary. 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 the designer are observed by the adversary, who then chooses an attack strategy. The initial network design and the subsequent attack together yield a probability distribution on the surviving networks 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. $\mathcal{D}$ chooses links between the nodes and allocates a defence budget $d \in \mathbf{N}$ across the nodes to protect the network. We let $\mathbf{d} = (d_1, d_2, ..., d_n)$ denote this allocation, where $d_i \geq 0$ and $\sum_{i \in N} d_i \leq d$.

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 $\mathcal{D}$.[9]

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, while the network such that $g_{ij} = 0$ for all i and j is called *empty*, and denoted g^e. A *core-periphery* network has two groups of nodes, N_1 and N_2. Nodes in N_1 constitute the periphery and have a single link each and this link is with a node in N_2; nodes in N_2 constitute the core and are fully linked with each other and with a subset of nodes in N_1. The *star* network is a special case of such an architecture in which the core contains a single node.

A sequence of nodes $i_{1,}.., i_k$ is said to constitute a path from node i to node j if and only

[9] 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).

if $g_{ii_1} = g_{i_1i_2} = ... = g_{i_{k-1}i_k} = g_{i_kj} = 1$. Two nodes are said to be connected if and only if there exists a path between them. A component of the network g 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 maximum component of g is a component with maximum cardinality in $\mathcal{C}(g)$.

Fix a network g on all N nodes. For any subset of nodes $N' \subset N$, a network g' on N' is said to form a sub-graph of g if $g_{ij} = 1$ whenever $g'_{ij} = 1$. We let $G(g)$ denote the set of sub-graphs of g.

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 $\mathcal{D}$ 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 $f(0) = 0$ and $f(n) = 1$.

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, strictly convex, $f(0) = 0$ 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*[10]

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 the number of individuals). The total payoff to all the 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

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

network g is:

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

Thus, the returns from a network are component additive. $\triangle$

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 from a pair of linked nodes is $4/n^2$. Generalizing, for a component C_k, with m nodes, the surplus is:

$$f(m) = \text{Prob}[i \in C_k \text{ demands good}] \ \times \ \text{Prob}[j \in C_k \text{ can supply good}] = \left[\frac{m}{n}\right]^2 \tag{3}$$

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:

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

$\triangle$

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 Mr. 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 probabilities across the different components, and is given by:

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

$\triangle$

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 $\mathcal{A}$ observes the network g and the allocation $\mathbf{d}$ chosen by $\mathcal{D}$ and then makes his attack decisions. We let $\mathbf{a} = (a_1, a_2, ..., a_n)$ denote the allocation chosen by $\mathcal{A}$, where $a_i \geq 0$ and $\sum_{i \in N} a_i \leq a$. We now develop a model of attack and defence on networks.

For simplicity, we use the standard Tullock (1980) contest function. The probability of successful attack on a node whenever $a_i \geq 0, d_i > 0$ is given by:

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

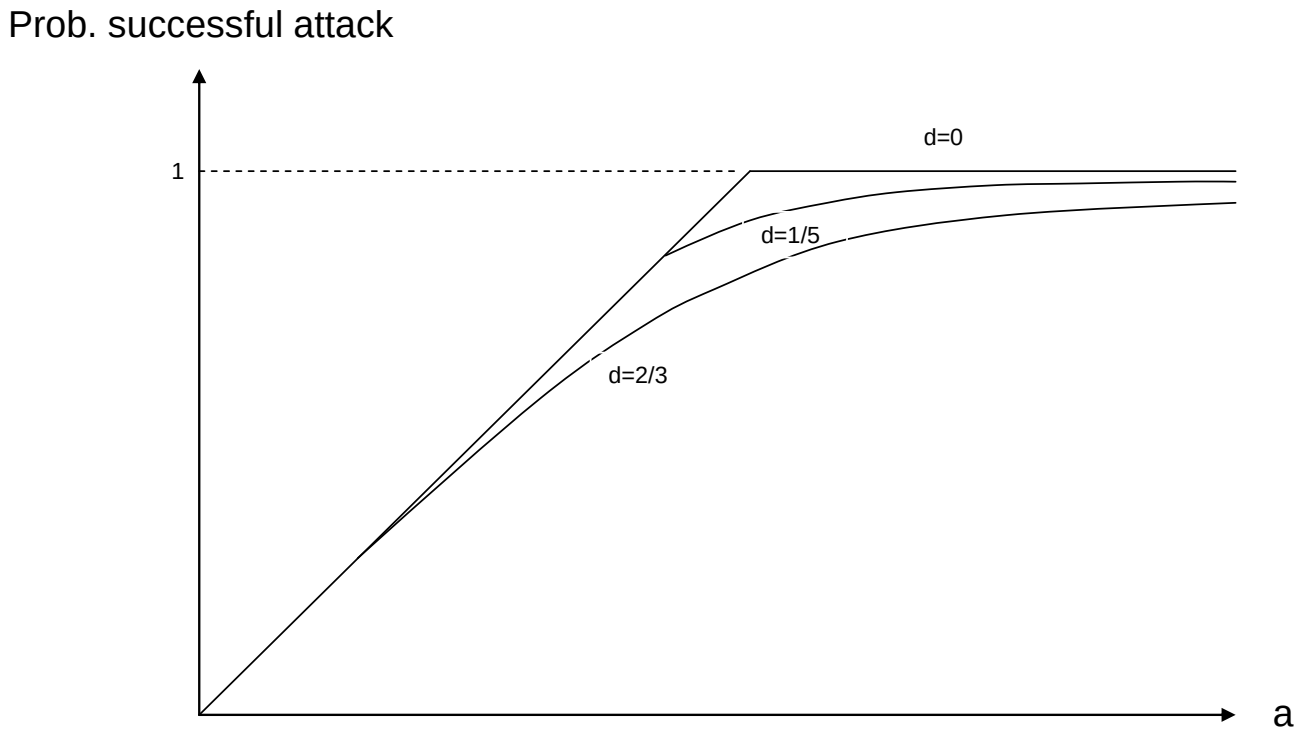

Figure 1: Probability of successful attack on a node

Next observe that, if $d_i = 0$, then the probability of successful attack is 1, for every $a_i > 0$. In many applications it is more reasonable to suppose that the probability of successful attack on a node is proportional to the attack resources allocated to it. For instance, it is implausible that an unprotected criminal or terrorist is detected with probability 1, with very small monitoring efforts. Similarly, it is implausible that an undefended city is captured with probability 1 through the deployment of a tiny enemy force.

These considerations motivate an extension of the Tullock (1980) formulation. We say that the probability of successful attack for $d_i \geq 0$, $a_i \geq 0$, is:

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

Observe that this formula always yields a number between 0 and 1, and therefore defines a probability. Figure 2 illustrates the behavior of our contest function as we vary the attack and defence allocation on a node. These considerations are summarized as follows.

Assumption A.2: *Suppose designer allocates d_i and adversary allocates a_i to node i. If*

$a_i + d_i > 0$, *then the probability of successful attack is given by* $\min\{a_i, \frac{a_i}{a_i+d_i}\}$. *If* $a_i + d_i = 0$, *then the probability of successful attack is* 0. *The success of attack is independent across nodes.*

Remark 1: We have also investigated a simpler and more general formulation of the contest function. The probability of successful attack on node i for $a_i, d_i > 0$ is given by $a_i^\gamma/(a_i^\gamma + d_i^\gamma)$, for some $\gamma > 0$. Suppose that player allocation on node i, $a_i, d_i \in \{0\} \cup X$, where $X = \{x \in \mathcal{R}_+ | x \geq 1\}$. Once this restriction on strategy space is in place, there is no longer the problem that the probability of attack is equal to 1 with very small attack resources. We can now dispense with the 'minimum' operator. Appendix C shows that our principal insights (Theorems 1 and 2) extend to a model with this general contest function.

