Robust Networks

Abstract

Connections between individuals facilitate the exchange of goods, resources and information and create benefits. However, these connections may be exploited by intelligent adversaries to spread their attacks as well. Links thus create value but also enable the spread of attacks which are harmful. How does this tension in the role of links shape network architecture and the nature of conflict?

We first study the pure design problem: a designer chooses a network and an adversary then chooses an attack strategy. An equilibrium network consists of equal size 'components' whose number grows (and size falls) with the attack budget of the adversary.

We then suppose that, in addition, the designer has defence resources to protect the network. If defence and attack budgets are small, relative to the number of nodes, the equilibrium network is a star; the designer allocates all resources to protect, while the adversary allocates all resources to attack, the central node. If defence and attack budgets are large then equilibrium networks are dense and dispersed defence allocation is more effective.

1 Introduction

Connections between individuals facilitate the exchange of goods, resources and information and create benefits. However, these connections can also be exploited by intelligent adversaries to spread their attacks, as the following examples illustrate.

- The Internet is a network of local area networks (LANs). The LAN's – examples include interconnected computers in a university or a firm – facilitate communication between member users. The LAN manager decides on connections between the computers and the security settings on them (anti-virus software, Firewall etc). A hacker chooses which computer(s) to attack with a virus. If a computer is infected, the virus may be forwarded to the mailing list – viz. the connections – of the infected host computer.[1]

- Criminals/terrorists communicate to coordinate their actions. Communication requires connections – knowledge of identity, skills, addresses and telephone numbers. But connections also make the gang more vulnerable to the police and intelligence agencies: the detection of one person may lead, via interrogation and traces of communication, to the capture of other gang members.[2]

What are the implications of this trade-off in connections for the design of networks and the nature of conflict in networks?

We consider a game with 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 this network design and protection plan and then allocates his attack resources across the nodes.[3] The probability of successful attack on a

[1] In 2009, it was estimated that roughly 10 million computers were infected with malware designed to steal online credentials. In Europe, the annual damages caused by malware are of the order of 9.3 billion Euros, while in the US the annual costs of identity theft are estimated at 2.8 billion USD (Moore, Clayton and Anderson (2009)). These large losses have fueled the demand for computer and network security. One indicator of the size of this market is the valuation of security firms: Intel bought McAfee in 2010, for 7.68 billion USD (bbc.co.uk; 19 August 2010).

[2] The conflict between covert intelligence activity, one the one hand, and terrorist and criminal organizations, on the other hand, is an old one. The September 11, 2001 attacks in New York and the ensuring public debate about terrorism has highlighted this conflict. The US government has allocated over 400 billion USD and new institutions – such as the Counter-terrorism section at the US department of Justice – have been created to combat terrorism (Baccara and Bar-Isaac (2008) and Zakaria (2008).

[3] Section 2 describes how our assumptions on conflict and timing of moves between designer and adversary are appropriate in the study of computer network security. A game with simultaneous moves appears to be more appropriate for the study of crime and covert intelligence. Section 4 develops analogues of our main results in a game with simultaneous moves.

node depends on the defence and attack resources allocated to it. Following Tullock (1980), we suppose that this probability of attack is increasing in the attack resources and decreasing in the defence resources allocated to the node. The other key element is the spread of attack from one node to a neighboring node: we suppose that the probability of spread of attack to a neighboring node is falling in the amount of defence resources allocated to the neighboring node.

The conflict between defence and attack resources and the spread of successful attacks eventually yields a collection of surviving network of (healthy) nodes and links. This surviving network defines the payoffs of the two players. The payoffs to the designer from a network are equal to the sum of the returns from the different components and the returns from a component are increasing and convex in its size.[4] We study the (sub-game perfect) equilibrium of the game between the designer and the adversary.

We begin with an analysis of the pure design game, i.e., when the designer has zero defence resources. As there are no defence resources, the only way in which the designer can contain the spread of a successful attack on a node is by separating it from other nodes: in other words, by breaking 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 on the relation between adversary budgets, on the one hand, and network size and success of the network, on the other hand, echo discussions in the popular press with regard to terrorist networks. 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... Terrorists have not been able to hit big, symbolic targets.... So they blow up bombs in cafes, marketplaces, and subway stations" (Zakaria (2008, p. 3))

We then turn to the study of the general problem of design and defence of networks which face an adversary. If defence and attack resources are small relative to the number of nodes then a star network in which the designer and adversary allocate all resources to the central

[4]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 – the empty network – 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.

node is an equilibrium (Theorem 2).[5] On the other hand, if defence and attack budgets are large then an equilibrium network is dense and dispersed defence allocation is more effective.

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 of a slight 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 choose to concentrate all his resources on the central node. Similar considerations suggest that it is attractive for the adversary to allocate all resources to attacking the central node. The probability of the entire network surviving is then simply equal to the probability of the central node surviving and this is proportional to the relative magnitude of defence and attack.

Next consider a strategy of creating a core-periphery network and defending the multiple central nodes.[6] Faced with such a network, the adversary can mimic the proportions on the central node of a star network.[7] The key theoretical observation is that the distribution of surviving nodes in the star network is a mean preserving spread of the distribution in the core-periphery network. Since the payoffs of the designer are convex in the size of a component, the designer obtains higher payoffs (and the adversary lower payoffs) in the star network. A comparison of Figures 3 and 4 illustrates this point. Our proof shows that the argument can be extended beyond core-periphery networks to cover general network architectures.

Theorem 2's prediction that resources will be concentrated at key nodes of the network provides us a possible explanation for the prominent role of the Firewall in internet security. In his well known book on security engineering, Ross Anderson writes, "The most widely sold solution to the problem of internet security is the Firewall. This is a machine that stands between a local network and the Internet, and filters out traffic that might be harmful. The idea of a 'solution in a box' has great appeal to many organizations, and is now so widely accepted that it is seen as an essential part of corporate due diligence" (Anderson (2008, p. 654)).

[5] A star network has a central node which is linked to all other nodes, and there are no other connections in the network. Figure 3 illustrate a star node.

[6] A *core-periphery* network structure has two groups of nodes, the core and the periphery. The core nodes are fully linked among themselves, while the periphery nodes have a single link with one of the core nodes. Figure 4 illustrates a core-periphery network with two core codes.

[7] So, for instance, suppose the designer has 4 units and adversary has 8 units of resources. If the designer allocates 3 and 1 to two central nodes of the core-periphery network the adversary can allocate 6 and 2 units to the two central nodes, respectively. The ratio on each of the two central nodes – one-half – mimics the ratio on the single central node of the star network.

The principal contribution of this paper is to propose and solve a model of conflict in networks. In doing so, we build on and contribute to two rich strands of research: the theory of networks and the theory of conflict/contests.[8]

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 distinguished tradition of research in communication networks, see e.g., Bolton and Dewatripont (1994), Radner (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.[9] 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. There are two main differences between our paper and their's: one, we study design and defence of networks, while they focus on the pure design problem. Two, in our paper networks facilitate communication and exchange while in theire paper, networks facilitate cooperation. 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; see e.g., Hirshliefer (1983), Tullock (1967, 1980), Sandler and Hartley (2007), Shubik and Weber (1981), Esteban and Ray (2010), Dixit (1987), Hirshleifer (1991), Skaperdas (1996), Baye (1998), Clark and Konrad (2007), and Kovenock and Roberson (2009). 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 as conflict functions and budgets vary. Our paper uses the Tullock contest 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.

[8]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). 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.

[9]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.

The rest of the paper is organized as follows. Section 2 presents our model of design, defence and attack in networks; it also elaborates on the computer network security application to illustrate the appropriateness of our assumptions. Section 3 contains the equilibrium analysis. Section 4 studies alternative timing of moves between the designer and the adversary and develops analogues of our main results to different settings. Section 5 concludes. Appendix A contains the proofs of our main results, while appendices B and C discuss and provide new results under alternative modeling assumptions.

