UNIVERSITY OF CAMBRIDGE
Faculty of Economics

Institute for New Economic Thinking

Cambridge-INET Institute

Cambridge-INET Working Paper Series No: 2014/02

Cambridge Working Paper in Economics: 1451

ATTACK, DEFENSE AND CONTAGION IN NETWORKS

Sanjeev Goyal
(University of Cambridge)

Adrien Vigier
(University of Oslo)

ABSTRACT

Connections between individuals facilitate the exchange of goods, resources and information and create benefits. These connections may be exploited by adversaries to spread their attacks as well. What is the optimal way to design and defend networks in the face of attacks?

We develop a model with a Designer and an Adversary. The Designer moves first and chooses a network and an allocation of defense resources across nodes. The Adversary then allocates attack resources on nodes; if an attack succeeds then the Adversary decides on how successful resources should navigate the network.

We obtain two principal results. One, we show that in a wide variety of circumstances a star network with all defence resources allocated to the central node is optimal for the Designer. Two, we identify conditions on the technology of conflict, network value function and the resource configuration for which networks with multiple hubs/components are optimal.

Attack, Defense and Contagion in Networks

Sanjeev Goyal [*]   Adrien Vigier[†]

January 18, 2014

Abstract

Connections between individuals facilitate the exchange of goods, resources and information and create benefits. These connections may be exploited by adversaries to spread their attacks as well. What is the optimal way to design and defend networks in the face of attacks?

We develop a model with a Designer and an Adversary. The Designer moves first and chooses a network and an allocation of defense resources across nodes. The Adversary then allocates attack resources on nodes; if an attack succeeds then the Adversary decides on how successful resources should navigate the network.

We obtain two principal results. One, we show that in a wide variety of circumstances a star network with all defence resources allocated to the central node is optimal for the Designer. Two, we identify conditions on the technology of conflict, network value function and the resource configuration for which networks with multiple hubs/components are optimal.

[*] Faculty of Economics and Christ's College, University of Cambridge. Email: sg472@cam.ac.uk

[†] Department of Economics, University of Oslo. Email: a.h.vigier@gmail.com

We are grateful to the editor, Philipp Kircher, and three anonymous referees for suggestions that have substantially improved the paper. We thank Marcin Dziubinksi for very helpful comments. We also thank Nizar Allouch, Murali Agastya, Mariagiovanna Baccara, Oliver Bätz, Parimal Bag, Francis Bloch, Heski Bar-Isaac, Yann Bramoulle, Antonio Cabrales, Diego Cerdiero, Indranil Chakravarty, Michiel de Jong, Matthew Elliot, Andrea Galeotti, Edoardo Gallo, Julien Gagnan, Aditya Goenka, Britta Hoyer, Matthew Jackson, Daniel Kovenock, Gilat Levy, Margaret Meyer, Francesco Nava, Volcker Nocke, Romans Pancs, Anja Prummer, Rony Razin, Bryony Reich, Pauline Rutsaert, Fernando Vega-Redondo, Rakesh Vohra, and participants at a number of seminars for comments. Sanjeev Goyal acknowledges support from a Keynes Fellowship, Cambridge-INET Institute and European Research Area Complexity-Net ('Resilience and interaction of networks in ecology and economics').

1 Introduction

Connections between individuals, cities, countries and computers facilitate the exchange of goods, resources and information and generate value. However, these connections may serve as a conduit for the spread of damaging attacks. The Internet reflects this tension clearly. Connectivity facilitates communication but is also used by hackers, hostile governments and firms, and 'botnet' herders to spread viruses and worms which compromise user privacy and jeopardize the functioning of the entire system.[1] As energy, communication, travel, consumer interaction increasingly adopt digital networks, cybersecurity has emerged as a major priority.[2] At the heart of these developments is the question of how to design and defend large scale networks.

In their influential paper on computer security, Staniford, Paxson and Weaver (2002) identify stealth worms and viruses as the main threats to security in computer networks. Using data from actual attacks, they argue that adversaries scan the network to explore its topology and the vulnerabilities of nodes, prior to attack. In the first instance, the objective is to deploy a worm on selected nodes in the network. Deployed worms then exploit communication between nodes to progressively take control of neighboring nodes in the network. The likelihood of capture of a node and the spread of the worm in a network depends on the strength of the worm, the topology of connections and on vulnerabilities of individual nodes. These considerations motivate the following theoretical model.

We consider a setting with two players: a Designer and an Adversary. The Designer moves first and chooses a network and an allocation of defense resources. The Adversary then allocates attack resources on nodes; if an attack succeeds then the Adversary decides on how successful resources should navigate the network. The model has three important ingredients: the value of the network, the technology of conflict between between defense and attack resources, and the spread of successful attack resources through the network.

We assume that the value of a network is increasing and convex in the number of intercon-

[1] In 2009, roughly 10 million computers were infected with malware designed to steal online credentials. The annual damages caused by malware is of the order of 9.3 billion Euros in Europe, while in the US the annual costs of identity theft are estimated at 2.8 billion USD (Moore, Clayton and Anderson (2009)). One indicator of the economic magnitude of the problem is the valuation of security firms: Intel bought McAfee in 2010, for 7.68 billion USD (bbc.co.uk; 19 August 2010).

[2] In the United States, the Department of Homeland Security (DHS) is responsible for cybersecurity. Its mission statement reads,"Our daily life, economic vitality, and national security depend on a stable, safe, and resilient cyberspace. We rely on this vast array of networks to communicate and travel, power our homes, run our economy, and provide government services."

nected nodes.[3] We model the conflict between defense and attack resources on a network node as a *Tullock contest*.[4] The contest defines the probability of a win for Designer and Adversary, as a function of their respective resources. The resources of the loser of the contest are eliminated, the winner retains his resources. In case the Adversary wins a contest on a node, the winning attack resources can move and attack neighboring nodes. The dynamics of conflict continue as long as both defense and attack resources co-exist. The initial network design and the conflict dynamics yield a probability distribution on surviving nodes, i.e., nodes that have not been captured by the Adversary. The Designer and Adversary are engaged in a zero sum game; so, given a defended network, we consider the minimum payoff of the Designer given all possible attacks. An *optimal defended network* maximizes this (minimum) payoff.

We obtain two principal results. One, we show that in a wide variety of circumstances the optimal defended network is a star network with all defence resources allocated to the central node (a CP-star). Two, we identify conditions on the technology of conflict, network value function and the resource configuration for which networks with multiple hubs/components are optimal.

The argument is developed in two steps. In the first step, we consider the class of connected networks.[5] The dynamics of conflict and contagion on the CP-star yield extremal outcomes: either (almost) all nodes survive and remain connected or all nodes are captured. Consider next a network with two equally defended hub nodes and an equal number of periphery nodes linked to either hub. Faced with this defended network, the Adversary can allocate resources to peripheral nodes in line with the defense resources allocated to the corresponding hub node. The dynamics of conflict and contagion can generate extremal outcomes (as in the CP-star) but they also generate intermediate outcomes in which one hub is captured but the other hub (and its peripheral nodes) survives. The expected number of surviving nodes is equal in the two scenarios, but the CP-star yields a mean-preserving spread distribution on surviving nodes. Since the network value function is convex in number of interconnected nodes, the CP-star thus generates greater expected payoffs for the Designer. Theorems 1 and 2 generalize these ideas to cover all connected networks.

[3] This is consistent with Metcalfe's Law (network value is proportional to the square of nodes) and Reed's Law (network value is exponentially increasing in nodes). Our assumption is also in line with the large theoretical literature on network externalities (Katz and Shapiro, 1985; Farrel and Saloner, 1986) and network economics (Bala and Goyal 2000a).

[4] Here we build on the rich literature on rent seeking and conflict, see Tullock (1980) and Hirshleifer (1995).

[5] Two nodes are connected if there is a path between them. A component is a maximal set of nodes that are connected. A network is connected if it contains only one component. Formal definitions are provided in Section 2.

In the second step, we allow for networks with multiple components. This leads us to study a situation where defense allocation, number of components and the architecture of individual components are all decision variables for the Designer. Theorem 4 characterizes network value functions for which optimal defended networks are connected and identifies circumstances when multiple defended components may be optimal. It says, roughly speaking, that if network value grows exponentially in the number of nodes then the CP-star is optimal, but if value grows at a slower rate (as in a polynomial function) then networks with multiple components may be optimal. Finally, Proposition 1 and section 4.2 explore the interaction between network value function, the technology of conflict and the resource configuration to identify circumstances when multiple hubs may be optimal.

The optimality of the CP-star is consistent with the practice of traffic monitoring at key nodes by security personnel (Anderson, 2010). In the context of Peer-to-Peer file sharing our results suggest that large networks, such as BitTorrent (where the same content can be obtained from many providers and hence the value function is no longer strongly convex), ought to contain multiple hubs.

Our framework of network design, defense and attack provides a useful way to think about a number of questions relating to networks that face threats. Section 4 shows that by varying our assumptions on network value functions, number of players, and the timing of moves we trace out an ensemble of models that can help us address questions in economic epidemiology, terrorist networks, modern warfare, finance and criminal activity.

Our paper contributes to two rich strands of theoretical research in economics: the theory of networks and the theory of conflict. The research on networks is concerned with the formation, structure and functioning of social and economic networks; for surveys of this work, see Goyal (2007), Jackson (2008), and Vega-Redondo (2007). To the best of our knowledge, the present paper is the first to study design and defense of networks that face an intelligent Adversary.[6]

In Baccara and Bar-Isaac (2008) information links between criminals facilitate cooperative play, but the detection of one criminal leads to the detection and punishment of connected others. This creates a trade-off between connections and vulnerability and suggests a similarity with the present paper. However, the models differ along a number of dimensions as they are motivated by very different applications. We highlight four differences. One, in our model the gains from large scale connectivity are key; by contrast, in their model the size of the network

[6] In an early paper, Bala and Goyal (2000b) study network formation with exogenous probability of link failure. Hong (2008) studies network formation model with strategic complementarities between linking and protection. These models do not have an active Adversary.

plays no essential role in defining network value. Two, we study conflict between defense and attack; by contrast, there are no defense resources in their paper. Three, the Designer moves first in our model; the Adversary moves first in their model. Four, we assume that links are undirected, they use a formulation with directed links. These differences are substantive and taken together lead to very different insights. Section 4.3 develops this point in greater detail.