We now describe our assumptions on how an attack spreads through a network. Our model of attack reflects two ideas. The first idea is that attack spreads through links, and so an attack from i can spread to j only if there is a path between the two nodes. The second idea is that likelihood of an attack spreading from node i to node j will depend on the defence resources allocated to j. To get our main arguments across, we assume a particularly simple threshold property of attack: attack on node i moves to a neighboring node j, if and only if $d_j < 1$. And, we shall say that a 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* $d_j < 1$.

Figure 2 illustrates the working of these assumptions and the key role of defence as a 'firewall' containing the spread of attacks in a network. Consider first a ring network in which d non-adjacent nodes are allocated a unit of defence resources each. If $\mathcal{A}$ allocates a unit of attack resources on one node between the protected nodes then attack spreads and eliminates all nodes except for the protected nodes. Consider next a star network with all defence resources allocated on the central node. Suppose $\mathcal{A}$ allocates one unit to a periphery nodes: then the a nodes are eliminated but there is no spread of the attack through the network, due to the presence of the protected central node.

Remark 2: We have also investigated a smoother model of spread of attack in which successful attack resources move to neighboring nodes and engage in general Tullock contests with

[11] Observe that if the designer's budget $d = 0$, then any path between two nodes is weak.

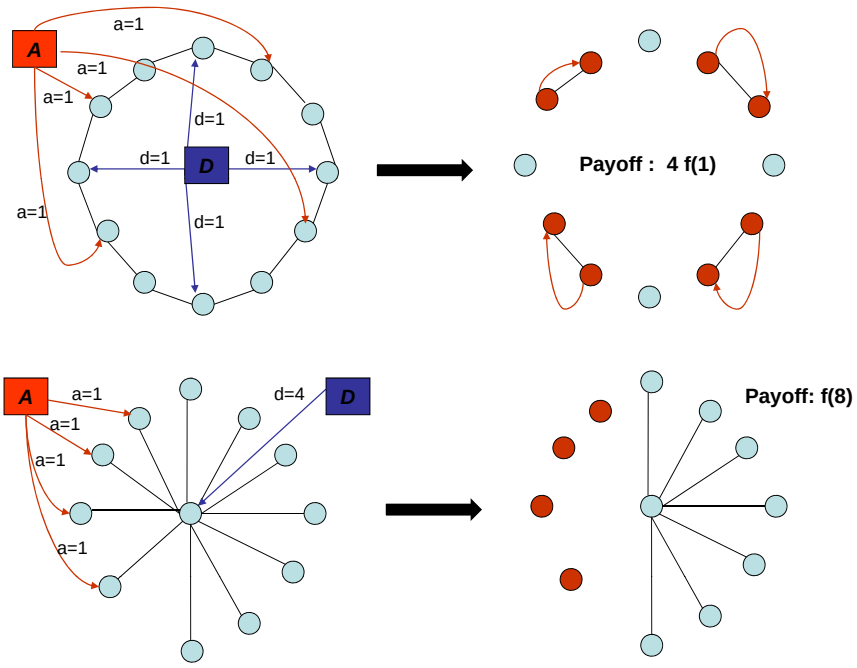

Figure 2: Attack & defence on ring and star; $n = 12, a = d = 4$.

defence resources on those nodes. Appendix D presents this model and develops conditions under which Theorem 1 and 2 extend.

We are now ready to define payoffs. First, $\mathcal{D}$ chooses a strategy $(g, \mathbf{d})$. This strategy is observed by $\mathcal{A}$ who then chooses an attack strategy $\mathbf{a}$. Given a network g, and allocations of resources $\mathbf{d}$ and $\mathbf{a}$, assumptions (**A.2**) and (**A.3**) define a probability distribution on the set of sub-networks $g' \in G(g)$ surviving the attack. Let this distribution be given by P. So $P(g'|\mathbf{a}, \mathbf{d}, g)$ is the probability of network g' surviving from the game of conflict with defence $\mathbf{d}$ and attack $\mathbf{a}$ played out on the network g. We require

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

The payoffs of $\mathcal{D}$ from choosing strategy $(g, \mathbf{d})$ when $\mathcal{A}$ chooses allocation $\mathbf{a}$ are:

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

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

3 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 the adversary 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 $C_k \in C(g)$ survives if and only if attack is unsuccessful on all of its nodes. The probability of this event is simply $\prod_{i \in C_k}(1 - a_i)$.[12] Then, component additivity implies that the payoff of $\mathcal{D}$ facing attack $\mathbf{a}$ is:

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

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 $\mathcal{A}$ 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 $a \leq n - 1$ then a robust network contains at least $a + 1$ maximum components, and at most one component which is smaller.*[13]

Proof: First, we establish that at most one node is attacked in a component. When $\mathcal{A}$ attacks two nodes with positive resources, there is positive probability of a state in which both nodes

[12] Here we are assuming that $a_i \leq 1$; this is without loss of generality as adversary will never set $a_i > 1$, given our assumption (**A.2**) and the hypothesis $d = 0$.

[13] 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.

are eliminated: this is wasteful as elimination of one node is sufficient to remove the entire component.

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 $a+1$ maximum components. Suppose this is not the case and let component C_1 denote a maximum component. 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' is maximum, 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 maximum size $\overline{s}$. 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 (weakly) smaller than the maximum 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 $\mathcal{D}$ to break up the network. So greater convexity of $f(.)$ induces robust networks which consist of fewer and larger components. The following result illustrates these effects. Define $x(a,\alpha) = \frac{\alpha a}{\alpha - 1}$.