2 A Model

We study a zero-sum two player game between a designer and an adversary. The designer has a collection of nodes and a budget for defense, while the adversary has an attack budget. The designer moves first and chooses links between the nodes and 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.[10] The initial network design and the subsequent conflict together define a probability distribution on surviving networks. We now set out the notation and the concepts which formally describe this game.

The designer: The designer, $\mathcal{D}$, has a collection of nodes $N = \{1, ..., n\}$, where $n \geq 2$; for expositional simplicity, suppose that n is an even number. $\mathcal{D}$ chooses links between the nodes and allocates a defence budget $d \in \mathcal{N}$ across the nodes to protect the network. Let $\mathbf{d} = (d_1, d_2, ..., d_n)$ denote this allocation, where $d_i \geq 0$ and $\sum_{i \in N} d_i \leq d$.

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 undirected, i.e. $g_{ij} = g_{ji}$. The links between the different pairs of individuals define a network g.

There is a path between two nodes i and j in network g if there exists a sequence of nodes $i_1, .., i_k$ such that $i_1 = i$, $i_k = j$ and $g_{i_1 i_2} = ... = g_{i_{k-1} i_k} = 1$. 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 maximally connected subset of nodes. $\mathcal{C}(g)$ is the set of components of g. We let $|C_k|$ indicate 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 $\mathcal{C}(g)$.[11]

[10] Section 4 studies alternative timing of moves and develops analogues of our main results, Theorems 1 and 2.

[11] The complete network, g^c, has $g_{ij} = 1, \forall i, j$. The empty network, g^e, has $g_{ij} = 0$ for all i and j. A core-periphery network has two types of nodes, N_1 and N_2. Nodes in N_1 constitute the periphery and have a

Following Myerson (1977), we assume that the value of a network is the sum of the value of the different components. Given n, let $f_n(m)$ denote the value from a component of size m. If $f_n(.)$ is decreasing, or increasing and concave, then an empty network maximizes value. As our interest is in the tension between the pressure to connect nodes to create value and the threat of contagion of attack via connections, it is reasonable to restrict attention to increasing and convex $f_n(.)$ functions. These considerations lead to:

Assumption A.1: *The value of network g is given by*

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

where $f_n(.)$ is increasing, strictly convex, $f_n(0) = 0$ and $f_n(n) = 1$.

By way of illustration, consider the following example.

Example 1 *A communication network*[12]

Suppose every individual has one piece of information with value 1, to everyone. A link between X and Y allows X to access Y's information as well as information which Y may have accessed via his links with others. In a network g, X has access to all others in his component C_k; his payoff is $|C_k|/n^2$ (where the denominator reflects a normalization). As there are $|C_k|$ nodes in the component, the total payoff in component C_k is $|C_k|^2/n^2$. The aggregate social payoff in a network is the sum of the payoffs from the different components:

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

The payoffs given in (2) satisfy (**A.1**). △

The adversary: The adversary $\mathcal{A}$ has $a \in \mathcal{N}$ attack resources. He observes the network g and the defence allocation $\mathbf{d}$ and then allocates his resources across the n nodes. Let $\mathbf{a} = (a_1, a_2, ..., a_n)$ denote the allocation chosen by $\mathcal{A}$, where $a_i \geq 0$ and $\sum_{i \in N} a_i \leq a$.

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. When the core contains a single node, we have a star network. For a a general introduction to network terminology, see Goyal (2007).

[12]This is a variation on the connections model, developed in Goyal (1993), Bala and Goyal (2000) and Jackson and Wolinsky (1996).

Consider the conflict between defence resources d_i and attack resources a_i which are allocated on node i. Our formulation builds on the Tullock (1980) contest function.[13] We would like the probability of successful attack to be small if the attack resources allocated are small, even if there is no defense available. This motivates:

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: Skaperdas (1996) shows that under reasonable axioms on the nature of conflict, the probability of successful attack on node i is $a_i^\gamma/(a_i^\gamma + d_i^\gamma)$, with $a_i, d_i > 0$ and $\gamma > 0$. Appendix B shows that if $a_i, d_i \in \{0\} \cup X$, where $X = \{x \in \mathcal{R}_+ | x \geq 1\}$, then our principal results (Theorems 1 and 2) obtain under this general contest function.

Consider next the spread of attack through a network. We suppose that an attack from i can spread to j only if there is a path between the two nodes. We next suppose that the likelihood of an attack spreading from node i to node j is declining in the defence resources allocated to any node on the path between i and j. In the basic model, we assume that a successful attack on node i spreads to a neighboring node k, if and only if $d_k < t$, for some $t \in \mathcal{R}$. And, for simplicity we set $t = 1$. A path between two nodes i and j is said to be *weak* if and only if $d_k < 1$ for all nodes on the path from i to j.

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.*

Remark 2: An alternative (and smoother) model of spread of attack would be as follows: successful attack resources move to neighboring nodes and engage in (general) Tullock contests with defence resources on those nodes. Appendix B develops conditions under which our main results, Theorem 1 and 2, extend to this setting.

Figure 1 illustrates conflict and the spread of attack across the network. The upper half of the figure presents a ring network in which d non-adjacent nodes are allocated one unit of defence resources each. If $\mathcal{A}$ allocates a unit of attack resources on a node between the protected nodes then attack spreads and eliminates all nodes except for the protected nodes. The lower half of the figure presents a star network with all defence resources allocated on the

[13] With this contest function, the probability of successful attack on a node i is $\frac{a_i}{a_i+d_i}$, for $a_i \geq 0, d_i > 0$.

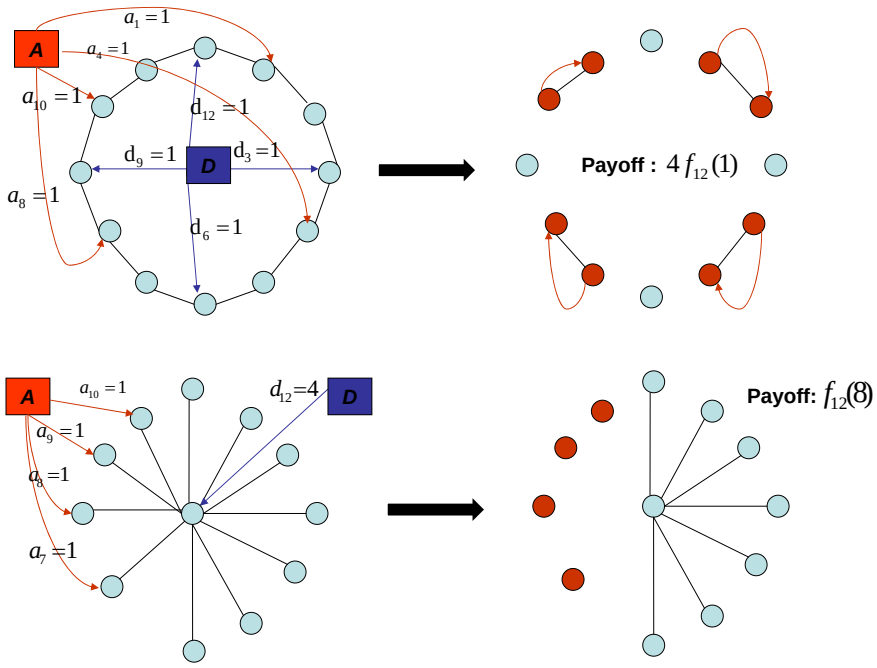

Figure 1: Defense and attack in a network: $n = 12, a = d = 4$.

central node. Suppose $\mathcal{A}$ allocates one unit to each of a periphery nodes: then the a nodes are eliminated but there is no spread of attack through the network (due to the presence of the central node which is protected above the threshold level).