The theory of contests studies allocation of resources in situations of conflict; see e.g., Baye (1998), Bier, Oliveros and Samuelson (2006), Garfinkel and Skaperdas (2012), Hart (2008), Konrad (2009), Kovenock and Roberson (2012) and Roberson (2006). Our paper extends this line of work along two dimensions: one, we locate individual contests 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 'battlefields'.

The problem of network design and defence has been extensively studied in electrical engineering and computer science; for an overview of this work see Alpcan and Basar (2011), Anderson (2011) and Roy et al. (2010). In an early paper, Cunningham (1985) looks at the problem of network design with conflict on links. A link is eliminated if the Adversary assigns more resources than the Designer (thus conflict is modeled as an all-pay auction). Network security is a very active field of research currently in these disciplines. Gueye, Walrand and Anantharam (2011) and Laszka, Szeszler, and Buttyan (2012) study a model in which the network operator chooses a spanning tree of a given network to route messages, and simultaneously, the attacker chooses an edge to be removed. Aspnes, Chang and Yampolskiy (2006) (and the literature that follows them) study protection choices by nodes faced with a viral infection; upon infecting a node, the virus travels through the network. Blume et al. (2011) study optimal and stable networks in a context where direct links bring benefits but infections travel through links with an exogenous probability. Our paper contributes to this literature by developing a general framework for the study of optimal network design and defence in a setting with strategic conflict and contagion dynamics. The analysis of this framework yields results on optimality of CP-star and multi-hub networks that are new.

The rest of the paper is organized as follows. Section 2 presents our model. Section 3 studies optimal defended networks. Section 4 discusses extensions and open research questions while Section 5 concludes. All proofs are presented in an appendix.

2 Model and application

We study a zero-sum game between a Designer and an Adversary. The Designer has a collection of nodes and defense resources, while the Adversary has attack resources. The Designer moves first and chooses links between the nodes and allocates resources across the nodes. The network and defense choices of the Designer are observed by the Adversary, who then chooses an attack strategy. The network design, the allocation of resources and the dynamics of conflict define a probability distribution on surviving nodes which in turn determines players' payoffs. We first set out the notation and concepts to formally describe this game and then relate our modeling assumptions to problems in computer security.

2.1 The Designer-Adversary game

The Designer has a collection of nodes $N = \{1, ..., n\}$, $n \geq 2$. He chooses links between the nodes and allocates $d \in \mathbb{N}$ resource units across the nodes. Let $\underline{d} = (d_1, d_2, ..., d_n)$ denote the vector of allocated resources, where $d_i \in \mathbb{N}$ 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 nodes and the links together define a network g. The network-defense pair $(g, \underline{d})$ defines a strategy for the Designer. The strategy such that g is a star network and all defense resources are allocated to the central node (a center-protected star) plays a prominent role in the paper. We will refer to this particular strategy as a *CP-star*, and denote it $(g^s, \underline{d}^s)$.

A path between two nodes i and j in network g is a sequence of nodes $i_{1,..}$, i_k such that $g_{ii_1} = g_{i_1 i_2} = ... = g_{i_{k-1} i_k} = g_{i_k j} = 1$. Two nodes are said to be connected if 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 $|C|$ indicate the cardinality (or size) of the component C. A maximum component of g is a component with maximum cardinality in $\mathcal{C}(g)$. A network with a single component is said to be connected.[7] A network g' on N' is a sub-network of g if and only if $N' \subset N$, and $g'_{ij} = 1$ implies $g_{ij} = 1$. We let $\mathcal{G}(g)$ denote the set of sub-networks of g.

[7] The complete network, g^c, has $g_{ij} = 1$, for all pairs (i, j). The empty network, g^e, has $g_{ij} = 0$ for all pairs (i, 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 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 general introduction to networks concepts and terminology, see Goyal (2007).

Following Myerson (1977), we assume that the value of a network is the sum of the value of the different components and that the value of any component is a function of its size only. Let the function $f : \mathbb{N} \to \mathbb{R}_+$ specify a value to component size. If f is decreasing then the value of a component is falling in its size: splitting the network enhances value. Similarly, if f is increasing and concave, then value from a group of nodes is maximized when they are in a collection of singleton components. Our interest is in the tension between the pressure to connect nodes to create value and the threat of contagion of attack via connections: so, in the benchmark model, we assume increasing and convex returns to size of component.

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

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

where f is (strictly) increasing and (strictly) convex.

Increasing and convex network value functions arise naturally in the large literature on network externalities (see e.g. Guye and Marbukh (2012); Katz and Shapiro (1985); Farrell and Saloner (1986)). In that literature, the value to a consumer from buying a product is related to the number of other consumers who buy the same product, i.e., belong to the same network. In its simplest form this gives rise to the quadratic form $f(n) = n^2$. Such a function also arises in the well known communications model in the literature on network economics (see e.g. Goyal (1993); Bala and Goyal (2000a)). The appendix presents the details of this derivation.

Given a defended network $(g, \underline{d})$, let K denote the subset of protected nodes and O the subset of unprotected nodes. Further, for $i \in N$ let $O_i \subset O$ denote the subset of unprotected nodes which can be reached from i through some path such that each node on that path lies in O. Similarly, let $K_i \subset K$ denote the subset of protected nodes which can be reached from i through some path such that each node on that path lies in O.

The Designer moves first and chooses a strategy $(g, \underline{d})$; this is observed by the Adversary, who then chooses a strategy $(\underline{a}, \Delta)$. The Adversary first allocates $a \in \mathbb{N}$ units across the nodes, $\underline{a} = (a_1, a_2, .., a_n)$, where $a_i \in \mathbb{N}$ and $\sum_{i \in N} a_i \leq a$.[8] The matrix $\Delta = \left(\delta_{ij}\right)_{i,j \in N}$ describes subsequent movements of (successful) attack resources. Row i in the matrix Δ

[8] The assumption of integer defence and attack allocations is made for expositional reasons. Our main result holds with continuous defense and attack resources under stronger assumptions on the contest function and on the dynamics of contagion; for completeness, this variant of our result with continuous actions and its

specifies a 'pecking order' on K_i: resources on node i relocate to node $j_1 \in K_i$ with $\delta_{ij_1} = 1$. If j_1 has already been captured, resources are relocated to node j_2 with $\delta_{ij_2} = 2$, and so forth.[9] The details of the dynamics of attack contagion are described later in this section.

Attack resources a_i and defense resources d_i located on a node i engage in a *contest* for control of the node. If $a_i + d_i > 0$ then, following Tullock (1980), we set

$$\text{probability of successful attack} = \frac{a_i^\gamma}{a_i^\gamma + d_i^\gamma} \tag{2}$$

where $\gamma > 0$. If a_i is 0 then the probability of successful attack is 0, irrespective of the value of d_i: a node is safe if it is not under attack. Skaperdas (1996) provides axiomatic foundations for the Tullock contest function. The parameter γ is referred to as the technology of conflict in the literature on conflict (Hirshleifer (1995)). Raising γ favors the side with more resources. In particular, an all-pay auction – where the side with more resources wins the contest for sure – is a special case of our model. An important feature of the contest function is that it is homogenous of degree 0 in resources. We further assume that all contests are statistically independent.

The discrete-time dynamics of attack then proceed as follows:

At time **t=0**: The attack begins with unprotected nodes. For all $i \in O$ such that $a_i > 0$ the Adversary (i) captures i, (ii) captures O_i and, (iii) relocates a_i attack resources to node $j = \arg\min_{k \in K_i}\{\delta_{ik}\}$.

At time **t=1**: Let N^1 denote the set of un-captured nodes at the beginning of period $t = 1$ and $\underline{a}^1$ the allocation of attack resources at that point in time (all attack resources now target protected nodes). A contest takes place at all i such that $a_i^1 > 0$, following the rules defined in (2).

1. If *attack succeeds at* i then the Adversary (i) eliminates all d_i defense resources located there, (ii) captures node i, (iii) captures any remaining node in O_i and, (iv) relocates

proof are presented in the online appendix.

[9] Our assumptions concerning the relocation of attack resources are intended to reflect scarcity of operational resources. All our paper's results carry through under the alternative assumption that attack resources replicate and spread to all neighbors simultaneously, following a successful attack on a node. This follows from the observation that (3) is unchanged under the alternative assumption, while for any defended network and allocation of attack resources the resulting payoff of the Designer is weakly less under the alternative assumption than in our model.

the a_i^1 attack resources to node $j = \arg\min_{k \in K_i \bigcap N^2}\{\delta_{ik}\}$. If $K_i \bigcap N^2 = \emptyset$ then the a_i^1 attack resources are eliminated.

2. If *defense succeeds at i* then the Designer eliminates all a_i^1 attack resources located there.

At time **t=2**: Let $\underline{a}^2$ denote the allocation of attack resources at the beginning of period $t = 2$, and N^2 the set of un-captured nodes. If $\underline{a}^2 = \underline{0}$ then the process terminates. Otherwise, it proceeds following the rules laid out as in period $t = 1$.

Given a defended network $(g, \underline{d})$ and attack strategy $(\underline{a}, \Delta)$, the dynamics of conflict described above yield a probability distribution on $\mathcal{G}(g)$. Let $\mathbb{P}(g'|g, \underline{d}, \underline{a}, \Delta)$ denote the probability that the sub-network g' is the residual network of surviving nodes after all conflicts have ended. Observe that, given the rules of the dynamics, all conflict must cease within a maximum of $a + d$ period. Define $\Pi^e(g, \underline{d}, \underline{a}, \Delta)$ to be the expected payoff of the Designer given defended network $(g, \underline{d})$ and attack strategy $(\underline{a}, \Delta)$. Then

$$\Pi^e(g, \underline{d}, \underline{a}, \Delta) = \sum_{g' \in \mathcal{G}(g)} \mathbb{P}(g'|g, \underline{d}, \underline{a}, \Delta)\Pi(g').$$

Figures 1 and 2 illustrate the nature of the dynamics of attack. In Figure 1, $n = 12$, $a = d = 4$. The Designer allocates all 4 units to the central node, while the Adversary allocates 1 unit each to four unprotected peripheral nodes. These attack units capture the 4 peripheral nodes and then simultaneously attack the central node. Given contest function, (2), the Designer and Adversary face an equal probability of winning. If the Designer wins the contest, the attack resources are eliminated. There are 8 surviving connected nodes. In case the Adversary wins, the central node is captured and the defense resources are eliminated. The attack resources then capture the remaining 7 undefended peripheral nodes. The expected payoff of the Designer is $f(8)/2$.