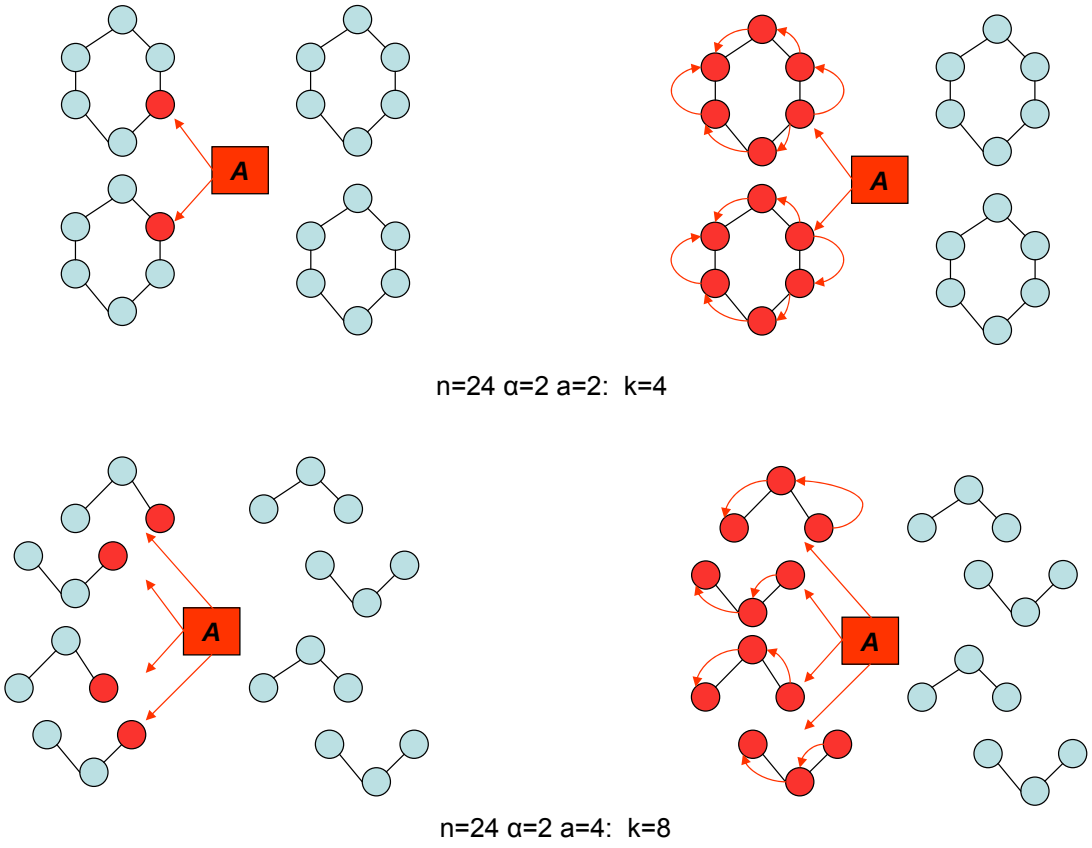

Figure 3: Robust networks: $n = 24$, α=2 and $a = 2, 4$

Proposition 1 *Suppose payoffs are given by*

$$f(m) = \left[\frac{m}{n}\right]^{\alpha} \quad \text{for } \alpha > 1. \tag{13}$$

*(**A.2**)-(**A.3**) hold, $a > 0$ and $d = 0$. If $a < n/2$, $x(a, \alpha) \in \{a + 1, .., n\}$ and divides n, then the unique robust network consists of $x(a, \alpha)$ equal size components.*

Figures 3 and 4 illustrate robust network structures, optimal attack strategies and surviving networks for $n = 24$, $\alpha = 2, 3$ and for adversary budgets $a = 2, 4$.

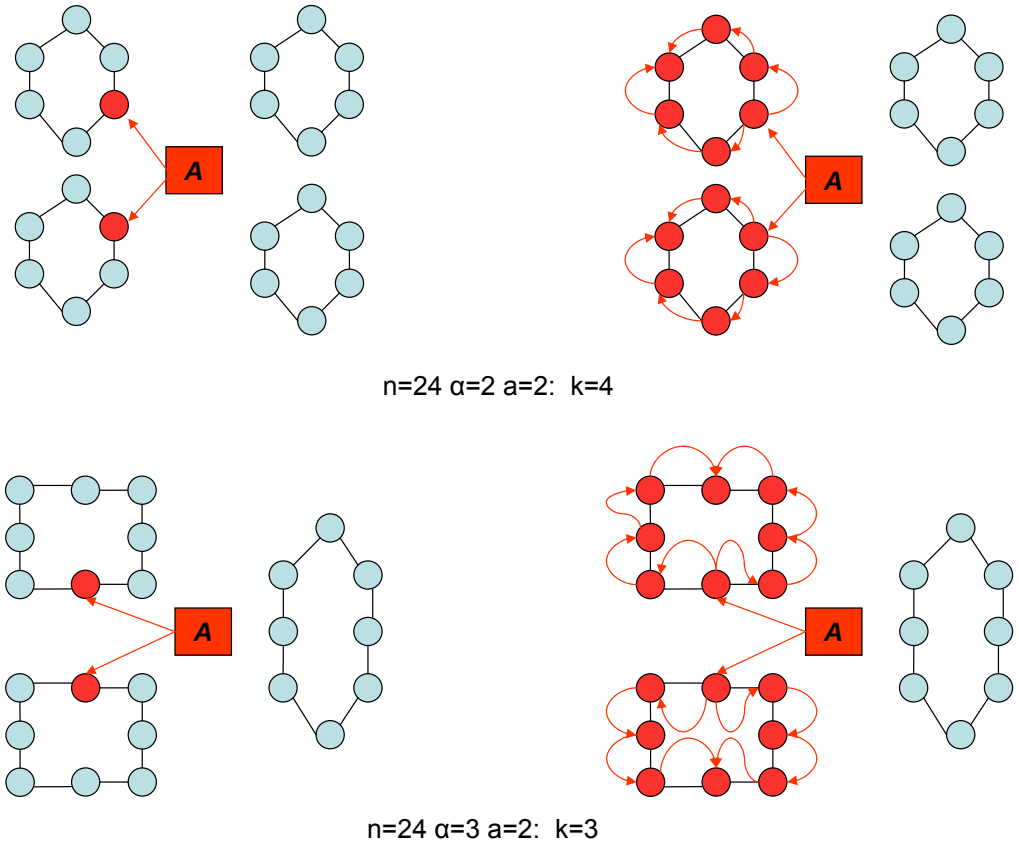

Figure 4: Robust networks: $n = 24$, α=2, 3 and $a = 2$

The proof is provided in Appendix A. Let $a < n/2$. Suppose $x(a, \alpha) \in \{a+1, .., n\}$; observe that if $\alpha = 2$ then this is true for all a. We first show that a network with $x(a, \alpha$ 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 maximum size is dominated by some network with equal size components. The claim then follows from Theorem 1.

Theorem 1 and Proposition 1 suggest that as the adversary budget grows, $\mathcal{D}$ responds by splitting the network into smaller components. In the context of Example 2 our results suggest that as the adversary budget grow, fewer tasks are carried out. In the context of Example 3, our results imply that larger adversary budgets imply less complex tasks are carried out by the network. These interpretations are 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 4 *Random attack and unequal components*

Suppose that $n = 4$, $f(1) = 0.01$, $f(2) = 0.05$, $f(3) = 0.50$, $f(4) = 1$. Uniform random attack is given by $q = a/4$, where a varies across $\{0, 1, 2, 3, 4\}$. The payoffs to $\mathcal{D}$ 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 are $(1-q)^3[0.50]+(1-q)[0.01]$ and the payoff from the connected network are $(1-q)^4[1]$. 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 or unequal.

△

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

Example 5 *Number of components: random vs. strategic attack*

Suppose component payoffs are given by (3). Consider uniform random attack $q = a/n$. It is easily checked that the payoff of $\mathcal{D}$ from a network with k equal size components is:

$$\frac{1}{k}\left[1-q\right]^{\frac{n}{k}} \tag{14}$$

At $q = 0$, the optimal network is trivially connected. Since payoffs are continuous in q complete connectedness 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 payoff of $\mathcal{D}$ is $(1-\frac{1}{n})^n \sim e^{-1} \sim \frac{1}{2.72} = 0.38$. Our computations in Proposition 1, on the other

hand, tell us that under strategic attack robust networks have two components and the payoff of $\mathcal{D}$ is $1/4 = 0.25$, *irrespective of the number of nodes.* Thus the number of components in the robust networks differ – 1 *vs* 2 – and the payoff of $\mathcal{D}$ 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 the budgets of the designer and adversary, the star network is robust. The designer allocates all his resources to defend the central node. By contrast, when defence and attack budgets are large relative to the number of nodes, then dense networks are robust and dispersed allocation of defence resources is dispersed.

Theorem 2 *Suppose (**A.1**)-(**A.3**) hold. Fix budgets $a, d \in \mathcal{N}$. For n sufficiently large, the star network is robust in the class of connected networks and the designer allocates all resources to the central node.*

The proof of this result is given in Appendix A. There are three steps in the proof. First, we show that given a star network in which all defence resources are allocated to the central node, the adversary allocate $c \leq a$ units of resource to the center and concentrates all remaining $a - c$ resources on $[a - c]$ periphery nodes. Spreading resources across more nodes implies a greater spread in the distribution of surviving nodes; by convexity of $f(.)$, spreading resources across many nodes raises the payoff of the designer and, correspondingly, lowers the payoff of the adversary.