Payoffs: A network g' on N' is a sub-graph of g if $g_{ij} = 1$ whenever $g'_{ij} = 1$. Let $G(g)$ denote the set of sub-graphs of g. Given a network g, and allocations $\mathbf{d}$ and $\mathbf{a}$, assumptions (**A.2**) and (**A.3**) define a probability distribution on the set of sub-graphs $g' \in G(g)$ which survive the attack. Let $P(g'|\mathbf{a}, \mathbf{d}, g)$ be the probability that network g' survives the conflict between $\mathbf{d}$ and $\mathbf{a}$ on network g. Then:

$$\sum_{g' \in G(g)} P(g'|\mathbf{a}, \mathbf{d}, g) = 1 \tag{3}$$

The expected payoff to $\mathcal{D}$ from strategy $(g, \mathbf{d})$ when $\mathcal{A}$ chooses $\mathbf{a}$ is:

$$\sum_{g' \in G(g)} P(g'|\mathbf{a}, \mathbf{d}, g) \left[\sum_{C_k \in \mathcal{C}(g')} f_n(|C_k(g')|) \right] \tag{4}$$

So, in Figure 1, the expected payoff to the designer in the ring network is $4f_{12}(1)$, and in the star network it is $f_{12}(8)$. We assume that the objectives of the designer and the adversary are perfectly opposed, i.e., the game is zero-sum. Our analysis focuses on the (sub-game perfect) equilibrium of this zero-sum game.

We conclude this section with an application: the aim is to show that our assumptions are appropriate for the study of conflict in networks.

2.1 Application: computer network security

The Internet is a network of local area networks (LANs). The local area networks – examples include interconnected computers in a university, a firm, or a government department – facilitate communication between users. The returns from joining a network are increasing in the number of users. This is in line with our assumption (**A.1**) of convex and increasing returns in network size.

The network manager decides on connections between computers and makes investments in security (anti-virus software, Firewall etc). Online criminals – such as hackers and 'botnet' herders – take this network structure and the security settings as given, when they attack the computers. This is consistent with the strategic options and the timing of moves in our game.

Attack on a computer may take the form of a virus or a worm (self-propagating malicious software). The likelihood of successful infection of the computer is lower the more sophisticated the security installations on it. This is reflected in our contest function formulation, (**A.2**). Upon successful infection of a computer, the virus delivers payload: it corrupts (progressively larger portions of) the host computer.

Having corrupted its host, the virus can attach itself and be forwarded to the mailing list of the infected host computer. The probability that the virus succeeds in infecting neighboring computers varies with the level of security installations on those computers. The transmission of virus via communication links and subsequent conflict between the virus and the security installed on neighboring computers is consistent with our model of spread of attack through a network (see (**A.3**) and the following remarks).

3 Equilibrium analysis

We start with a consideration of the pure design problem, i.e., when $d = 0$. The main result is that, in equilibrium, the adversary targets at most one node in each component and the network consists of equal size components whose number grows (and size falls) as the attack budget increases. We then turn to the problem of design with positive defence resources. If defence and attack resources are small, relative to the number of nodes, then the star is an equilibrium network; the designer assigns all resources to defend, and the adversary assigns all

resources to attack, the central node. If defence and attack budgets are large then equilibrium networks are dense and dispersed defence allocation is more effective.

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 $\prod_{i \in C_k}(1 - a_i)$.[14] Then, component additivity implies that the payoff of $\mathcal{D}$, facing attack $\mathbf{a}$, is:

$$\sum_{C_k \in \mathcal{C}(g)} f_n(|C_k(g)|) \prod_{i \in C_k(g)} (1 - a_i) \tag{5}$$

Our first result characterizes optimal attack strategies and equilibrium networks in this design game.

Theorem 1 *Suppose (**A.1**)-(**A.3**) hold and $d = 0$. (i) If $a < n/2$ then, in equilibrium, the adversary targets at most one node in any component; the network contains at least $a + 1$ maximal components and at most one component which is smaller. (ii) If $n/2 \leq a < n - 1$, the empty network is the unique equilibrium outcome. (iii) If $a \geq n$ then, in equilibrium, the adversary eliminates all nodes. Any network can be sustained in equilibrium.*

Proof: If $a \geq n$ then the adversary can always eliminate all nodes, irrespective of the structure of the network. So, it follows that the designer earns a payoffs of 0 irrespective of the network. Hence, any network can be sustained in equilibrium. Now let us take up the case of $a \leq n-1$.

First, we note that 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.

Second, 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 maximal, 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' as compared to g, since by doing so she saves the node which has been isolated. This contradicts the hypothesis that g is optimal.

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

Third, 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 maximal components, and it now follows from the convexity of $f_n(.)$ that payoffs to $\mathcal{D}$ are strictly increased by this move.

Fourth, at most one node is attacked in a component. Observe there are always more components than adversary budget. So each component is assigned at most one unit of attack resource. If $\mathcal{A}$ attacks two nodes, there is positive probability of a state in which both nodes are eliminated and a corresponding state in which neither is eliminated: this is wasteful as elimination of one node is sufficient to remove the entire component.

Finally, observe that if $a \geq n/2$ then $\mathcal{A}$ can always eliminate every component with 2 or more nodes. Hence, the empty network is the unique equilibrium outcome. ■

So, loosely speaking, the number of components is (weakly) increasing and the size of each component is falling in the budget of the adversary. An increase in the returns from size of component is reflected in an increase in 'convexity' of the returns function. Given a fixed adversary budget, such an increase in convexity makes larger components more attractive at the margin. Figure 2 illustrates these comparative statics for a specific returns function: $f_n(m) = m^\alpha/n^\alpha$, with $\alpha > 1$ reflecting the 'convexity' of the function. In this figure, k refers to the number of components.

We now turn to the general problem of designing and defending a network which faces an intelligent adversary. For simplicity, we assume that the value of a node becomes insignificant as the number of nodes grows:

Assumption A.L Given $x \in \mathcal{R}$,

$$\lim_{n\to\infty} f_n(n-x) = 1 \tag{6}$$

Our main result in the design, defence and attack game is:

Theorem 2 *Suppose (**A.1**)-(**A.3**) and (**A.L**) hold. Fix $a, d \in \mathcal{N}$ and consider the set of connected networks. For n sufficiently large, there is an equilibrium with the star network; and in this equilibrium, the designer and adversary allocate all resources to the central node.*

The proof of this result is given in Appendix A. It builds on three arguments. *First*, we show that given a star network in which all defence resources are allocated to the central