Figure 2 illustrates the dynamics on the complete network, with $n = 4$ and $a = d = 1$. The Designer allocates his resource to node 1, while the Adversary allocates his to node 2. Since node 2 is undefended, it is captured at time $t = 0$, followed by undefended nodes 3 and 4 which are linked to it. At time $t = 1$, the attacking unit then spreads to node 1. Given (2), Designer and Adversary win with equal probability. The expected payoff of the Designer is $f(1)/2$.

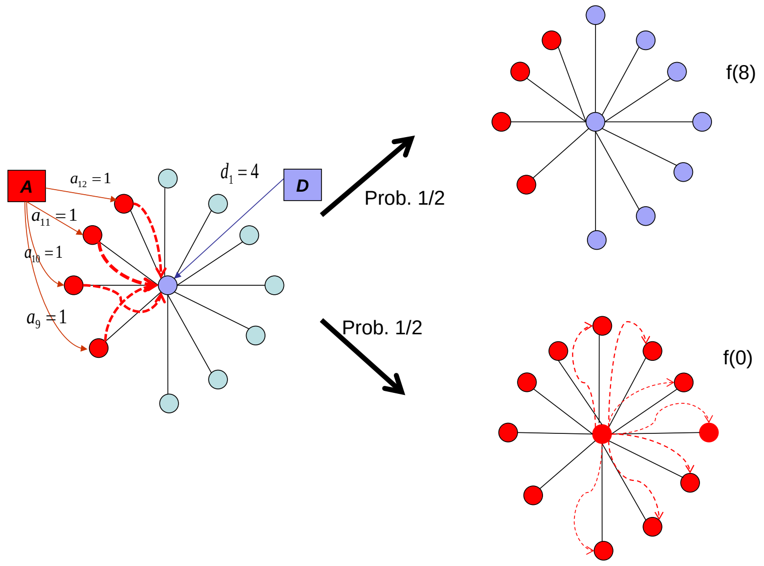


Figure 1: Dynamics of attack in a CP-star: $n = 12$, $a = d = 4$.

Let $\overline{\Pi}^e(g, \underline{d})$ denote the minimum expected payoff of the Designer playing strategy $(g, \underline{d})$:

$$\overline{\Pi}^e(g, \underline{d}) = \min_{\underline{a}, \Delta} \Pi^e(g, \underline{d}, \underline{a}, \Delta).$$

Since we are considering a zero-sum game, the minimax theorem applies and motivates the following definition of optimal defended networks:

Definition 1 *A defended network $(g, \underline{d})$ is optimal if $\overline{\Pi}^e(g, \underline{d}) \geq \overline{\Pi}^e(g', \underline{d}')$ for all defended networks $(g', \underline{d}')$.*
Given $\epsilon > 0$, a defended network $(g, \underline{d})$ is ϵ-optimal if $\overline{\Pi}^e(g, \underline{d}) \geq (1 - \epsilon)\overline{\Pi}^e(g', \underline{d}')$ for all defended networks $(g', \underline{d}')$.

2.2 Application: computer network security

We discuss the problem of first best design and defence in Peer-to-Peer (P2P) networks. A P2P network is an overlay computer network built on top of the physical computer network topology. The most popular use of such networks is file sharing; examples include BitTorrent, Gnutella, G2, and eMule. The returns from joining the network are increasing in the number of users. The appendix presents a simple model of communication networks to illustrate how this increasing value in number of users generates a network value function that is increasing and convex in component size.

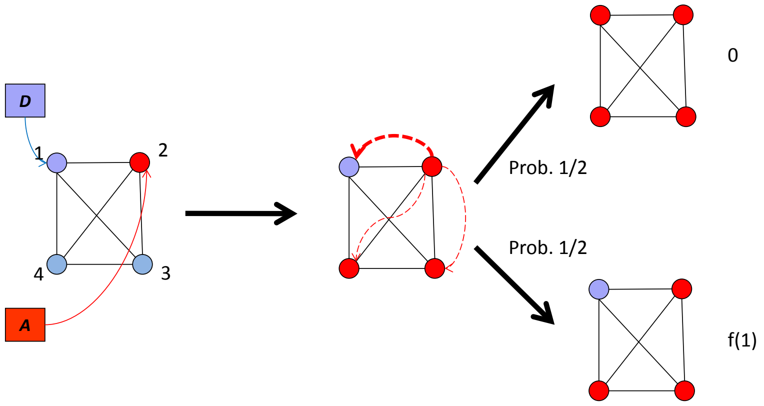


Figure 2: Dynamics of attack in a complete network: $n = 4$, $a = d = 1$.

Online criminals, such as hackers and 'botnet' herders take the topology and security of a P2P network as given when they attack hosts taking part in the network. These adversaries generally prepare their attack, after scanning the network to assess its topology and security; in their well known paper, Staniford, Paxson and Weaver (2002) elaborate on the different mechanisms available for such scanning and highlight the growing efficacy of scanning. This is in line with our assumption that the Adversary is aware of the topology and the defense allocations, prior to choosing the attack. Adversary knowledge of the network and vulnerabilities of nodes is assumed in the computer science and electrical engineering literature; see e.g., Saia et al (2002) and Suto et al. (2012). A theoretical reason for this assumption is that sometimes the interest is in understanding the behavior of the system in the worst possible case; assuming complete knowledge of the network enables the most effective attack. So a network that survives such an attack is especially attractive.

One of the main threats to P2P networks are self-propagating malicious software called stealth worms. Worms are typically deployed through viruses or other forms of malware. The quality of the malware and the number of deployments depend on the resources – programming skills, capital, and number of programmers – available to the hacker; in our model, this is captured by a. Hosts on the P2P network install security software and may employ security personnel to monitor traffic; the quality of this software and the degree of monitoring depends on the resources available for defence; in our model this is reflected in d.

The likelihood of successful infection of a host is higher the more sophisticated is the malware and the greater attention is devoted by the Adversary to a node. Similarly, the likelihood falls in the security resources – quality of security software and attention of specialized personnel – assigned to it. These features of the conflict between security and attacks is reflected in our contest function formulation (2).

Deployed worms propagate through the network by progressively taking control of neighboring hosts. The worm replicates and then attaches itself to packages of data sent between connected hosts; see Staniford, Paxson and Weaver (2002) (as noted in the previous section, our analysis and main results extend to a model where successful attack resources replicate themselves). The probability that the worm succeeds in infecting neighboring hosts varies with the level of security installations on them and the quality of malware being used. This transmission of a worm via communication links, the (relative) immobility of security installations, and the subsequent conflict between the virus and the security installed on neighboring hosts is consistent with our formulation of contagion dynamics.

3 Optimal defended networks

The Designer has two instruments at his disposal to sustain network value: strategic deployment of defense resources and creation of links. In particular, the Designer chooses the number and architecture of the components and the allocation of defense resources across nodes. This optimization problem is complicated and for expositional clarity it is convenient to proceed in steps. We start by solving the problem of optimal architecture and defense at the level of a single component.[10] We then consider the pure problem of number of components, in the absence of any defense resources. Finally, we combine the insights and present a result on optimal defended network where defense allocation, architecture of individual components and the number of components are all decision variables of the Designer.

Discrete optimization problems are marked by divisibility issues. We circumvent these difficulties here by assuming that $a/d \in \mathbb{N}$. The case $a < d$ is discussed separately at the end of Sections 3.1 and 3.3. A variant of our main result with continuous defence and attack strategies is stated and proved in the accompanying on-line appendix.

[10] This problem is of independent interest in situations where attacks are rare. The convexity of the network value function implies that the network is connected: defense resources are then primarily used to maximize operational capability in the rare event of attack. We thank the editor for this remark.

3.1 Connected networks

Suppose that the Designer chooses a CP-star. In this case the Adversary's best response is to allocate one resource unit to exactly a periphery nodes. The a periphery nodes are captured and the attack resources then mount a concerted attack on the central node. If attack succeeds on the central node, all remaining periphery nodes are subsequently captured. If attack fails, the Designer is left with $n - a$ connected nodes. The expected payoff of the Designer in a CP-star is

$$\overline{\Pi}^e(g^s, \underline{d}^s) = \frac{d^\gamma}{d^\gamma + a^\gamma} f(n-a). \tag{3}$$

The idea of a *mimic* attack strategy plays an important role in our analysis. Let $a = xd$, for some $x \in \mathbb{N}$, and consider a defended network $(g, \underline{d})$. For each node i such that $d_i > 0$ the Adversary allocates one resource unit to exactly x times d_i nodes in O_i – the unprotected neighbourhood of i – thereafter relocating each of these resource units to node i. Formally, given defended network $(g, \underline{d})$, say that $(\underline{a}, \Delta)$ *mimics* defence if and only if there exists a set of a distinct nodes, $\{j_1, ..., j_a\}$, such that:

1. $\{j_1, ..., j_{\frac{a}{d}d_{i_1}}\} \in O_{i_1}$;

   $\{j_{\frac{a}{d}d_{i_1}+1}, ..., j_{\frac{a}{d}d_{i_1}+\frac{a}{d}d_{i_2}}\} \in O_{i_2}$; ...;

   $\{j_{\frac{a}{d}d_{i_{k-1}}+1}, ..., j_{\frac{a}{d}d_{i_{k-1}}+\frac{a}{d}d_{i_k}}\} \in O_{i_k}$.

2. $\delta_{j_s i_1} = 1, \forall\ s$ s.t. $s \leq \frac{a}{d}d_1$;

   $\delta_{j_s i_2} = 1, \forall\ s$ s.t. $\frac{a}{d}d_1 + 1 \leq s \leq \frac{a}{d}d_1 + \frac{a}{d}d_2$; ...;

   $\delta_{j_s i_k} = 1, \forall\ s$ s.t. $\frac{a}{d}d_{k-1} + 1 \leq s \leq \frac{a}{d}d_{k-1} + \frac{a}{d}d_k$.

Figure 3, with $a = d = 4$, illustrates a mimic attack strategy. The network is a core-periphery network with two hubs such that the Designer allocates 2 units of defense to each hub. In the mimic strategy, the Adversary allocates 2 resource units to peripheral nodes connected to one hub and 2 resource units to peripheral nodes connected to the other hub. In the first instance, the Adversary captures these 4 peripheral nodes. The resources then target their respective hub nodes.