Second, we show that given an allocation of $c > 0$ units to the central node, it is optimal for the designer to allocate all resources to protecting the central node. Diverting a subset of

resources to some periphery nodes ensures their survival, but it comes at the cost of lowering the probability of survival of the central node. Given convexity of $f(.)$, for large n, the cost is greater than the benefit.

So consider the star network with all defence resources allocated to the central node. Suppose the adversary optimally allocates $c \geq 0$ units to attack the central node. Then the payoffs to the designer are:

$$\frac{d}{d+c} f(n-(a-c)). \tag{15}$$

Third, we show that, in any network other than the star and with any any allocation of defence resources, the payoff to the designer is bounded above by the payoff given in equation (15). To illustrate why this is true, suppose $a = d = 4$ and let $c = a$ be the optimal attack of the adversary. Consider a core-periphery network with two nodes in the core, each with $d/2$ units of defence resources, and $n/2 - 1$ nodes in their respective periphery. It is easily checked that the distribution of surviving nodes in a star network with all resources allocated to the central node is a mean-preserving spread of the distribution of surviving nodes in the former core-periphery network whenever $\mathcal{A}$ allocates half his budget on each node in the core. Figures 5 and 6 illustrate these outcomes. Moreover, the surviving nodes in a center-protected star constitute a component. Since $f(.)$ is increasing and convex, the payoff of $\mathcal{D}$ from a star network with all resources allocated to the central node is strictly higher than the maximum payoff he can hope to obtain in the core-periphery network with two nodes in the core. The arguments in our proof show that this intuition can be generalized to cover arbitrary connected networks and allocations.

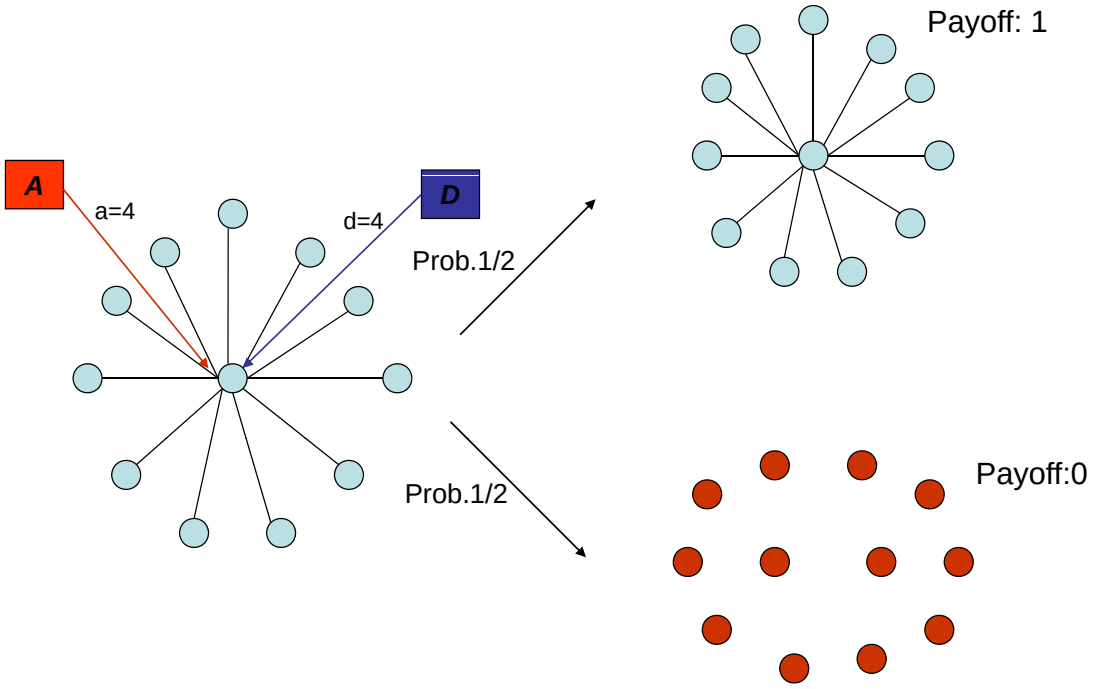

Figure 5: Attack & defence on a star

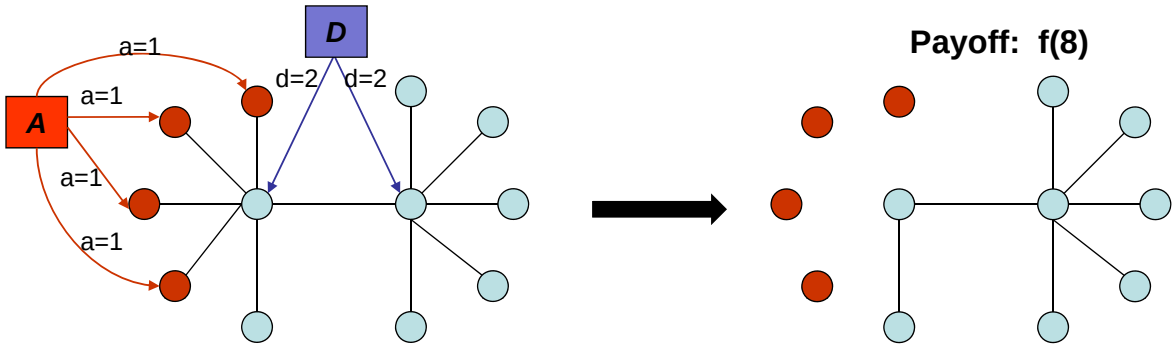

Figure 6: Core-periphery network: attacking the periphery nodes

The statement of the theorem is silent on the optimal strategy of the adversary. We now discuss how the optimal attack strategy depends on the nature of the reward function $f(.)$. We illustrate this by deriving the optimal attack strategy in Example 1 (the connections model) and Example 3 (the complementary skills and coordination model), respectively. Observe that, in the connections model, as n gets large the effect of any finite set of players on payoffs becomes insignificant. Consider a star network with all defence resources allocated on the central node. The payoffs to the designer if the adversary attacks the center with all resources are equal to $d/(d+a) < 1$. By contrast, if the adversary allocates all a units to the periphery nodes then the central node is safe; the payoff to the designer is $f(n-a)$, which converges to 1 as n grows large. So the optimal strategy for the adversary is to allocate all attack resources to the central node.

Next consider the complementary skills and coordination example. In the star network if both the designer and the adversary allocate all resources to the central node the payoff to the designer is $d/(d+a)$. By contrast, if the adversary assigns 1 unit each to a distinct peripheral nodes then the payoff to the designer is given by $f(n-a) = 1/2^a$. So, for instance, if $d = a \geq 2$ then $1/2^a < d/d+a$ and the optimal attack strategy of the adversary is to target a peripheral nodes.

Empirical research in networks highlights the importance of hubs in real world networks (see e.g., Goyal (2007) and Jackson (2008)). In an influential paper, Albert, Jeong and Barabasi (2000) highlight the vulnerability of hub-spoke architectures to strategic attacks: 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. In an interesting class of economic environments – where payoffs from exchange are increasing and convex – networks with hubs are robust. Thus our work provides an efficiency based justification for the salience of hubs in real world networks which face adversaries.

We have so far focused on the case where budgets are small relative to the number of nodes. This appears to us to be the natural model for large networked systems – such as infra-structural networks in countries, large criminal organizations and computer networks.

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.