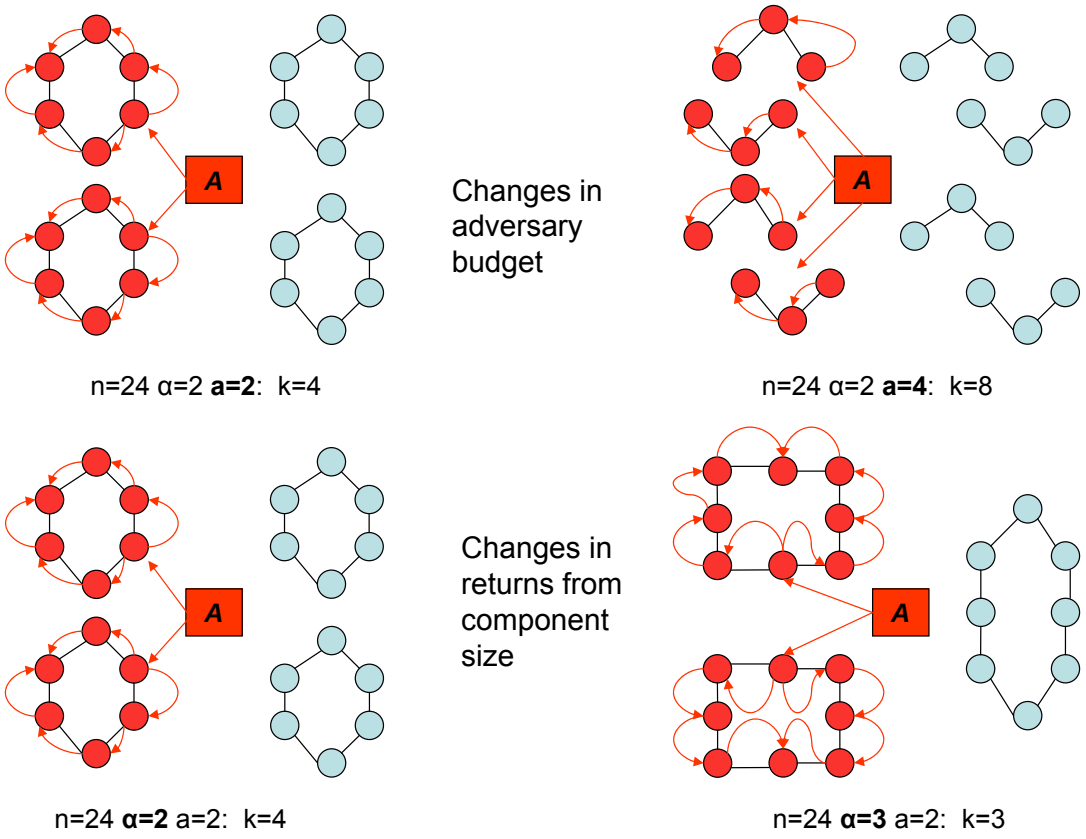

Figure 2: Equilibrium networks: $f_n(m) = (m/n)^\alpha$, $n = 24$, α=2,3 and $a = 2, 4$

node, if $\mathcal{A}$ allocates c units of resource to the center, then he will allocate the remaining $a - c$ units on $a - c$ peripheral nodes, in other words, it is optimal for adversary to concentrate the attack on $a - c$ peripheral nodes. Observe that, due to (**A.3**), a center-protected star ensures connectedness of all surviving nodes. So, spreading resources across more peripheral nodes yields a mean-preserving spread in the distribution of interconnected nodes. Given the convexity of $f_n(.)$, it then follows that spreading resources on more peripheral nodes raises the payoff for the designer and, correspondingly, lowers the payoff for the adversary.

The *second* step shows that the designer ensures himself a payoff of $d/(d + a)$ in a star network. Suppose $\mathcal{D}$ allocates all his resources to protect the central node. If $\mathcal{A}$ allocates k units of resource to peripheral nodes, then he eliminates k of them for sure. But this allocation of k units away from the central node lowers the probability of eliminating it correspondingly. Under assumption (**A.L**), for large enough n, the marginal value of eliminating a single node is very small compared to the value of a slight increase in the probability of eliminating the central node (and hence all nodes). Thus it is optimal for $\mathcal{A}$ to allocate all resources to the central node. But this implies that, with a star network, $\mathcal{D}$ can ensure himself a payoff of $d/(d + a)$.

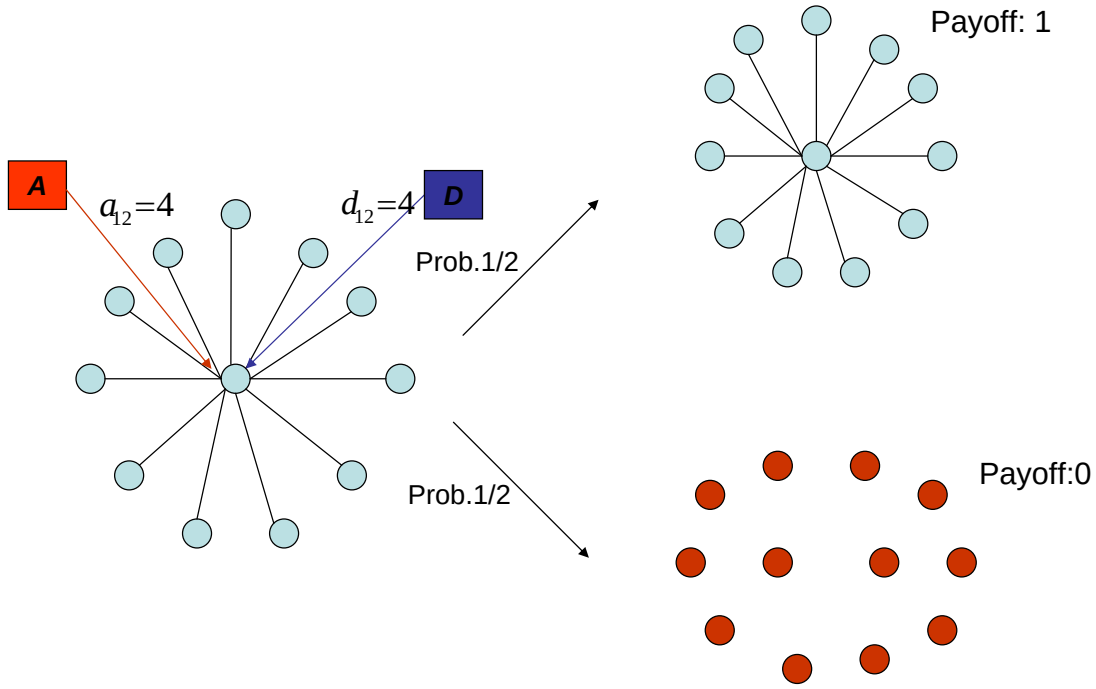

Figure 3: Attack & defence on a star

The *third* step constitutes the heart of the proof: we show that in any core-periphery network with all core nodes protected, the payoff of the designer is less than $d/(d+a)$. Fix a core-periphery network with k core nodes and consider a defence strategy $d = (d_1, d_2, ..d_k)$. The adversary has a feasible strategy in which he allocates $a_i = d_i \times (a/d)$ to each of these core nodes. This configuration of defence and attack resources creates k identical and independent lotteries each of which has probability of defense success $d/(d+a)$. We then apply Theorem 1 in Rothschild and Stiglitz (1970) to conclude that the probability distribution of the number of surviving interconnected nodes in the star network is a mean-preserving spread of the probability distribution of surviving nodes under k identical and independent contests in the core-periphery network. Given that the returns function $f_n(.)$ is convex, it then follows from standard arguments that the designer prefers the star network to such a core-periphery network. Figures 3 and 4 illustrate this argument with $n = 12$ and $a = d = 4$.[15]

This argument is then extended beyond core-periphery networks and shows that the payoff in any connected network is (weakly) lower than $d/(d+a)$. Roughly speaking, there are two main points here. One, it is better for the designer to connect every pair of protected nodes. This rules out the state in which subsets of protected nodes are separated into distinct

[15]Observe that the probability distribution of surviving component sizes for the 2-core network is Prob(0 nodes) = 1/4, Prob(6 nodes) = 1/2, and Prob(12 nodes) = 1/4. The probability distribution for the star network is Prob(0 nodes) = 1/2 and Prob(12 nodes) = 1/2. The latter distribution is a mean-preserving spread of the former one.