Mimic strategies do not always exist. The following remark, which follows immediately from Hall's theorem (see e.g. Bollobas (1998)), is a building block for the main results of our paper.

Remark 1 *Given defended network $(g, \underline{d})$, a mimic attack strategy exists if and only if the following condition holds:*

$$\left| \bigcup_{s=1}^{k'} O_{i_s} \right| \geq \frac{a}{d} \sum_{s=1}^{k'} d_{i_s} \quad \forall \{i_1, i_2, ..., i_{k'}\} \subset K \tag{4}$$

By way of illustration, suppose that $n = 12$, $a = d = 4$: the network has two hubs, with the first hub being linked to 9 peripheral nodes and the second hub being linked to one peripheral node. If the Designer allocates 2 units to each hub, then no attack strategy can mimic defense in this defended network.

We now state our first main result.

Theorem 1 *Assume that (A.1) holds, $a/d \in \mathbb{N}$, $n > a+1$ and consider the class of connected networks. Then an optimal network is either the CP-star or a defended network violating (4).*

The first observation is that if a defended network $(g, \underline{d}) \neq (g^s, \underline{d}^s)$ permits a mimic strategy $(\underline{a}^m, \Delta^m)$ then there is an upper bound on the maximum expected payoff of the Designer:

$$\overline{\Pi}^e(g, \underline{d}) \leq \Pi^e(g, \underline{d}, \underline{a}^m, \Delta^m).$$

The second – and key – observation is that the CP-star induces a distribution on the number of surviving nodes that is a mean-preserving spread (and with all surviving nodes connected) of the distribution yielded by any defended network $(g, \underline{d})$ that satisfies (4) and hence permits a mimic attack $(\underline{a}^m, \Delta^m)$. By convexity of f, this implies that

$$\overline{\Pi}^e(g^s, \underline{d}^s) > \Pi^e(g, \underline{d}, \underline{a}^m, \Delta^m).$$

Combining these inequalities gives us the desired result, for any defended network satisfying (4).

We illustrate this point with the help of Figure 3 discussed above. Recall that the network has two hubs and the Designer allocates 2 resource units to each hub, while the Adversary mimics defence. There are four possible outcomes of the two contests on the hubs: either both hubs survive, both hubs are captured or one hub survives and the other is captured. Given the equal resources engaged in contests, it follows that the first two outcomes each arise with probability 1/4. The two outcomes define terminal states of the dynamics, represented at the top and the bottom end of Figure 3. There is a probability 1/2 that one of the hubs survives and the other is captured. This is represented in the middle of the Figure 3. Capture of a hub triggers the capture of its respective peripheral nodes. All attack resources then target the surviving hub, inducing a second round of contests. With probability 1/2 the hub survives the attack, and with probability 1/2 it is captured. If the hub is captured then this triggers the capture of the remaining peripheral nodes. This brings to an end the dynamics of conflict.

The probability density P on surviving nodes is: with probability 1/2 all nodes are captured, with probability 1/4 four nodes survive and with probability 1/4 eight nodes survive. Observe that this distribution is first order stochastically dominated by the distribution P' such that with probability 1/4 all nodes are captured, with probability 1/2 four nodes survive and with probability 1/4 eight nodes survive. But P' is in turn second order stochastically dominated by the distribution P'' in which all nodes are captured with probability 1/2, and eight nodes survive with probability 1/2. Finally, notice that P'' is the distribution facing the Designer if he chooses a CP-star.

Theorem 1 suggests that defended networks violating (4) may be attractive for the Designer, since they preclude the use of mimic strategies by the Adversary. Observe, for instance, that in a setting where $n = 3$, $f(n) = n^2$ and $a = d = 2$, a CP-star yields expected payoff 1/2 for the Designer. It is easy to see on the other hand that the complete network with two protected nodes (which violates (4)) yields at least 1. This shows that, at least in some circumstances, defended networks violating (4) may dominate a CP-star. We next explore the generality of this observation.

The key to the problem is the number of nodes, n. As n grows, structures violating (4) start losing their attractiveness. To make our argument precise, we proceed by assuming that $\frac{f(n-1)}{f(n)}$ converges, as $n \to \infty$. Define

$$\ell = \lim_{n\to\infty} \frac{f(n-1)}{f(n)}.$$

Given that f is an increasing function, ℓ is either equal to 1 or it is less than 1. To get a sense of what these limit values imply, note that if $f(n) = n^2$, then the limit $\ell = 1$; if, on the other hand, $f(n) = 2^n$, then the limit $\ell = 1/2$. Thus, roughly speaking, the limit $\ell = 1$ corresponds to polynomial functions, while $\ell < 1$ corresponds to exponential functions.

We are now ready to present our main result concerning optimal connected networks.

Theorem 2 *Assume that (A.1) holds, $a/d \in \mathbb{N}$ and $n > a + 1$. Let $\epsilon > 0$ and consider the class of connected networks. There exists n_0 such that, for all $n > n_0$:*

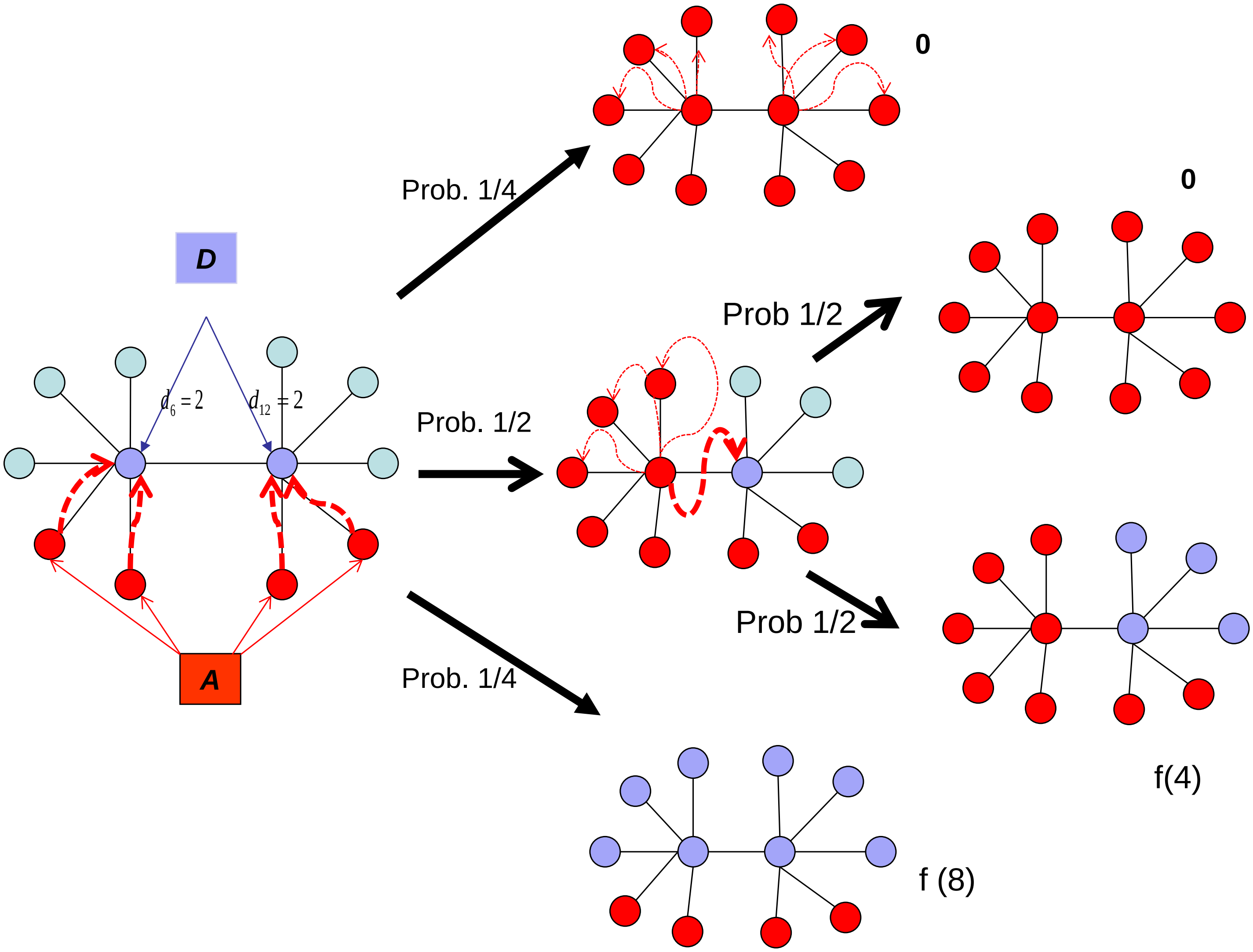


Figure 3: Mimic attack on two-hub network: $n = 12$, $a = d = 4$.

1. *If $\ell < 1$ the CP-star is uniquely optimal.*
2. *If $\ell = 1$ the CP-star is ϵ-optimal.*

If $\ell < 1$, the marginal contribution of a single node to network value is bounded away from zero. In this case, as n grows, spreading resources becomes increasingly risky for the Designer. All protected nodes must then have $d_i < d$, and the Adversary can focus his attack on a single node. If the unprotected neighborhood of that node contains a large enough fraction of all nodes, the Adversary can very effectively disrupt the network by targeting this node. This yields us part one of the result. If $\ell = 1$, the marginal contribution of a single node to network value vanishes, and so the Adversary can always approximate a mimic strategy by allocating all or part of his attack resources to the protected nodes themselves. The payoff from networks that violate (4) will therefore approximate the payoffs from corresponding networks that do respect that property. Combining this observation with Theorem 1 then yields us part two of the result.

Theorem 2 is a powerful result. It holds for all payoff functions which satisfy (A.1): so the result does not depend on the curvature of f. The result holds for all γ in the Tullock contest function: so the conclusion is robust with respect to the technology of conflict. The result holds for all resource configurations between the Designer and the Adversary such that $a/d \in \mathbb{N}$.

We now take up resource configurations a and d that violate this restriction. Define a defended core network as one in which (i) $x \in \{1, ...d\}$ nodes are protected, (ii) these x nodes constitute a connected sub-graph, and (iii) the $n - x$ unprotected nodes (if any exist) each have a unique link and this link is to a protected core node. It is possible to show that *a defended core network is optimal in the class of connected networks.* The proof is provided in the appendix.