Proposition 2 *Consider the payoffs in the connections model. Suppose (**A.2**)-(**A.3**) hold, resource allocations take integer values only, and consider the set of connected networks.* (i) *If $d = 1$ then, for all a, the star network is robust.* (ii) *If $a = 1$ then, for $d < n$, the star network is robust. For $a = 1$, $d = n$ 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 $\mathcal{A}$ 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, α' is defended. If $\mathcal{A}$ 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 $\mathcal{A}$'s strategy is feasible, the star is robust.

The argument for the second part with $a = 1$, $0 \leq d < n$, builds on two observations. Fix some network g and defence allocation $\mathbf{d}$. 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 with all defense resources allocated to the center, $1/(d+1)$. On the other hand, the set of nodes to which attack eventually spreads will be smaller than in the case of the case of the center-protected star (where all n nodes are eliminated). Nonetheless, there always exists a node j such that successful attack on this node exposes n_j unprotected nodes to the spread of attack where $n_j \geq n(d_j/d)$. So the maximum

payoff to $\mathcal{D}$ from such a network and defence is

$$\frac{d_j}{d_j+1}f(n)+\frac{1}{d_j+1}f(n-\frac{d_j}{d}n) \tag{16}$$

Indeed, we can write the difference in payoff between this network and allocation of defence and the center-protected star as:

$$\frac{d_j}{d_j+1}f(n)+\frac{1}{d_j+1}f(n-\frac{d_j}{d}n)-\frac{d}{d+1}f(n) \tag{17}$$

So the attractiveness of network g and a possibly dispersed defence $\mathbf{d}$ 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 is robust.

Next we explain why the star is not robust for $d = n$. If $\mathcal{D}$ protects every node and $\mathcal{A}$ attacks the central node, the payoff to $\mathcal{D}$ is

$$\frac{n-1}{2}f(1)+\frac{1}{2}f(n) \tag{18}$$

By contrast, in the ring network, with all nodes protected, the payoff to $\mathcal{D}$ is

$$\frac{1}{2}f(n-1)+\frac{1}{2}f(n) \tag{19}$$

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 $\mathcal{D}$ is at most $f(n-1)$ which is clearly smaller than the payoff from the ring.

We turn finally to the case where the budgets of *both* players are large relative to the number of nodes. The problem of attack and defence in networks with large resources on both sides is complex and we have been unable to characterize robust networks for general a, d and n. The following example (with $n = 4$) provides a characterization of robust (connected) networks.

Example 6 *Design, attack and defence with large budgets*

Consider the payoffs in the connections model. Suppose $n = 4$, (**A.2**)-(**A.3**) hold, and allocations take integer values only. The connected networks with 4 nodes are the line network (g^l), the star network ($g^{\rm s}$), the star with one pair of periphery nodes linked ($g^{\rm sp}$), the ring

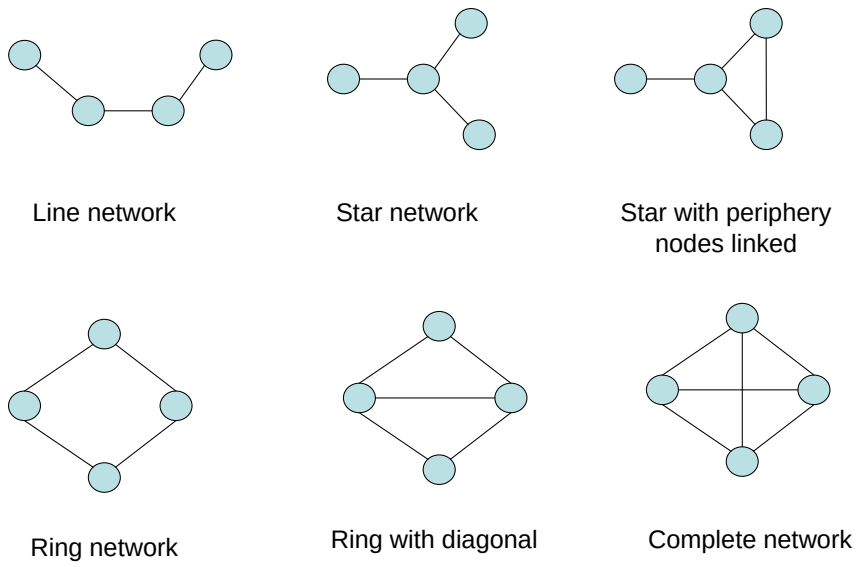

Figure 7: Connected networks: n=4

network (g^{r}), the ring network with one diagonal link (g^{rd}) and the complete network (g^{c}). Figure 7 illustrates these networks.

Appendix A provides the complete set of computations for robust networks, optimal defence and attack. We would like to highlight two points. Our first observation is about the attractiveness of a dispersed defence strategy for a designer with large defence resources. To see this let us look at $d = 4$ and $a = 1$, and the 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. 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, a dispersed defence strategy is better for $\mathcal{D}$ than the strategy of concentrating all resources on the hub node. Figure 8 illustrates this argument.

Our second observation is about the architecture of robust networks. The star network is robust when attack and defence resources are small but the complete network is robust when these resources are large. To get a sense of the factors underlying this, let us compare the star network with the complete network. Fix $d = 4, a = 1$. $\mathcal{D}$ can do better than the star with dispersed defence, by ensuring connectedness of the peripheral nodes in the event that attack on the center is successful. In particular, if $\mathcal{D}$ protects all nodes in the complete network then

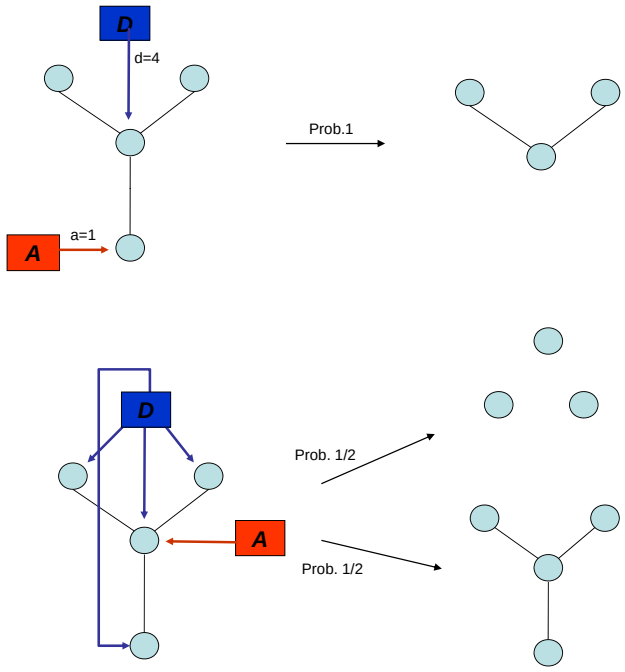

Figure 8: The attractions of dispersed defence

his minimum payoff is $f(3)/2+f(4)/2 = 25/32$. This shows that the complete network strictly dominates a star. Figures 9 illustrates these considerations. The matrix below summarizes our computations for robust architectures.

	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}

△

We conclude this section with three remarks.