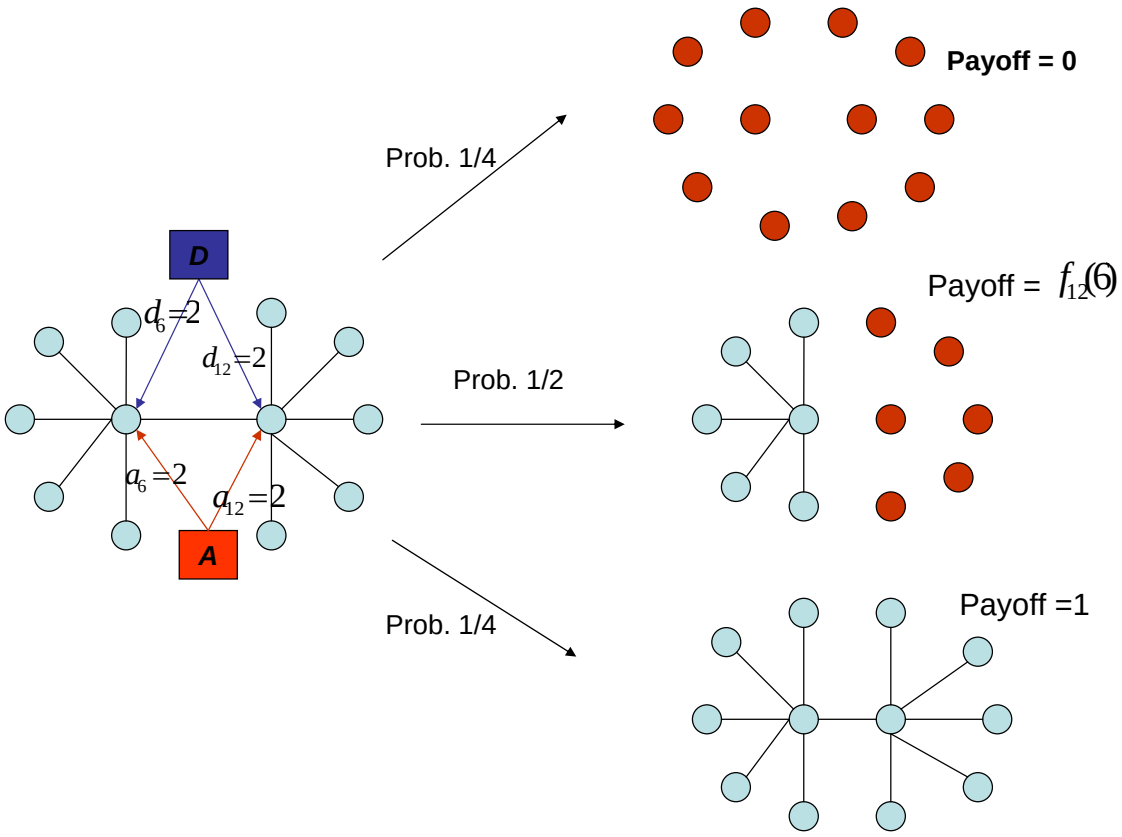

Figure 4: Core-periphery network: attacking the core nodes

components. Two, we show that it is better for the designer to connect a non-protected node to only one protected node and not to connect it to any unprotected node. This minimizes the probability that the unprotected node is indirectly attacked. These two points, taken together, imply that in a strategy which protects k nodes, the designer is best off with a core-periphery network.

Empirical work on networks has highlighted the salience of hub-spoke structures (see e.g., Barabasi (1999) and Goyal (2007)). In an influential paper, Albert, Jeong and Barabasi (2000) argue (using a set of simulations) that hub-spoke architectures are robust to random attacks but vulnerable to strategic attacks: the adversary can significantly reduce a hub-spoke network's functionality by removing only a few hub nodes. Our result, on the other hand, highlights the attractiveness of hub-spoke architectures from the point of view of defence. Successful defence of a few hub nodes contains the spread of attacks through the network. Thus our work provides an efficiency based foundation for the salience of hubs in real world networks which face adversaries.

We now comment on three hypotheses in Theorem 2: *one*, the vanishing value of a node as the number of nodes grows, (**A.L**), *two*, the connectedness of the network, and *three*, the role of the relative magnitude of resources and the number of nodes.

Let us *first* take up assumption (**A.L**). This assumption rules out applications in which a single node or a few nodes remain important when the total number of nodes grows. The literature on contests and conflict studies games in which each node is critical for the functioning of the system at large. They are referred to as *weakest link games*. Recent treatments include Kovenock and Roberson (2009) and Clark and Konrad (2007). Suppose $f_n(x)$ takes on value 0 for all $x < n$ and takes value 1 for $x = n$. Let us refer to this as assumption (**A.1′**). Observe now that the connectedness of the network is necessary for the designer. On the other hand, once a single node is infected the designer's payoff is zero and no further damage is possible via the spread of attack. Thus the weakest link game does not exhibit the tension – between the benefits and the costs of connections – which is the focus of our paper. Next, observe that, if $a \geq 1$, then optimal defense involves a dispersed allocation of resources.[16]

The *second* hypothesis concerns the connectedness of the network. If $f_n(.)$ is sufficiently convex then $\mathcal{D}$ will always choose a connected network. If, on the other hand, $f_n(.)$ is (approximately) linear then multiple components will be more attractive for the designer.[17]

The *third* hypothesis is about the relative magnitude of the defence and attack resources and the number of nodes. We have been unable to characterize equilibrium networks and defence and attack strategies for all values of a, d, and n. To make progress we study a game with a specific returns function taken from Example 1: $f_n(m) = m^2/n^2$. Proposition 1 in Appendix A characterizes equilibrium networks for general n and for $a = 1$, $d \geq 1$, and when $d = 1$ and $a \geq 1$, respectively. The following example highlights some of the issues which arise when budgets are large relative to the number of nodes.

Example 2 *The attractions of dense networks and dispersed defence.*

Consider the communication network example with payoffs given as in (2). Suppose that $n = 4$, $d = 4$ and $a = 1$. Fix 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 $f_4(3) = 9/16$. 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

[16] As the number of nodes gets large, the defence resources assigned to any node become progressively smaller, and the payoffs of the designer fall accordingly. This is in line with the findings of Clark and Konrad (2007). For a result on equilibrium networks and optimal defence and attack strategies in the absence of Assumption (**A.L**), refer to the working paper version of our paper, Goyal and Vigier (2010).

[17] In an earlier version of our paper we developed a set of sufficient conditions on the returns function under which the equilibrium network is connected, see Goyal and Vigier (2010).

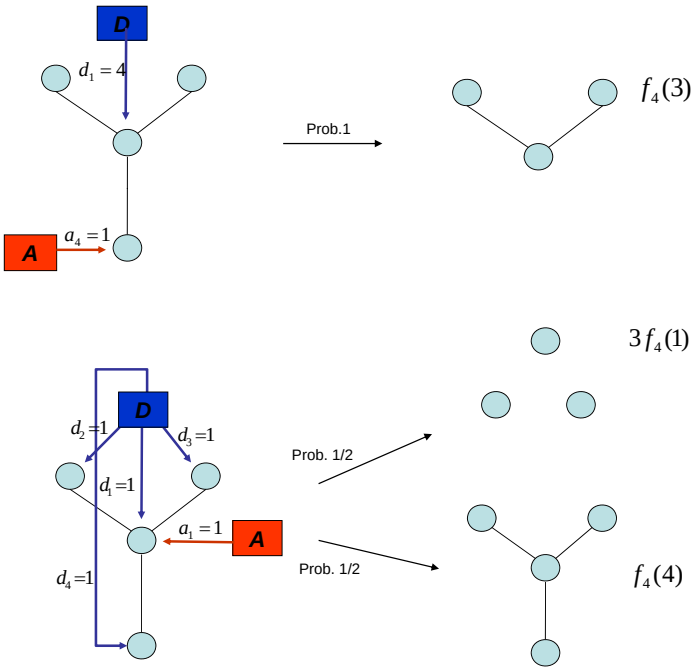

Figure 5: The attractions of dispersed defence

of $\mathcal{D}$ is $3f_4(1)/2 + f_4(4)/2 = 19/32$. So, a dispersed defence strategy is better for $\mathcal{D}$. Figure 5 illustrates this point.