This raises the question of how many core nodes are optimal. We do not have a complete answer to this question, but optimal networks now depend on resources and the technology of conflict. To make this point in the simplest way, we focus on a special class of symmetric core-periphery networks. In these networks, the protected core constitutes a clique (i.e., a completed subgraph) and every node in the core has an equal number of peripheral nodes. Denote the set of such networks by Λ. Recall that for $a \geq d$, and under the assumptions of Theorem 2, the CP-star is always optimal within Λ. By contrast, for $a < d$ the Designer may be tempted to exploit his resource advantage by spreading the defence and adapting the network. Whether he actually wishes to do so crucially hinges on γ. Large values of γ

favor concentration of defence and star like architectures, while low values of γ favor dispersed defence and more sprawled out architectures.[11]

Proposition 1 *Suppose that (A.1) holds, $a < d$, and $n > d + a$. Then, within Λ:*

1. *If γ is large, a CP-star is optimal.*
2. *If γ is small, the optimal defended network is either a CP-star or has d nodes in the core. In particular, if $a = 1$ then a core with $d > 1$ nodes strictly dominates the CP-star.*

These observations conclude our analysis of optimal connected networks. We now turn to the study of networks with multiple components.

3.2 Number of components

When the Designer has no defense resources, attack on a node induces the capture of the component to which it belongs. So, the only way to sustain network value is to separate the nodes into distinct components. This allows us to focus on the pure problem of number of components in optimal networks. The following result provides a characterization of the optimal number of components.

Theorem 3 *Assume that (A.1) holds and $d = 0$. (i) If $a < n/2$ then the optimal network contains at least $a+1$ maximal components and at most one component which is smaller. (ii) If $n/2 \leq a < n - 1$, then the empty network is the unique optimal network. If $a \geq n$ then every network yields payoff 0 to the Designer, and is optimal.*

If $a \geq n$ then the Adversary can always capture all nodes, so the Designer earns zero payoff irrespective of the network. Similarly, if $a \geq n/2$ then the Adversary can always capture any component with two nodes or more. So the interesting case is $a < n/2$. Observe now that there must be at least $a + 1$ components, else the payoff of the Designer is 0. A network with $a + 1$ components on the other hand guarantees the Designer strictly positive payoff. Finally, the Adversary will always prioritize the largest components. As a consequence, making some components larger than others is self-defeating for the Designer.

Theorem 3 sets lower bounds on the number of components; the precise number of components depends on the convexity of the payoff function. To gain further insights, we work

[11] We thank Michiel de Jong for drawing our attention to the optimality of multiple hub nodes in this case.

with a class of network value functions $f(n) = n^{\beta}$, where $\beta > 1$. We interpret β as a measure of the convexity of the network value function. Define $C(a, \beta) = \frac{\beta a}{\beta - 1}$. Observe that $C(a, \beta)$ is increasing in the quantity of attack resources, a, and falling in the parameter of convexity, β.

Proposition 2 *Assume that (A.1) holds, $d = 0$, and suppose $f(n) = n^{\beta}$, where $\beta > 1$. If $C(a, \beta) \in \{a + 1, .., n\}$ and divides n, then the unique equilibrium network consists of $C(a, \beta)$ equal size components.*

Figure 4 illustrates the comparative statics with respect to attack resources and convexity of the network value function. We take $n = 24$. First, consider the effects of varying the attack resources. Here we set $\beta = 2$. The optimal number of components increases from 4 to 8, as we increase attack resources from 2 to 4. Second, consider the effects of convexity. Here we set $a = 2$. The optimal number of components falls from 4 to 3 as we raise the curvature by moving from $\beta = 2$ to $\beta = 3$.

When the Designer has no defense resources, his choice of optimal networks revolves around the number of components. Optimal networks contain equal size components whose number is falling in the convexity of the value function and increasing in the quantity of attack resources.

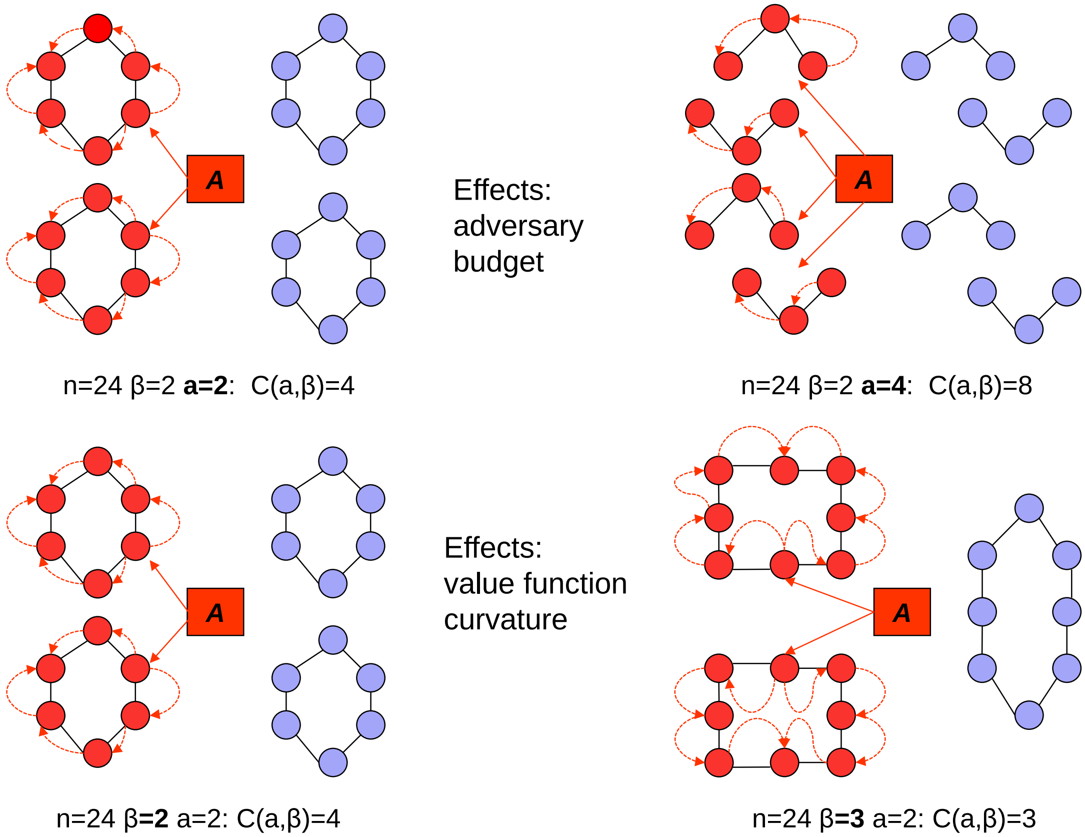
Sections 3.1 and 3.2 have covered the pure cases of optimal connected networks and optimal number of components, respectively. We now combine these insights and study optimal defended networks in a setting where defense allocation, architecture of individual components and the number of components are all decision variables for the Designer.

3.3 The general optimization problem of the Designer

A remarkable feature of Theorem 1 and Proposition 1 is that they make no assumptions on the degree of convexity of the network value function. However, Proposition 2 shows that the curvature of the network value function – f – is a crucial determinant of the number of components. It is, after all, the convexity of f which creates the tension between the pressure to connect nodes to create value and the threat of contagion via connections. Our next result builds on these results to characterize the circumstances under which optimal defended networks are connected.

Theorem 4 *Assume that (A.1) holds, $a/d \in \mathbb{N}$, $n > a + 1$, and let $\epsilon > 0$.*

1. *If $\ell < 1$ there exists n_0 such that, for all $n > n_0$, the CP-star is ϵ-optimal among all defended networks.*



n=24 β=2 **a=2**: C(a,β)=4

n=24 β=2 **a=4**: C(a,β)=8

n=24 β**=2** a=2: C(a,β)=4

n=24 β**=3** a=2: C(a,β)=3

Figure 4: Optimal networks: $f(m) = (m)^{\alpha}$, $n = 24$, β=2,3 and $a = 2, 4$

2. If $\ell = 1$ then optimal defended networks may contain multiple components.

When $\ell < 1$ network value grows exponentially in the number of nodes. In this case, the loss in value from splitting the network into multiple components can be made arbitrarily large, by suitably raising n. By contrast, when $\ell = 1$, optimal networks may consist of multiple components. Our proof exploits the resource configuration and the technology of conflict: if $a > d$, for large values of γ defense in a CP-star is not effective and this renders the alternative of separation of nodes into distinct components more attractive.

Theorem 4 covers resource configurations that satisfy $a/d \in \mathbb{N}$. When $a < d$, Proposition 1 and Theorem 4 taken together establish the following. For $\ell < 1$: if γ is large then the CP-star is ϵ-optimal, while if γ is small then a multiple-hubs network may be ϵ-optimal. If $\ell = 1$, then optimal networks may be disconnected. We conclude by noting that in the latter case, the optimal architecture and defence allocation for individual components is characterized by our results in section 3.1.

4 Discussion

Our framework of network design, defense and attack provides a useful way to think about a number of questions relating to networks that face threats. This section shows that by varying the network value function, the number of players, and the timing of moves we trace out an ensemble of models that can accommodate a wide range of applications. A complete analysis of these alternative models is outside the scope of the present paper; the exploratory analysis undertaken here suggests that arguments developed in the proofs of Theorems 1-4 can be applied to other games and also serves to bring out new insights that are consistent with empirical and applied work.

4.1 Decentralized linking and defense

In the benchmark model there is one Designer and one Adversary. In large scale computer networks, there are typically many players who can choose links and security.[12] Similarly, in social contexts, the spread of diseases depend on interaction and vaccination choices of individuals (Geoffard and Philipson (1997), Kremer (1996), Pongou and Serrano (2009)). In

[12] For a discussion on the theory of decentralized information networks, see Garicano (2000) and van Zandt (1999).

financial networks, banks make choices on linkages with other banks and also choose investments and level of reserves (Acemoglu, Ozdaglar and Tahbaz-Salehi (2013), Allen and Gale (2000), Cabrales, Gottardi and Vega-Redondo (2012), and Elliott, Golub and Jackson (2012)).

There are two natural variants within the decentralized decision making context. The first scenario involves a single Designer who chooses links but many players/nodes that choose security. This may correspond to the case where a central authority chooses an infrastructure while individual nodes choose defense or security levels. Individual security choices will generally create externalities on others (as in models of vaccination and epidemics). So the problem is to design a network in which these externalities are mitigated. The second scenario involves many players choosing links as well as security; here coordination problems arise in addition to the externalities present in the first scenario.