One, we observe that Theorem 2 and the other results in this section all maintain the assumption that the network is connected. Proposition 1, 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 $\mathcal{D}$ 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 strategies. If, on the other hand, $f(.)$ is close to being linear then multiple components will be better. Proposition 3 in Appendix B

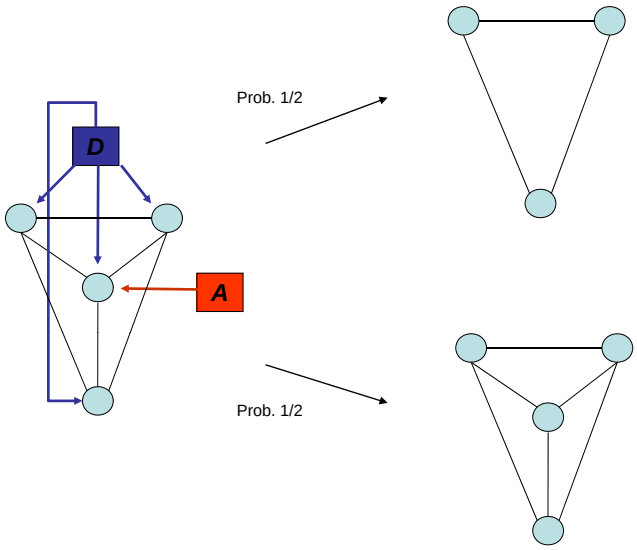

Figure 9: Defence and attack on complete network

develops a set of sufficient conditions on the returns function $f(.)$ which ensure connectedness of the robust network.

Two, we have assumed that $\mathcal{A}$ observes the allocation of defence resources made by $\mathcal{D}$ 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 4 in Appendix B presents this result in a setting with different orders of move and allows for multiple components.

Three, the hub-spoke network remains robust in the face of uniform random attacks. The following example illustrates this point.

Example 7 *Design and defence against random attack*

Consider the connections model and fix designer budget $d > 0$ and adversary budget $a > 0$. As before define uniform random attack as the case where every node is attacked by an equal amount of resource: $a_i = a/n$, $\forall i \in N$.

Define ϵ as a positive but small number. We say that a network is ϵ-efficient if it attains a payoff which is at most ϵ less than the payoff in an efficient network.

We will show that for any $\epsilon > 0$, there is a number of nodes $n(\epsilon)$ such that for all $n \geq n(\epsilon)$ the hub-spoke network attains a payoff of $1 - \epsilon$ or more. Since 1 is the maximum attainable

payoff for the designer, the hub-spoke network is ϵ-efficient.

Suppose the designer allots all his resources to the central node. Then the payoff from the hub-spoke network faced with a uniform random attack is

$$\frac{d}{d+a/n}\sum_{k=0}^{n-1}\binom{n}{k}\left[\frac{a}{n}\right]^k\left[1-\frac{a}{n}\right]^{n-1-k}f(n-k)$$
$$\geq \frac{d}{d+a/n}f(n-a) \tag{20}$$

where the inequality follows from the convexity of $f(.)$.

Next observe that for fixed a and d, $d/[d+a/n]$ converges to 1 as n gets large. Similarly, for fixed a, $f(n-a)=(n-a)^2/n^2$ converges to 1 as n gets large. So there is an $n(\epsilon)$ such for all $n \geq n(\epsilon)$, $d/[d+a/n] \geq \sqrt{1-\epsilon}$ and $f(n-a) \geq \sqrt{1-\epsilon}$. It now follows from equation (20) that for $n \geq n(\epsilon)$, the payoff to the designer from a hub-spoke network is greater than $1-\epsilon$.

△

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 a 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, the designer allocates all his budget to protecting the central node. The adversary may spread his resources across the central node and some peripheral nodes. On the other hand, if the budgets are large then denser networks with dispersed defence allocations are robust.

6 Appendix A: Proofs

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

$$f(s)(\frac{n}{s} - a) \tag{21}$$

It is easily checked that for $f(.)$ given by (4), this is maximized at $s = \frac{n(\alpha-1)}{a\alpha}$.

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

$$f(\overline{s})(\beta - a) + f(s) \tag{22}$$

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

$$f(\overline{s})(\beta - a) + \frac{s}{\overline{s}} f(\overline{s}) \tag{23}$$

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

$$f(\overline{s})(\frac{n}{\overline{s}} - a) \tag{24}$$

But by the first step, for $f(.)$ given by (4) 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 maximum size. By Theorem 1, it follows that a network with $\frac{\alpha a}{\alpha-1}$ equal size components is robust.

■

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 $n_1 I_1 + .. + n_k I_k$.*

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 $n_1 I_1 + n_2 I_2$, respectively. Define $1 - \delta = \alpha$. Then

$$F(x) = \begin{cases} \alpha & \text{if } 0 \leq x < n \\ 1 & \text{if } x = n \end{cases} \tag{25}$$

and

$$G(x) = \begin{cases} \alpha^2 & \text{if } 0 \leq x < n_1 \\ \alpha & \text{if } n_1 \leq x < n_2 \\ 1 - \delta^2 & \text{if } n_2 \leq x < n \\ 1 & \text{if } x = n \end{cases} \tag{26}$$

So, using Theorem 1 in Rothschild and Stiglitz (1970), $(n_1 + n_2)I_1$ is a mean-preserving spread (MPS) of $n_1 I_1 + n_2 I_2$ if and only if

$$\alpha - \alpha^2 = 1 - \delta^2 - \alpha \tag{27}$$

or, substituting for δ

$$\alpha - \alpha^2 = 2\alpha - \alpha^2 - \alpha \tag{28}$$

So the result holds for $k = 2$. Next, suppose the result holds up to $k \geq 2$. We want to show that it also holds for $k + 1$.

Observe that 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 $X + Q$.

But then setting $X = n_1 I_1 + n_2 I_2 + ... + n_k I_k$, $Y = (n_1 + n_2 ... + n_k)I_1$, $Q = n_{k+1} I_{k+1}$, using the result for $k = 2$ and the induction step, it follows that $(n_1 + n_2 + .. + n_{k+1})I_1$ is a MPS of $n_1 I_1 + n_2 I_2 + n_3 I_3 .. + n_{k+1} I_{k+1}$.

■

Proof of Theorem 2: 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. So the proof will focus on the case where both designer and adversary have positive budgets.

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, given some adversary allocation c to the central node, we show that with a center-protected star the designer can earn $df(n - (a - c))/(c + d)$.

Last, we show that, for any other network and for any other defence strategy, the payoff of the designer is bounded above by $df(n-(a-c))/(c+d)$.

Step 1: Consider a star network and suppose $\mathcal{D}$ allocates all his resources to protecting the central node. 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

$$(1-a_{i_1})(1-a_{i_2})f(m+2) + [a_{i_1}(1-a_{i_2}) + a_{i_2}(1-a_{i_1})]f(m+1) + a_{i_1}a_{i_2}f(m) \tag{29}$$

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

$$(1-\overline{a})f(m+2) + \overline{a}f(m+1) + a_{i_1}(\overline{a} - a_{i_1})[f(m+2) - 2f(m+1) + f(m)] \tag{30}$$

By convexity of $f(.)$, $f(m+2) - 2f(m+1) + f(m) > 0$. So the 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 establish that it is optimal for the designer to assign all resources to the central node. Suppose the designer allocates all resources to the hub node and the adversary best responds by allocating c units to the central node. Then the payoff to the designer is:

$$\frac{d}{d+c} f(n-(a-c)). \tag{31}$$

To fix ideas, first suppose that the designer allocates $x \leq d-1$ units to the central node, and $d-x$ units to $d-x$ peripheral nodes. If the adversary allocates c units to the central node and $a-c$ units to the peripheral nodes, then for large n, the payoff to the designer is given by:

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

Observe that this strategy is feasible for the adversary, and so the payoff to the designer is bounded above by this expression. More generally, allocating all resources to the central node is optimal if

$$\frac{d}{d+c} f(n-(a-c)) > \frac{x}{x+c} f(n-(a-c)) + \frac{c}{x+c} f(d-x). \tag{33}$$

Rearranging terms and simplifying, this is equivalent to:

$$\frac{d-x}{d+c} > \frac{f(d-x)}{f(n-(a-c))}. \tag{34}$$

Assumption (**A.1**) says that f is convex, increasing and that $f(n)=1$. So, for large enough n, the right side converges to 0; so it is optimal for the designer to allocate all resources to the central node. This argument can be extended to also cover the allocation of defence budget in which designer spreads his budget across more widely (and not just $d-x$ nodes).