Next, consider a complete network and suppose $d = 4, a = 1$. $\mathcal{D}$ can do better with the complete network. If $\mathcal{D}$ protects all nodes in the complete network then his minimum payoff is $f_4(3)/2 + f_4(4)/2 = 25/32$. This shows that the complete network with dispersed protection strictly dominates a star. Figures 6 illustrates this point.[18]

△

4 Discussion: timing of moves

In the basic model, we study a sequential move game in which the designer moves first, followed by the adversary. In some contexts, such as crime and terrorism, the designer seeks to conceal the network structure while intelligence agencies carry out covert operations to identify and eliminate members. The simultaneous move game appears to be a more natural framework for the analysis of this problem.[19]

[18] For a complete characterization of equilibrium architectures in this example, see the working paper version of our paper, Goyal and Vigier (2010).

[19] An alternative is the game in which the designer first chooses a network and the designer and adversary then simultaneously choose the defence and allocation resources on the network. The pure design problem is exactly the same as in the benchmark model; the defence and attack game is explored, and a partial analogue

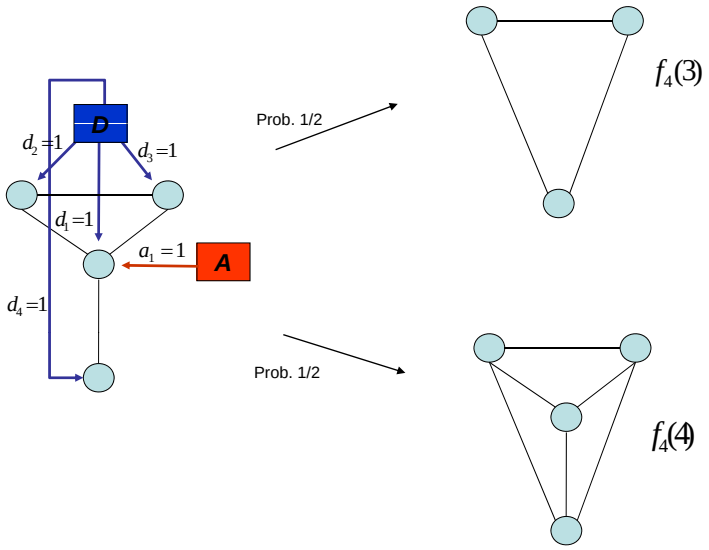

Figure 6: Defence and attack on complete network

We start with a consideration of the pure design problem (the defence budget is 0). Here we establish a partial analogue of Theorem 1: in the class of equal component networks, in equilibrium, the designer randomizes assignment of nodes across components; moreover, the number of components is increasing and the size of each component is falling in the budget of the adversary. The adversary's strategy involves mixing across different nodes. We then study the game of defence, design and attack and prove a partial analogue to Theorem 2: a strategy of mixing across different star networks with all resources devoted to the central node, is approximately optimal. The adversary's randomizes his attacks across across nodes.

Let us start with the case $d = 0$: The designer chooses a network and the adversary allocates a units of resources (for simplicity, we restrict attention to integer allocations) across nodes. As in the basic model, suppose that assumptions (**A.1**)-(**A.3**) hold. A mixed strategy of the designer is a probability distribution, σ, on the set of networks and defence allocations $\mathcal{G} \times \mathbf{D}$. Similarly, the mixed strategy of the adversary, ρ, is a probability distribution on the set of attack allocations $\mathbf{A}$. The expected payoffs to $\mathcal{D}$ from strategy σ when $\mathcal{A}$ chooses ρ are:

$$\sum_{(g,\mathbf{d})\in \text{supp } \sigma;\mathbf{a}\in \text{supp } \rho} \sigma(g,\mathbf{d})\rho(\mathbf{a}) \sum_{g'\in G(g)} P(g'|\mathbf{a},\mathbf{d},g) \left[\sum_{C_k\in\mathcal{C}(g')} f_n(|C_k(g')|) \right] \tag{7}$$

to Theorem 2 is developed, in Appendix C.

Networks with equal size components provide a useful analytical benchmark.[20] Suppose the adversary allocates one unit of resource to a nodes chosen uniformly at random from the set of nodes. By (**A.2**)-(**A.3**) the probability that a component of size m survives is:

$$\left(\begin{array}{c} n-m \\ a \end{array} \right) \Big/ \left(\begin{array}{c} n \\ a \end{array} \right) = \frac{(n-m)!}{(n)!} \frac{(n-a)!}{(n-m-a)!} \tag{8}$$

The total payoffs of the designer from a network with all components having equal size s is then:

$$\frac{n}{s} \frac{(n-s)!}{(n)!} \frac{(n-a)!}{(n-s-a)!} f_n(s) \tag{9}$$

For n large this becomes, approximately:

$$\frac{n}{s} n^{-s} (n-a)^s f_n(s) = \frac{n}{s} f_n(s) (1 - \frac{a}{n})^s \tag{10}$$

Abstracting from integer constraints, the optimal size, s^*, solves the following equation:

$$s^* \left[f_n'(s^*) + f_n(s^*) \ln(1 - \frac{a}{n}) \right] - f_n(s^*) = 0. \tag{11}$$

For this to be a maximum, the second order condition must be satisfied: $f''(s) + sf'(s) \ln(1 - a/n) + f_n(s) \ln(1 - a/n) < 0$. It is now easy to see that there is an equilibrium in which the designer randomizes uniformly over all networks with equal size components s^*, while the attacker allocates one unit of resource to a nodes chosen uniformly at random.

The optimality of mixing across links suggests that flexible networks are attractive for criminal and terrorist organizations. This is in line with the influential argument that in modern conflict, flexible networks – which permit quick reconfiguration of connections between individuals – will be an important organizational form (see e.g., Arquilla and Ronfeldt (1996, 2001)).

Now consider the effect of a change in adversary budget a on equilibrium component size

[20] Equal size components is a restriction in this setting. Convexity in returns to size makes a network with unequal size components attractive in the simultaneous game. Suppose $n = 4$, $a = 1$, $f_4(1) = f_4(2) = 0$, $f_4(3) = \frac{1}{2}$, and $f_4(4) = 1$. It can be shown that in the unique equilibrium the designer randomizes uniformly over all networks with two components of respective size 3 and 1, while the attacker allocates his unit resource uniformly at random. In the sequential move game any attempt to exploit convexity by choosing unequal components is defeated by the elimination of the larger component by the adversary. When players move simultaneously the designer can conceal the identity of nodes in the larger components.

s^*.

$$\frac{\mathrm{ds}^*}{\mathrm{da}} = \frac{s^* f_n(s^*) n^{-1}(1-a/n)^{-1}}{f''(s^*) + s^* f'(s^*)\ln(1-a/n) + f_n(s^*)\ln(1-a/n)} \tag{12}$$

The numerator is positive while the denominator is negative (from the second order condition). So a growth in adversary budget leads to a fall in component size. This result is consistent with the observation that terrorist networks have had to "... scatter, work in small cells", as the intelligence budgets have grown (Zakaria 2008).

We turn next to the game with positive defence resources. Define σ^* as a mixed strategy of the designer which randomizes uniformly over all star networks with defense resources concentrated on the central node. We will compute a lower bound on the designer's payoff from such a strategy.

Let ρ denote a mixed strategy for the adversary. Given the integer allocation restriction, the maximum number of nodes with $a_i > 0$ is equal to a. So, given σ^*, the maximum probability that the central node is attacked is $\frac{a}{n}$. Now observe that the payoff is at least $f_n(n-a)$ if the central node is not attacked. So it follows that if $d \geq 1$ then the minimum expected payoff of the designer given σ^* is at least $(1-\frac{a}{n})f_n(n-a)$. Next observe that if (**A.L**) holds then:

$$\lim_{n\to\infty}(1-\frac{a}{n})f_n(n-a) = 1 \tag{13}$$

So, as n becomes large, σ^* ensures the designer a return close to the maximum payoff attainable. The intuition is straightforward: by randomizing over all possible star networks, the designer conceals the identity of the central node from the adversary. For fixed attack budget, a, as n grows, the probability that the adversary successfully locates the central node declines and eventually goes to zero.