Our results, Theorems 1-4, are useful for the study of the decentralized problem as they set out the first best (or the planner) solution. This solution is a first step in the study of questions such as what is the the price of decentralization of links and of security (i.e., the difference between the social welfare attained in the first best and the expected welfare attained in the decentralized equilibrium).

4.2 Richer network value models

In the benchmark model, network value is strictly increasing and convex in number of nodes in a component. If the network value function is increasing but concave then network value can be enhanced by splitting up any component with multiple nodes into a collection of isolated nodes. Hence, the empty network would maximize value. The presence of contagious threats reinforces this pressure and the empty network remains optimal. There is, however, a range of possible alternatives between concave and convex network value functions. In particular, in some settings the marginal value of connections is initially increasing but then dissipates sharply. The aim of the example below is to draw out an implication of such network value functions for our arguments in Theorems 1 and 2.

Suppose that $n = 12$, $a = d = 2$, and the network value function is as follows:

$$f(n) = \begin{cases} n^2 & \text{for } 0 < n \leq 6 \\ 36 + 0.2(n-6) & \text{for } 6 < n \leq 12. \end{cases} \tag{5}$$

The probability distribution of the surviving nodes under CP-star is: probability 1/2 for 10 surviving nodes and probability 1/2 for 0 surviving nodes. The expected payoff from the CP-star is:

$$\frac{1}{2}f(10) = \frac{1}{2}(36 + 0.8) = 18.4. \tag{6}$$

Next consider the two hubs network (as in Figure 3). The expected payoff from a two hub network depends on the attack strategy of the Adversary. It may be checked that the Adversary prefers to attack periphery nodes attached to distinct hub nodes. The probability distribution of surviving nodes under the two hubs network is: probability 1/4 for 10 surviving nodes, probability 1/2 for 5 surviving nodes and probability 1/4 for 0 surviving nodes. The expected payoff of the Designer is:

$$\frac{1}{4}f(10) + \frac{1}{2}f(5) = 21.7. \tag{7}$$

Thus the two hub network dominates the CP-star.

The move from the CP-star network to the two hub defended network creates the following trade-off: the probability of 10 nodes surviving goes down from 1/2 to 1/4, but the probability of 5 nodes surviving goes up from 0 to 1/2. As the network value function is eventually linear, most of the potential network value is attained with the few initial nodes. So the increase in probability of 5 hubs surviving is more attractive for the Designer. If a significant part of the network value is attainable with a subset of the resources then multi-hub networks may be optimal.

An implicit assumption in the benchmark model is that there are no congestion effects; so traffic flows equally well through a single hub as through multiple hubs. In actual practice, both in computer networks as well as other infrastructure networks, it is likely that congestion effects are important. Large congestion costs will create a pressure toward multiple paths and the creation of multiple hubs. A general analysis of optimal networks in the presence of significant congestion costs remains an important open problem for future research.

4.3 Alternative timing of moves

In the benchmark model, we studied a sequential move game in which the Designer moves first, followed by the Adversary. In this section we show that by varying the order of moves, we can accommodate a variety of new applications.

Adversary moves first, followed by Designer: In some settings the Adversary is constrained to commit itself to a policy which is publicly observable. This may be due to political, legal or organizational reasons; a prominent instance is public policy with regard to crime.[13]

So suppose the Adversary moves first and chooses to allocate his budget $a \in \mathbb{N}$ across N nodes. The Designer observes this allocation and then chooses a network. To fix ideas suppose that $\ell = 1$. The Designer can then isolate all the nodes which are being attacked and constitute a component with the remaining un-captured nodes. A maximum of a nodes can be targeted: so the minimal payoff of the Designer is $f(n-a)$. As $\ell = 1$, it follows that for any $\epsilon > 0$, there is a $\bar{n}$, such that $f(n'-a) \geq (1-\epsilon)f(n')$, for all $n' \geq \bar{n}$. In other words, the Designer can ensure himself an expected payoff which is arbitrarily close to what he could attain in the absence of any Adversary.

This timing of moves allows us to relate our paper to Baccara and Bar-Isaac (2008) more closely. In their paper, attack resources are continuous variables and they suppose that $a_i \in [0,1]$. Fix $a = 1$ and suppose that $f(n) = n^2$. Consider the case of symmetric allocation $a_i = 1/n$. The payoff from a connected network is then simply the probability that it is not successfully attacked, which is $(1-1/n)^n f(n)$. It is possible to verify that as n gets large, the connected network dominates networks with multiple equal components. On the other hand, Baccara and Bar-Isaac (2008) show that, for small a_i, a network with binary cells is optimal. Clearly, in our setting a collection of binary cells is very unattractive.

This discussion abstracts from defense allocation: a more complicated design would involve protecting a subset of the attacked nodes and possibly linking these nodes. But this is a second order problem, given the high payoffs already attained.

The Simultaneous Game: In some contexts it may be possible to conceal the network structure and defense allocations: leading examples are criminal and terrorist networks and covert political protest movements. In addition to the government, the Adversary often includes intelligence agencies and secret services. These organizations may be able to keep their actions covert. These considerations motivate a game in which the Designer and Adversary make all choices simultaneously.[14] We have carried out a preliminary analysis of this game.

[13] See Baccara and Bar-Isaac (2008) for a detailed discussion of the reasons for such commitment.

[14] A referee has drawn our attention to a paper by Gueye and Marbukh (2012) who study a game in which the Designer picks a spanning tree from a network while an Adversary picks a link to delete. The aim of the Adversary is to maximize the loss to the Designer. They show the existence of mixed strategy equilibrium and their analysis highlights the role of link between-ness in understanding strategic behavior. The simultaneous game being considered in this section shares the same order of moves but there are crucial differences between the papers: we study optimal network defense and design and contagion plays a key role in our model. By

The details are presented in the appendix.

The analysis shows that, in equilibrium, both the Designer and the Adversary exploit simultaneity by mixing their strategies. Moreover, this opportunity for disguising the network will enable the Designer to earn higher payoffs as compared to the benchmark sequential model analyzed in Section 3. Our finding with regard to mixing by the Adversary echoes recent research on the practical value of mixed strategies as highlighted in the recent work of Tambe (2011) with the Los Angeles Police Department. On the other hand, our finding on the mixing by the Designer suggests that flexible networks are attractive for criminal and terrorist organizations. This is consistent with the prominent role of flexible networks – that permit quick reconfiguration of connections – in modern insurgencies (Arquilla and Ronfeldt, (1996, 2001), Zakaria (2008)).

Design followed by conflict: In some applications, the network is a physical object, e.g., transport or telecommunication infrastructure. Such a network takes time to build, is not easy to modify in the short run and is very visible. The resources of Designer and the Adversary represent personnel and equipment. These considerations motivate a model in which the Designer sets up a network; this network is observed by the Adversary and the two players then simultaneously choose the allocation of resources on this network.

Our analysis proceeds by way of an example about core-periphery networks: it shows that the Adversary and Designer have an incentive to mimic their resource allocations. This mimic behavior allows us then to exploit the mean preserving spread arguments developed in Theorem 1 and 2 to demonstrate that the star is optimal.

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. Figure 6 illustrates core-periphery networks with $n = 12$.

Recall that $\ell = \lim_{n\to\infty} \frac{f(n-1)}{f(n)}$. We are now ready to state the following result.

Proposition 3 *Assume that (A.1) holds. Let $a, d > 0$, $a/d \in \mathbb{N}$ Suppose $\ell = 1$. Then for large enough n, the star is optimal in the class of regular core-periphery networks.*

The proof is presented in the appendix. In the star network, given that $\ell = 1$ and n is large, there is a (Nash) equilibrium in which $\mathcal{D}$ allocates all resources to the central node and

contrast, in their work the Designer chooses only the spanning tree from a given network and there is no threat of contagion.

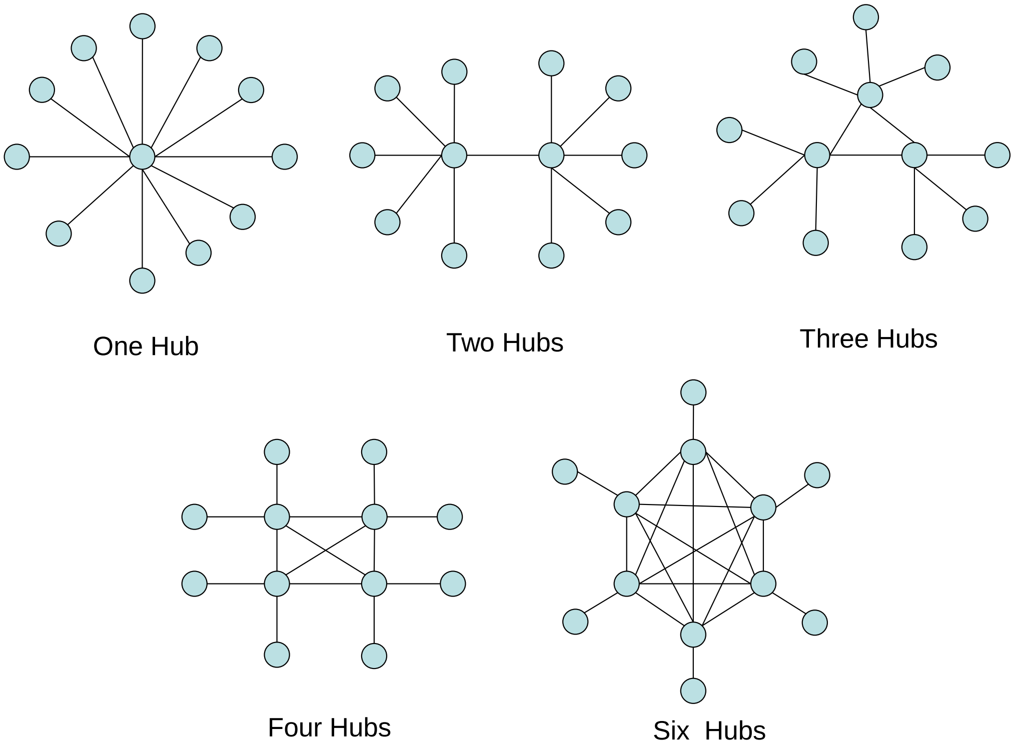


Figure 5: Regular core-periphery networks: $n = 12$.

the Adversary allocates all resources to peripheral nodes. The key step in the proof shows that in case of multiple hubs, it is optimal for $\mathcal{D}$ to allocate equal resources to each hub and for the Adversary to adopt a mimic strategy. The optimality of the mimic strategy lies in the nature of the conflict technology: it exhibits decreasing returns. The best response to equal allocations by the opponent is a mimic allocation. Given this equilibrium it then follows from arguments in Theorem 1-2 that the probability distribution of surviving nodes in the star is a mean preserving spread of the distribution obtained under a multiple core-periphery network. The result then follows from the assumption that f is convex.