Step 3: For network $g \neq g^s$, the payoff to the designer is bounded above by (31). Let us start with the case where adversary allocates all resources to the central node, i.e., $c=a$ in the star network. We establish that designer payoff in any other network is bounded above by $d/d+a$.

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

$$\min\{a_j, \frac{a_j}{a_j+d_j}\} \geq \min\{a_j, \frac{a_j}{a_j+d}\} = \frac{a}{a+d} \tag{35}$$

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 $a_i = \frac{a}{d} d_i$, $i \in \{1, .., k\}$. Notice that, for all i in $\{1, .., k\}$:

$$\min\{a_i, \frac{a_i}{a_i + d_i}\} = \frac{a_i}{a_i + d_i} = \frac{a}{a + d} \tag{36}$$

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))

$$E[f(n_1 I_1 + .. + n_k I_k)] \leq E[f((n_1 + .. + n_k) I_1)] = \frac{d}{a + d} \tag{37}$$

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 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 (37), this shows that the resulting payoff of $\mathcal{D}$ is less than he can guarantee himself with a star.

We now consider the case where the adversary allocates $c < a$ units to the central node. The designer's payoff in the star network is equal to $f(n - (a - c))d/(d + c)$. We establish that

payoff in any other network and with any defence strategy is bounded above by this payoff.

First, consider defence allocations which leave $(a-c)$ or more nodes unprotected. The adversary now has available the following feasible strategy: allocate $a-c$ units to $a-c$ unprotected nodes and c units among the protected nodes in such a way that the allocation on the protected nodes mimics the ratio c/d. In other words, allocate $a_i = d_i[c/d]$ units to every protected node with defence allocation d_i. Now it is straightforward to verify, using a variant of the arguments in step 3 above, that the design's payoff is bounded above by $f(n-(a-c))d/(d+c)$.

Second consider defence allocations in which all nodes have $d_i < 1$. Consider the minimally protected node; observe that the defence allocation on this node will approach 0 as n grows. The adversary can therefore eliminate this node with probability close to 1. The payoff to the designer approaches 0, and this is strictly below the payoff attainable in the star network with protected star.

Finally, consider defence allocations in which there are two or more nodes such that $d_i \geq 1$ while all other nodes are either unprotected or protected with defence budgets smaller than 1 (the case of one protected node with $d_i \geq 1$ is straightforward and omitted). A variant of the above argument now shows that there exist at least $a-c$ weakly protected nodes which can be eliminated with probability close to 1. Following our arguments in step 3 above, the payoff to the designer from the dispersed defence allocation across two or more nodes is strictly smaller than the payoff from protecting a central node in the star network. Putting together these observations yields the upper bound on designer payoffs as claimed.

■

Proof of Proposition 2: (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

$$\frac{1}{1+t} f(n-a+t) \tag{38}$$

Next, consider an arbitrary connected network g', in which a single node, α' say, is defended. Consider the following attack by $\mathcal{A}$: allocate t units of resource to attack node α', and 1 unit of resource to $a-t$ other nodes. Clearly, the resulting payoff of $\mathcal{D}$ is at most (38). 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

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

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

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

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

$$\frac{d_j}{d_j+1}f(n) + \frac{1}{d_j+1}f(n-\frac{d_j}{d}n) - \frac{d}{d+1}f(n) \tag{41}$$

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

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

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

$$\left[d_j(d+1)d^2-(d_j+1)d^3\right]+(d+1)(d-d_j)^2 \tag{43}$$

This simplifies to

$$(d-d_j)^2+dd_j(d_j-d) \tag{44}$$

And further to:

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

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

$$d(d-d_j)(1-d_j) \tag{46}$$

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

$$\frac{1}{2}f(n-1)+\frac{1}{2}f(n) \tag{47}$$

Consider next an arbitrary network, (say) g. If all nodes nodes are protected the maximum payoff to $\mathcal{D}$ is clearly given by (47). 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 (47). So both the complete and ring networks are robust.

■

Example 6: The cases with $d = 1$ or $a = 1$ are covered by Proposition 2. Next, it is easy to show that it is payoff weakly dominant for the designer to have: (i) no unprotected nodes between two protected nodes, and (ii) not have any link between two unprotected nodes. We

will exploit these observations in obtaining our characterization. Throughout, we refer to the resulting payoffs of $\mathcal{D}$.

d=2, a=2: The $\mathcal{A}$ can always reduce the payoff of designer to 1/4 by targeting two unprotected nodes. Moreover, the optimal strategy of $\mathcal{A}$ in the face of a star with all defence resources allocated to the central node (henceforth, center-protected (CP) star) is to target two peripheral nodes. This induces designer payoff equal to 1/4 and so the CP star is optimal.

d=2, a=3: The optimal strategy of $\mathcal{A}$ with a CP star is to target three peripheral nodes. The payoff to the designer is 1/16. With two defended nodes, the network is a line and the optimal strategy of $\mathcal{A}$ is to target two unprotected nodes and one defended node. The payoff to designer is $5/32 > 1/16$; thus two defended nodes and a line network is optimal.

d=2, a=4: The optimal strategy of $\mathcal{A}$ with a CP star consists in targeting all nodes. The payoff induced is 1/24. With two defended nodes, the optimal strategy of $\mathcal{A}$ again consists in targeting all nodes. The payoff induced however is $3/32 > 1/24$. 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: General order of moves

This appendix considers networks containing multiple components and 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. 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 $\mathcal{DDA}$ 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 orders of move and an arbitrary number of components.

Proposition 3 *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 at least $f(n)/2$. The argument is completed by showing that the maximum payoff of $\mathcal{D}$ in networks with two or more components is no more than $f(n)/2$.

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

$$\frac{1}{2}f(\overline{m} + \underline{m}) \geq 1/2f(\overline{m}) + f(\underline{m}) \tag{48}$$

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

$$\frac{1}{2}f(\frac{m}{2} + x) + f(\frac{m}{2} - x) \tag{49}$$

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 (48) 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, 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

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

By (48), this in turn is, at most

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

and, repeating the argument, at most

$$\frac{1}{2}f(n_k + n_{k-1} + .. + n_1) = \frac{1}{2}f(n) \tag{52}$$

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 4 *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 paths between i and at least half the nodes in the network which do not contain node j. We will refer to such a paths as j-independent paths.* 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 4.

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 4: 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: Generalized contest function

This section extends the model to allow for a generalized contest function on individual nodes. We show that the principal insights contained in Theorem 1 and Theorem 2 extend to this model.

Let us start by defining the general contest function: given a defence d_i and an attack a_i on node i, if $a_i + d_i > 0$ then the probability of successful attack is given by:

$$\frac{a_i^\gamma}{a_i^\gamma + d_i^\gamma} \tag{53}$$

for $\gamma > 0$. If $a_i = d_i = 0$, then the probability of successful attack is equal to 0. Let us refer to this as assumption (**A.2′**).

In an influential paper, Skaperdas (1996) showed that this is the unique contest function satisfying a set of plausible axioms on conflict. Observe that as γ grows we approximate a threshold attack function: attack is certain to succeed if and only if $a_i > d_i$, while it is certain to fail if and only if $a_i < d_i$. The probability of success is 1/2 at $a_i = d_i$, for all $\gamma > 0$

A simple way to circumvent the problem of very small attacks is to suppose that defence and attack can only take a value of 0 or values greater than 1. So $a_i, d_i \in \{0\} \cup X$, where $X = \{x \in \mathcal{R}_+ | x \geq 1\}$. Let this restriction of the strategy set be denoted by (**A.0**).