This derivation also illustrates the designer's incentives to conceal the network design from the adversary. Recall, in the sequential model, the equilibrium payoff to the designer is given by $d/(d+a)$ (see Theorem 2, and the discussion in section 3). By contrast, in the simultaneous move game the designer can hope to earn (almost) 1, the maximum attainable payoff, in the *absence* of any threat from the adversary!

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 and the equilibrium networks consist of equal size components. The number of components grows and size falls as the attack budget of the adversary increases.

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 an equilibrium. In equilibrium, the designer allocates all his budget to protecting the central node while the adversary allocates all resources to attacking the same node. On the other hand, if the budgets are large then denser networks with dispersed defence allocations arise in equilibrium.

6 Appendix A

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

$$F(x) = \begin{cases} \alpha & \text{if } 0 \leq x < n \\ 1 & \text{if } x = n \end{cases} \tag{14}$$

and

$$G(x) = \begin{cases} \alpha^2 & \text{if } 0 \leq x < n_1 \\ \alpha & \text{if } n_1 \leq x < n_2 \\ 1 - \delta^2 & \text{if } n_2 \leq x < n \\ 1 & \text{if } x = n \end{cases} \tag{15}$$

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

$$\alpha - \alpha^2 = 1 - \delta^2 - \alpha \tag{16}$$

or, substituting for δ

$$\alpha - \alpha^2 = 2\alpha - \alpha^2 - \alpha \tag{17}$$

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: Cases where $a = 0$ or $d = 0$ are trivial. If $a = 0$ then there is no attack and any connected network yields payoff 1. If $a \geq 1$ and $d = 0$ then by (**A.1**)-(**A.3**) any connected network yields zero payoff in equilibrium. 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 (a star with all defense resources allocated to the central node) there exists $c \in \mathcal{N}$ such that the optimal response of $\mathcal{A}$ consists in allocating c units of resource to the central node and exactly 1 unit of resource to $a - c$ periphery nodes. Two, we show that given a center-protected star the attacker allocates all resources to the center, resulting in designer payoff $\frac{d}{d+a}$. Last we show that, for any other network and allocation of defense resources the payoff of the designer is at most $\frac{d}{d+a}$.
Step 1: Consider a center-protected star 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, 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 hand 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. The payoff of $\mathcal{D}$ (conditional on M) can be written as

$$(1 - a_{i_1})(1 - a_{i_2}) f_n(m + 2) + [a_{i_1}(1 - a_{i_2}) + a_{i_2}(1 - a_{i_1})] f_n(m + 1) + a_{i_1} a_{i_2} f_n(m) \tag{18}$$

Assume up to relabeling $a_{i_1} \leq a_{i_2}$ and let $\overline{a} = a_{i_1} + a_{i_2}$. The former expression becomes

$$(1-\overline{a})f_n(m+2) + \overline{a}f_n(m+1) + a_{i_1}(\bar{a} - a_{i_1})[f_n(m+2) - 2f_n(m+1) + f_n(m)] \tag{19}$$

By convexity of $f_n(.)$, $f_n(m+2) - 2f_n(m+1) + f_n(m) > 0$. So the expression is increasing in a_{i_1}, 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. Observe that, by hypothesis, the designer allocates no resources to protect the peripheral node: so it is never optimal to raise attack resource allocation on such a node beyond 1 unit.

Putting together these observations, we conclude 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 for large enough n).

Step 2: In a center-protected star the attacker allocates all resources to the center, resulting in designer payoff $\frac{d}{d+a}$.

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

$$\frac{d}{c+d} f_n(n-a+c) \tag{20}$$

Differentiating with respect to c yields

$$-\frac{d}{(c+d)^2} f_n(n-a+c) + \frac{d}{(c+d)} f_n'(n-a+c) \tag{21}$$

Notice next that since $f_n(.)$ is increasing and convex, for all $c < a$:

$$f_n(n-a+c) \leq f_n(n-a+c) + (a-c)f_n'(n-a+c) \leq 1 \tag{22}$$

Hence, by (**A.L**):

$$\lim_{n\to\infty} f_n'(n-a+c) = 0, \ \forall\, c < a \tag{23}$$

The derivative in (21) thus tends to $-\frac{d}{(c+d)^2} f_n(n) < 0$ as n becomes large. So the best-response of $\mathcal{A}$ to a center-protected star consists in setting $c = 0$, i.e. allocating all resources

to the central node. The resulting payoff of $\mathcal{D}$ is then $d/(a+d)$.

Step 3: If the designer does not choose a center-protected star then he obtains at most payoff $\frac{d}{d+a}$.

Consider an arbitrary (connected) network g, and arbitrary allocation of defence resources across nodes in this network. We will find an attack strategy which leaves the designer with payoff at most $d/(a+d)$.

In what follows we let $X(\mathbf{d}) = \{i_1, ..., i_k\}$ denote the set of nodes for which $d_i \geq 1$, and $O = N - X$. We next consider separately the cases where $k = 0$ and $k \geq 1$.

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

$$\min\{a_j, \frac{a_j}{a_j + d_j}\} \geq \min\{a_j, \frac{a_j}{a_j + d}\} = \frac{a}{a+d} \tag{24}$$

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

Next suppose $k \geq 1$ and consider the attack strategy which sets $a_{i_t} = \frac{a}{d} d_{i_t}$, $t \in \{1, .., k\}$. Notice that, for all t in $\{1, .., k\}$:

$$\min\{a_{i_t}, \frac{a_{i_t}}{a_i + d_{i_t}}\} = \frac{a_{i_t}}{a_{i_t} + d_{i_t}} = \frac{a}{a+d} \tag{25}$$

Let $O_{i_t} \subset \big(O \cup \{i_t\}\big)$, $t = 1, ..., k$ denote the subset of nodes which can be reached from i_t through a path lying in $O \cup \{i_t\}$. Construct the sequence of sets $(N_{i_t}(g, \mathbf{d}))_{1 \leq t \leq k}$ recursively as follows:[21]

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

Let $n_{i_t} = |N_{i_t}|$, $t = 1, ..., k$. Note that by connectedness of g, $\bigcup_{t=1}^{k} N_{i_t} = N$. Moreover, by construction of the sets it follows from (**A.3**) that nodes in N_{i_t} survive only if i_t survives, for all $t \in \{1, .., k\}$. So the distribution of the total number of surviving nodes is first order stochastically dominated by that of $n_{i_1} I_1 + .. + n_{i_k} I_k$, where $\{I_1, .., I_k\}$ denote a set of independent Bernoulli random variables such that $P(I_t = 1) = \frac{d}{a+d}$, for all $t \in \{1, .., k\}$. Given increasing and convex $f(.)$, the expected payoff of $\mathcal{D}$ is thus at most $E[f_n(n_{i_1} I_1 + .. + n_{i_k} I_k)]$.