This result suggests that the star with protected center is an attractive configuration for the Designer in settings beyond the benchmark model; a general characterization of optimal networks and defense remains an open problem.

5 Concluding remarks

Connections between individuals facilitate the exchange of goods, resources and information and create benefits. These connections may serve as a conduit for the spread of attacks and negative shocks as well. This paper studies the optimal design and defence of networks that face threats.

We develop a model with a Designer and an Adversary. The Designer moves first and chooses a network and an allocation of defense resources. The Adversary then allocates attack resources on nodes and determines how successful attacks should navigate the network. The model has three important ingredients: the value of the network, the technology of conflict between defense and attack resources, and the spread of attack through the network. We assume that the value of a network is increasing and convex in the number of interconnected nodes. We model the conflict between defense and attack resources on a node as a Tullock contest. If attack resources are successful in a contest, they can spread to neighboring nodes and cause contagion.

We obtain two principal results. One, we show that in a wide variety of circumstances a star network with all defence resources allocated to the central node is optimal for the Designer. Two, we identify conditions on the technology of conflict, network value function and the resource configuration for which networks with multiple hubs/components are optimal.

Empirical work on networks draws attention to the prominence of the hub-spoke network architecture (see e.g., Goyal, 2007; Newman, 2010). In an influential paper, Albert, Jeong and Barabasi (2000) argue that these architectures are vulnerable to strategic attacks since the Adversary can significantly reduce their functionality by removing only a few hub nodes. By contrast, our work highlights the attractiveness of these architectures in a setting where defence resources are scarce and network value is convex.

6 Appendix

Example 1 *Communication networks (Goyal (1993); Bala and Goyal (2000a))*

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; his payoff is $|C|$. As there are $|C|$ nodes in the component, the total payoff in component C is $|C|^2$. The aggregate social payoff in a network is the sum of the payoffs from the different components:

$$\sum_{C \in \mathcal{C}(g)} |C|^2. \tag{8}$$

The payoffs given in (8) satisfy assumption (*A.1*).

■

Lemma 1 *Let* $\{I_1, .., I_k\}$*,* $k \geq 2$*, denote a set of i.i.d. Bernoulli random variables with mean in* $(0, 1)$*. If* $f : \mathbb{R} \to \mathbb{R}$ *is convex then*

$$\mathbb{E}[f\big(\sum_{t=1}^{k} n_t I_t\big)] < \mathbb{E}[f\big((\sum_{t=1}^{k} n_t) I_1\big)].$$

Proof: Note first that it is enough to show that $(n_1 + .. + n_k)I_1$ is a mean-preserving spread of $n_1 I_1 + .. + n_k I_k$ (see e.g. Rothschild and Stiglitz (1970)).

Let $\tau = \mathbb{P}(I_i = 1)$, and 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 $1 - \tau = \alpha$. Then

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

and

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

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

$$\alpha - \alpha^2 = 1 - \tau^2 - \alpha$$

or, substituting for τ, if and only if

$$\alpha - \alpha^2 = 2\alpha - \alpha^2 - \alpha$$

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

Observe that that if Y is a MPS of X, for any random variable Q independent of X and Y, then $Y + Q$ is a MPS of $X + Q$.

But then setting $X = n_1I_1 + n_2I_2 + ... + n_kI_k$, $Y = (n_1 + n_2... + n_k)I_1$, $Q = n_{k+1}I_{k+1}$, 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 $n_1I_1 + n_2I_2 + n_3I_3.. + n_{k+1}I_{k+1}$.

■

Proof of Theorem 1: Let $(g, \underline{d}) \neq (g^s, \underline{d}^s)$ denote an arbitrary (connected) defended network satisfying (4). We will show that there exists a strategy $(\underline{a}, \Delta)$ such that $\Pi^e(g, \underline{d}, \underline{a}, \Delta) < \frac{d^\gamma}{d^\gamma + a^\gamma} f(n - a)$. Since, by (3), the right-hand side of this inequality is the payoff achieved by the Designer with a CP-star, this will establish the statement of the theorem.

Let here $K = \{i_1, ..., i_k\}$ denote the subset of protected nodes in $(g, \underline{d})$.

Case 1: $k = 1$

Since $(g, \underline{d}) \neq (g^s, \underline{d}^s)$ we can find two nodes in O with a link between them. By allocating one resource unit to one of these nodes we can then find an attack $(\underline{a}, \Delta)$ such that $\Pi^e(g, \underline{d}, \underline{a}, \Delta) \leq \frac{d^\gamma}{d^\gamma + a^\gamma} f(n - a - 1)$.

Case 2: $k > 1$

Construct the sequence of sets $\left(N_{i_s}\right)_{1 \leq s \leq k}$ recursively as follows:

$$N_{i_1} = O_{i_1} \ ; \ N_{i_2} = O_{i_2} - N_{i_1} \ ; \ ... \ ; \ N_{i_k} = O_{i_k} - \bigcup_{s=1}^{k-1} N_{i_s}$$

Let $n_{i_s} = |N_{i_s}|$, $s = 1, ..., k$. Note that by connectedness of g, $\bigcup_{s=1}^{k} N_{i_s} = O$.

Suppose first that $n_{i_s} \geq \frac{a}{d} d_{i_s}$, $\forall s$, and attack mimics defense in such a way that one resource unit is allocated to exactly $\frac{a}{d} d_{i_s}$ nodes in N_{i_s}, each of these resource units thereafter relocating to node i_s. Let Π^e denote the resulting expected payoff of the Designer.

Observe that, since $N_{i_s} \subset O_{i_s}$, a necessary condition for nodes in N_{i_s} to survive the attack is that i_s itself survives the attack. So the distribution of the total number of surviving nodes is first order stochastically dominated by that of $(n_{i_1} + 1 - a_{i_1})I_1 + .. + (n_{i_k} + 1 - a_{i_k})I_k$, where $\{I_1, .., I_k\}$ denotes a set of independent Bernoulli random variables such that $P(I_s = 1) = \frac{d^\gamma}{d^\gamma + a^\gamma}$, $\forall s \in \{1, .., k\}$. Since f is increasing and convex we have

$$\Pi^e \leq \mathbb{E}[f\big(\sum_{s=1}^{k}(n_{i_s} + 1 - a_{i_s})I_t\big)].$$

But, using Lemma 1:

$$\mathbb{E}[f\big(\sum_{s=1}^{k}(n_{i_s} + 1 - a_{i_s})I_s\big)] < \mathbb{E}[f\big((\sum_{s=1}^{k} n_{i_s} + 1 - a_{i_s})I_1\big)] = \mathbb{E}[f\big((n-a)I_1\big)].$$

Hence

$$\Pi^e < \frac{d^\gamma}{d^\gamma + a^\gamma} f(n-a).$$

Finally, since $(g, \underline{d})$ satisfies (4), it follows from Remark 1 in the text that we can always find a sequence of a nodes such that the first $\frac{a}{d} d_{i_1}$ of these nodes belong to O_{i_1}, the next $\frac{a}{d} d_{i_2}$ of these nodes belong to O_{i_2}, and so on up to k. So, by relabeling appropriately, the previous steps can be repeated in the case where $n_{i_s} < \frac{a}{d} d_{i_s}$ for some s.

$\blacksquare$

Proof of Theorem 2: By Theorem 1: $\overline{\Pi}^e(g^s, \underline{d}^s) > \overline{\Pi}^e(g, \underline{d})$ for any defended network satisfying (4). So we are only left to compare the performance of the CP-star with that of a network violating (4).

Case 1: $\ell < 1$

We will show that for any (connected) defended network $(g, \underline{d})$ violating (4) and for n large enough: $\overline{\Pi}^e(g, \underline{d}) < \frac{d^\gamma}{d^\gamma + a^\gamma} f(n-a)$.

Let $\epsilon' > 0$ such that $\ell' = \ell + \epsilon' < 1$. We can find n_0' such that $\frac{f(n-1)}{f(n)} < \ell'$, $\forall n \geq n_0'$. Then by induction $f(m) < (\ell')^{n-m} f(n)$, $\forall n \geq m \geq n_0'$.

Consider next a (connected) defended network $(g, \underline{d})$ violating (4). Let $i \in K$ such that $|O_i| \geq \frac{n-k}{k}$. Since $k \geq 2$, note that $d_i < d$. Suppose all attack resources are allocated in O_i, thereafter relocating to node i. Let Π^e denote the resulting expected network value. We have $\Pi^e \leq \frac{d_i^\gamma}{d_i^\gamma + a^\gamma} f(n-a) + \Gamma$, where $\Gamma \leq f(n - \frac{n-k}{k}) = f(1 + \frac{k-1}{k} n)$. Note from the remark above that for n large enough $f(1 + \frac{k-1}{k} n) < (\ell')^{\frac{n}{k}-1} f(n)$. Thus, for n large enough, $\Gamma < (\ell')^{\frac{n}{k}-1} f(n)$ and, finally:

$$\Pi^e < \frac{d_i^\gamma}{d_i^\gamma + a^\gamma} f(n-a) + (\ell')^{\frac{n}{k}-1} f(n). \tag{9}$$

Now let $\epsilon'' > 0$ such that $\ell'' = l - \epsilon'' > 0$. We can find n_0'' such that $\frac{f(n-1)}{f(n)} > \ell''$, $\forall n \geq n_0''$. Then by induction $f(n-a) > (\ell'')^a f(n)$, $\forall n \geq n_0'' + a$. For n large enough (9) now yields

$$\Pi^e < \left(\frac{d_i^\gamma}{d_i^\gamma + a^\gamma} + (\ell')^{\frac{n}{k}-1} (\ell'')^{-a} \right) f(n-a) \tag{10}$$

The first bracketed term in (10) is less than $\frac{d^\gamma}{d^\gamma + a^\gamma}$, since $d_i < d$, while the second term tends to 0 as n becomes large. We thus obtain $\Pi^e < \frac{d^\gamma}{d^\gamma + a^\gamma} f(n-a)$ for n large enough.