Let us characterize robust networks with this generalized contest function. We first observe that when $d = 0$, attack spreads instantaneously across a component. Given the restrictions on the allocation of attack resources, $a_i \in \{0\} \cup X$, $\mathcal{A}$ will assign one unit of resource to at most one node in every component. It is now possible to show that the arguments in the proof of Theorem 1 carry over directly.

Consider the design, defence and attack game next.

Theorem 2′: *Suppose assumptions (**A.0**)-(**A.1**) (**A.2′**) and (**A.3**) hold. Fix budgets $a, d \in \mathcal{N}$ and let $a \geq d$. For n sufficiently large, the star network is robust in the class of connected networks. In equilibrium, the designer allocates all his resources to the central node.*

Sketch of proof: The proof follows along the lines of Theorem 2, except in step 3, where it is significantly simpler, as the designer can protect at most d nodes and so $n - d$ nodes will be unprotected. For large n, $n - d > a - c$.

Step 1 follows as a consequence of (**A.0**). Step 2 shows that optimal defence involves allocating all defence resources to the central node. So the payoff to $\mathcal{D}$ in a center-protected star is $d^\gamma f(n - (a - c))/d^\gamma + a^\gamma$.

Step 3 is the key step in the proof and demonstrates that the payoff to $\mathcal{D}$ in an arbitrary network with multiple protected nodes is bounded above by $d^\gamma f(n - (a - c))/d^\gamma + a^\gamma$, where the adversary assigns c units to attack the central node. We observe that $\mathcal{A}$ can construct a strategy which mimics attack success rates at each of the defended nodes and also respects $a_i \in \{0\} \cup \{x \in R_+ | x \geq 1\}$, so long as $a \geq d$. This strategy, following the arguments in Theorem 2, leads to a stochastically dominating distribution of surviving networks as compared to the center-protected star network. Since payoffs of $\mathcal{D}$ are convex, the resulting

expected payoff of $\mathcal{D}$ is bounded above by $d^\gamma f(n-(a-c))/d^\gamma + a^\gamma$. Putting together steps 1-3 completes the proof.

■

9 Appendix D: Richer model of spread of attack

We now turn to a richer model of the spread of attack. The key feature is that successful attack resources at a node i now engages in contests with defence resources at neighboring node j. This formulation also addresses the discontinuity in indirect attack at $d = 1$ in the basic model.

Fix a network g and an allocation of defence $(d_1, ..., d_n)$ and attack resources $(a_1, a_2, ..., a_n)$. Suppose that in the contest at node i attack a_i prevail over the defence d_i. Then attack resources a_i are available for further attacks on neighboring nodes. If they fail in their attack then they are neutralized and removed from the network, while defence resources d_i remain intact. This first round of contests defines a set of captured nodes and a profile of residual defence resources which we refer to as $(d'_1, ..., d'_n)$. Observe that $d'_i \leq d_i$.

Now consider the follow up round of contests. Once attack resources a_i prevail over defence resources d_i at node i, they move to a neighboring node j. Similarly, surviving attack resources at other nodes also move to a neighboring node (the choice of node could be random or it is due to deliberate choice by the adversary). If the neighboring node j is defended with positive resources $d_j > 0$, then the resource engages in contest on the node. If not, it moves instantaneously across the undefended node to a neighboring node. This flow proceeds until it encounters a defended node. If there are no defended nodes then the attack takes over the network and the process ends at the empty network. If there are defended nodes then the flows of attack will stop within a finite set of steps and yield a new allocation of attack resources in the network, $(a'_1, a'_2, ..., a'_n)$. These attack resources engage in contests with defence resources at the different nodes $(d'_1, d'_2, ..., d'_n)$. The outcome of contests at nodes $1, 2, 3, .., n$ is in turn defined by the contest function (53).

This process continues way until there is no attack resource left or all nodes have been successfully attacked. We summarize the spread of attack as follows.

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 adversary relocates (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 resources 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.

Since any surviving attack resources move to un-captured nodes immediately, 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 $a + d$.

Anticipating the optimal attack strategy of $\mathcal{A}$, $\mathcal{D}$ chooses a network g and a defence strategy $\mathbf{d}$ to maximize his ex-ante expected payoffs at the start of the process.

Let us first consider the case $d = 0$, i.e., the pure design and attack problem. Observe that when $d = 0$, attack spreads instantaneously across a component. Given the restrictions on the allocation of attack resources, $a_i \in \{0\} \cup X$, $\mathcal{A}$ will assign one unit of attack to at most one node in every component. The arguments in the proof of Theorem 1 now carry over directly.

Let us turn to the design, defence and attack game.

Theorem 3 *Suppose assumptions (**A.0**), (**A.1**), (**A.2'**) and (**A.3'**) hold. Fix $a, d \geq 0$. For large enough n and γ,[14] the star network is approximately robust in the class of connected networks. The designer allocates all his resources to protecting the central node.*

Sketch of Proof: In a connected network, if $d < a$, then for large enough γ, $\mathcal{A}$ can prevail by starting with an assignment of all resources to a single node and then moving all resources across nodes one at a time until all defence resources are eliminated. So every node including the star network yield $\mathcal{D}$ zero payoff.

Next consider $a < d$. Under the assumption on allocations, $\mathcal{D}$ can protect a maximum of d nodes. So for, for large enough n, it must be the case that $\mathcal{A}$ can always eliminate a nodes. So the maximum payoff for $\mathcal{D}$ is $f(n-a)$. Consider the star network and suppose $\mathcal{D}$ allocates all resources to the central node. Given large enough γ, since $a < d$, $\mathcal{D}$ can ensure that the central node is protected with probability close to 1. Thus $\mathcal{D}$ earns a payoff *approximately* equal to $f(n-a)$ in a star network with defended central node.

[14] In the dynamic model where the adversary can choose where to re-allocate victorious attack resources, it may sometimes be more attractive to attack a node sequentially with small amount of resources rather than allocate all resources in a single attack. Large γ makes this sequencing unattractive and permits us to use a simple argument which is presented below.

Finally, consider $a = d$. Given that γ is large, allocating defence resources to two or more nodes allows $\mathcal{A}$ to eliminate each of the defended nodes, with probability close to 1. Thus a strategy of protecting multiple nodes yields a payoff of 0 to $\mathcal{D}$. Next consider a strategy of protecting a single node. In the star network with protected center, clearly the best strategy for $\mathcal{A}$ is to first target a peripheral nodes and then move the successful resources in a coordinated attack with all a units on the central node. This yields $\mathcal{D}$ a payoff of $f(n-a)/2$. Observe that the payoff to $\mathcal{D}$ is at most $f(n-a)/2$ in any network with a single protected node. So the star is robust.

■

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

11. 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.

12. Esteban, J. and D. Ray (2010), A model of ethnic conflict, *Journal of European Economic Association*, forthcoming.

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. (2008), *Social and economic networks.* Princeton University Press. Princeton. New Jersey.

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

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

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

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

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

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

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

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

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

37. Szentes, B. and R. W. Rosenthal (2003), Three-Object Two-Bidder Simultaneous Auctions: Chopsticks and Tetrahedra, *Games and Economic Behavior*, 44, 1, 114-33.

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

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

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

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

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

43. 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.

44. 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.

45. Vega-Redondo, F. (2007), *Complex social networks.* Cambridge University Press. Cambridge, England.

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