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

Convexity of $f(.)$ and Lemma 1 then show (see e.g., Rothschild and Stiglitz (1970))

$$E[f_n(n_{i_1}I_1 + .. + n_{i_k}I_k)] \leq E[f_n((n_{i_1} + .. + n_{i_k})I_1)] = \frac{d}{a+d} \tag{26}$$

Hence, by step 2, we have found an attack strategy such that the resulting payoff of $\mathcal{D}$ is less than he can guarantee himself with a center-protected star. This completes the proof of the Theorem. ■

Proposition 1 *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 is an equilibrium network.* (ii) *If $a = 1$ then, for $d < n$, the star is an equilibrium network. For $a = 1$, $d = n$ the complete and ring networks are equilibrium networks, and payoff dominate the star network.*

Sketch of Proof: We provide a sketch of the proof; the details are available from the authors upon request. 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

$$\frac{d_j}{d_j+1}f(n) + \frac{1}{d_j+1}f(n - \frac{d_j}{d}n) \tag{27}$$

Indeed, we can write the difference in payoff between this network and allocation of defence

and the center-protected star as:

$$\frac{d_j}{d_j+1}f(n)+\frac{1}{d_j+1}f(n-\frac{d_j}{d}n)-\frac{d}{d+1}f(n) \tag{28}$$

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

$$\frac{n-1}{2}f(1)+\frac{1}{2}f(n) \tag{29}$$

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

$$\frac{1}{2}f(n-1)+\frac{1}{2}f(n) \tag{30}$$

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. ■

7 Appendix B

This section extends the model to allow for a generalized contest function on individual nodes and then also considers richer models of spread of attack. We show that the principal insights contained in Theorem 1 and Theorem 2 extend to progressively general settings. 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:

$$\frac{a_i^\gamma}{a_i^\gamma+d_i^\gamma} \tag{31}$$

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 equilibrium 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′**), (**A.3**) and **A.L**) hold. Fix budgets $a, d \in \mathcal{N}$ and let $a \geq d$. For n sufficiently large, in equilibrium the network is a star, and the designer and adversary allocate all their resources to the central node.*

Sketch of proof: Steps 1 and 2 follow directly as before. 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 $d^\gamma/d^\gamma + a^\gamma$. 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 inequality allows the adversary to mimic the ratio and respect the integer allocation constraint). 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 (strictly) above by $d^\gamma/d^\gamma + a^\gamma$. Putting together steps 1-3 completes the proof. ■

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 (31).

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 2″: *Suppose assumptions (**A.0**), (**A.1**), (**A.2′**), (**A.3'**) and (**A.L**) hold. Fix*

$a, d \in \mathcal{N}$ and consider the set of connected networks. For n sufficiently large, the star network is 'almost' optimal for the designer.

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 $f_n(n-a)$. 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 $f_n(n-a)$ in a star network with defended central node.

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 $f_n(n-a)/2$. Observe that the payoff to $\mathcal{D}$ is at most $f_n(n-a)/2$ in any network with a single protected node. So the star is optimal. ■

8 Appendix C: Design followed by conflict

Following a suggestion by one of the referees, we present here an analysis of the game in which the designer first chooses a network and the designer and adversary then simultaneously choose the defence and allocation resources on the network. Let us refer to this as the $\mathbf{D}^*$ game.

Define a k-regular core periphery network as a core periphery network in which there are k core nodes and each core node is connected to $(n-k)/k$ peripheral nodes.

Proposition 2 *Consider the $\mathbf{D}^*$ game. Suppose assumptions (**A.1**)-(**A.3**) and (**A.L**) hold and restrict attention to regular core periphery networks. Given d and a, for large enough n, the star is an equilibrium network; in equilibrium the designer and adversary assign all their resources to the central node.*

Sketch of Proof: Fix some defense budget d and some attack budget a and let n be large.

First, in star network, the unique equilibrium is one in which designer allocates d to central node and adversary allocates a to central node.

Second suppose $k = 2$. Then there is an equilibrium (in the set of pure strategies) in which designer allocates $d/2$ to each core node and periphery allocates $a/2$ to each core node. Number the core nodes 1 and 2.

Suppose that adversary allocates $a/2$ to each core node. Consider an allocation $d_1 = d/2 + x$, $d_2 = d/2 - x$. We will show that it is optimal for the designer to set $x = 0$.

Observe that under (**A.L**) it is not optimal to allocate any resources to the periphery nodes. Next, consider allocations on the two core nodes. For simplicity assume that $d \geq 2$; the case where it is smaller than 2 is uninteresting under our threshold assumption (**A.3**). Observe that there are four states of the world: both core nodes are defended, both are attacked successfully, and two states corresponding to the case where only one of them is attacked successfully. The payoff to the designer from this strategy is given by:

$$f_n(\frac{n}{2}) \left[\frac{2da}{(d+2x+a)(d-2x+a)} \right] + f_n(n) \frac{d^2 - 4x^2}{(d+2x+a)(d-2x+a)} \tag{32}$$

We ask how a change in x affects the payoff. Differentiating this payoff with respect to x, we get:

$$f_n(\frac{n}{2}) \left[\frac{-8x}{(d+2x+a)^2(d-2x+a)^2} \right] + f_n(n) \left[\frac{(-8x)((d+2x+a)(d-2x+a) - (d^2 - 4x^2))}{(d+2x+a)^2(d-2x+a)^2} \right]. \tag{33}$$

Simplifying, we get

$$\left[\frac{-8x}{(d+2x+a)^2(d-2x+a)^2} \right] \left[-f_n(\frac{n}{2}) + f_n(n)(a^2 + 2ad) \right]. \tag{34}$$

This expression in negative if

$$\frac{f_n(n)}{f_n(n/2)}) > \frac{1}{a^2 + 2ad}. \tag{35}$$

So the designer allocates resources equally to the two core nodes if this inequality is satisfied. For $a, d \geq 1$, this inequality is satisfied for all functions $f_n(.)$ which satisfy (**A.1**).

Now consider optimality of the adversary's strategy in the face of an equal split of defence resources $d/2$ between the two core nodes. As before under (**A.L**), it follows that the adversary will not find it optimal to assign any resources to the peripheral node.

The payoff to an attack strategy $a/2 + x, a/2 - x$ is given by:

$$f_n(\frac{n}{2}) \left[\frac{2da}{(d + 2x + a)(d - 2x + a)} \right] + f_n(n) \frac{d^2}{(d + 2x + a)(d - 2x + a)} \tag{36}$$

It is easily checked that the denominator is falling in x. So it follows that the designer's payoff is increasing in x and is minimized at $x = 0$.

This proves that in a regular core periphery network with 2 core nodes, there exists an equilibrium in which both the adversary and the designer assign equal resources to the two core nodes.

This argument can be generalized to cover $k \geq 2$ nodes: fix equal allocations for the protected nodes for the designer and adversary some allocation of $k - 2$ nodes. Now use the $k = 2$ arguments to show show that the adversary gains by equalizing the allocations on the remaining two nodes. Then repeat the exercise for the designer. ■

9 References

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

2. Anderson, R. (2008), *Security Engineering.* Second Edition. Wiley.

3. Arquilla, J. and D. Ronfeldt (1996), *The Advent of Netwar* (RAND: Santa Monica, CA).

4. Arquilla, J. and D. Ronfeldt (2001), *Networks and Netwars: The Future of Terror, Crime, and Militancy* (RAND: Santa Monica, CA).

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

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

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

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

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

10. 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.

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

12. Clark, D. J. and K. A. Konrad (2007), Asymmetric conflict: weakest link against best shot. *Journal of Conflict Resolution*, 51, 457-469.

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

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

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

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

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

18. Goyal, S. and A. Vigier (2010), Robust networks, *mimeo*, University of Cambridge.

19. 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.

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

21. Hirshleifer, J (1983), From Weakest-Link to Best-Shot: The Voluntary Provision of Public Goods, *Public Choice*, 41, 3, 371-386.

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

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

24. Jackson, M. O. (2008), *Social and economic networks.* Princeton University Press. Princeton. New Jersey.

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

26. Kovenock, D. and B. Roberson (2009), The optimal defence of a networks of targets, *mimeo*, University of Iowa and University of Miami.

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

28. Moore, T., R. Clayton and R. Anderson (2009), The economics of online crime, *Journal of Economic Perspectives*, 23, 3, 3-20.

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

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

31. Powell, R. (2009), Sequential non-zero sum Blotto: allocating defence resources prior to attack, *Games and Economic Behavior*, 67 2, 611-615.

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

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

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

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

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

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

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

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

40. 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.

41. 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.

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

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