Case 2: $\ell = 1$

Let $(g, \underline{d})$ denote a defended network violating (4). Notice first that we can find n_0 such that $f(n-a) \geq (1-\epsilon) f(n)$ for all $n > n_0$. Consider an attack strategy such that $a_i = \frac{a}{d} d_i$, $\forall i \in K$. Let Π^e denote the resulting expected network value. It follows from the proof of Theorem 1 that $\Pi^e < \frac{d^\gamma}{d^\gamma + a^\gamma} f(n)$. So for $n > n_0$: $\frac{d^\gamma}{d^\gamma + a^\gamma} f(n-a) \geq (1-\epsilon) \frac{d^\gamma}{d^\gamma + a^\gamma} f(n) > (1-\epsilon)\Pi^e$. But this implies $\frac{d^\gamma}{d^\gamma + a^\gamma} f(n-a) > (1-\epsilon) \overline{\Pi}^e(g, \underline{d})$.

■

Claim 1 *Within the class of connected networks, a defended core network is optimal.*

Proof: Observe first that if there exists a path of unprotected nodes between two protected nodes, i and j say, then adding a link between i and j is without loss for the Designer. We can thus restrict attention to defended networks in which the set of protected nodes constitute a connected sub-graph, and such that any pair of protected nodes connected through a path of unprotected nodes are also directly linked.

Now suppose there does exist a path of unprotected nodes between two protected nodes, i and j say. Observe that, by the previous step, these unprotected nodes play no role in connecting i and j (or any other pair of protected nodes). So the alternative network in which all of these unprotected nodes have a single link to node i induces no loss for the Designer. Indeed, in the new network, these unprotected nodes' survival is contingent only on node i's survival, whereas in the old network it was contingent on node i and node j's survival. So any outcome in which node j is captured but node i is not in fact induces a strict gain for the Designer.

We are only left to show that a link between two unprotected nodes is never optimal. Suppose that i and j are unprotected and have a link between them. By the previous step there exists a unique protected node k connected to i and j through a path of unprotected nodes. Let I denote the set of unprotected nodes connected to k through a path of unprotected nodes. It is then immediate to see that the alternative network in which all nodes in I have a single link to node k weakly dominates the original network.

■

Proof of Proposition 1: For the first part, observe that as $\gamma \to \infty$ any contest involving $d_i > a_i$ results in certain success for the Designer.

For the second part, observe first that as $\gamma \to 0$ any contest involving $a_i > 0$ and $d_i > 0$ results in equal success probabilities for attack and defence.

We next show that any defended network in Λ with x nodes in the core and $1 < x < a+1$ is strictly dominated by the CP-star. Indeed, observe that in this case the Adversary's best response involves allocating (at least) one unit of resource to the periphery of each core node. The probability of success at each contest is $\frac{1}{2}$. This is also the probability of success given a CP-star. The central argument of Theorem 1 regarding MPS thus applies, and shows that the CP-star strictly dominates any such defended network.

Finally, we show that any defended network in Λ with x nodes in the core and $a < x < d$ is strictly dominated by d nodes in the core. In this case, the Adversary's best response involves allocating exactly one unit of resource to the periphery of a core nodes. But it is then easy to see that the resulting distribution of captured nodes FOSD that resulting with d nodes in the core. Since f is increasing, d nodes in the core therefore dominates.

Thus for γ small, the optimal defended network is either a CP-star or has d nodes in the core.

If $a = 1$, the Adversary's always attacks one periphery node. In that case, the distribution of captured nodes in a CP-star FOSD that resulting with d nodes in the core. Once again, since f is increasing, d nodes in the core dominates.

■

Proof of Theorem 3: First, we note that there must be at least $a + 1$ components: if the number of components is fewer than $a + 1$, then the Adversary can set $a_i = 1$ for one node in each component and thereby ensure that the Designer earns zero payoff. A network with $a + 1$ components on the other hand, guarantees the Designer strictly positive payoff as at least one component survives any attack.

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 best response, the Adversary must capture 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 captured as part of the best response by the Adversary. But then the Designer 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.

Finally, we show that at most one component has size strictly smaller than the maximum size $\overline{s}$. Suppose we can find two such components. The Designer 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 while, due to convexity of f, payoffs to the Designer are strictly increased by this move.

■

Proof of Proposition 2: Consider a network g consisting of equal size components, and let m denote this size. Using arguments from Theorem 3 we find

$$\overline{\Pi}^e(g) = f(m)(\frac{n}{m} - a). \tag{11}$$

Simple algebra establishes that (11) is maximized at $m = \frac{n(\beta-1)}{a\beta}$.

Next, consider a network g' with all but one component having maximum size m', and one component of size s, $0 < s < m'$. Let $b = \frac{n-s}{m'}$ denote the number of maximum components in g'. By optimality of the Adversary's strategy:

$$\overline{\Pi}^e(g') = f(m')(b-a) + f(s). \tag{12}$$

Observe then that by convexity of f

$$\overline{\Pi}^e(g') < f(m')(b-a) + \frac{s}{m'}f(m').$$

Substituting for b and simplifying then yields

$$\overline{\Pi}^e(g') < f(m')(\frac{n}{m'} - a).$$

So, by the first step, a network with $\frac{\beta a}{\beta-1}$ equal size components dominates any network in which one component has less than maximum size. By Theorem 3, it then follows that a network with $\frac{\beta a}{\beta-1}$ equal size components is in fact optimal.

■

Proof of Theorem 4: For the first part, it follows from Theorem 2 that, for large n, we only need to compare the performance of the CP-star with that of unconnected defended networks.

Consider therefore defended network $(g, \underline{d})$, with g unconnected. Let C denote the largest component in g, $n_C = |C|$, and d_C the total amount of resources allocated to nodes in C. Note in particular that $d_C \leq d$, while $n_C < n$.

Suppose first $n_C \leq \frac{n}{2}$. As in Theorem 2, choose $\ell' < 1$ and n_0' such that $f(m) < (\ell')^{n-m} f(n)$, $\forall n \geq m \geq n_0'$. For $n > 2n_0$ and irrespective of attack the network value is then bounded above by $2(\ell')^{\frac{n}{2}} f(n)$. Now, again as in Theorem 2 choose $\ell'' > 0$ and n_0'' such that $f(n-a) > (l'')^a f(n)$, $\forall n \geq n_0''$. Then for n large enough and irrespective of attack the network value is bounded above by $2(\ell')^{\frac{n}{2}}(\ell'')^{-a} f(n-a)$. A comparison with (3) establishes that the CP-star dominates $(g, \underline{d})$, for large enough n.

Assume henceforth $n_C > \frac{n}{2}$. By Theorem 2 we can find an attack on g with resulting expected network value Π^e such that

$$\Pi^e \leq \frac{d_C^\gamma}{d_C^\gamma + a^\gamma} f(n_C - a) + f(n - n_C).$$

But for $n \geq n_0' + n_C$ we have $f(n - n_C) \leq (\ell')^{n_C} f(n)$. Since $\ell' < 1$ and $n_C > \frac{n}{2}$, we obtain:

$$\Pi^e < \frac{d_C^\gamma}{d_C^\gamma + a^\gamma} f(n_C - a) + (\ell')^{\frac{n}{2}} f(n).$$

Using the fact that $f(n) < (\ell'')^{-a} f(n-a)$ for $n \geq n_0''$ we then have, for n large enough:

$$\Pi^e < \frac{d_C^\gamma}{d_C^\gamma + a^\gamma} f(n_C - a) + (\ell')^{\frac{n}{2}}(\ell'')^{-a} f(n-a).$$

Finally, $n_C < n$, and so

$$\Pi^e < \left(\frac{d_C^\gamma}{d_C^\gamma + a^\gamma} + (\ell')^{\frac{n}{2}}(\ell'')^{-a}\right) f(n-a). \tag{13}$$

The first bracketed term in (13) is at most $\frac{d^\gamma}{d^\gamma + a^\gamma}$, since $d_C \leq d$, while the second term tends to 0 as n becomes large. This completes the proof of the first part of the proposition.

For the second part, suppose $f(n) = n^2$, $d = 1$, $a = 2$. In the class of connected networks, Theorem 2 tells us that CP-star is optimal. In the CP-star network, the expected payoff is:

$$\overline{\Pi}^e(g^s, \underline{d}^s) = \frac{1}{1 + 2^\gamma}(n-2)^2. \tag{14}$$

Now let $(g, \underline{d})$ denote a defended network consisting of two components of equal size; suppose one component is a star with defended central node. We then have

$$\overline{\Pi}^e(g, \underline{d}) = \frac{1}{2}(\frac{n}{2} - 1)^2. \tag{15}$$

If $n \geq 4$, then $\overline{\Pi}^e(g, \underline{d}) > \overline{\Pi}^e(g^s, \underline{d}^s)$, for all $\gamma > 5$.

■

Analysis of the simultaneous move game: A mixed strategy of the Designer is a probability distribution, σ, on the set of networks and defense allocations. The mixed strategy of the Adversary, ρ, is a probability distribution on the set of attack allocations.[15] The expected payoff to $\mathcal{D}$ from strategy σ when $\mathcal{A}$ chooses ρ is:

$$\sum_{(g,\underline{d}) \in \text{supp } \sigma; \underline{a} \in \text{supp } \rho} \sigma(g, \underline{d})\rho(\underline{a}) \sum_{g' \in G(g)} P(g'|\underline{a}, \underline{d}, g) \left[\sum_{C_k \in \mathcal{C}(g')} f(|C_k(g')|)\right] \tag{16}$$

Consider a network with large n and two hubs who are each linked to $(n-2)/2$ nodes. Suppose for simplicity that $a = d = 2$ and that Designer allocates 1 resource unit to each hub. Suppose the Adversary targets one peripheral node each attached to different hubs. Then the probability distribution of surviving nodes is $\bar{P}$: probability 1/2 for 0 surviving nodes, probability 1/4 for $(n-2)$ surviving nodes and probability 1/4 for $(n/2 - 1)$ surviving nodes. Figure 3 illustrates the dynamics and this distribution. Next suppose that the Adversary targets 2 peripheral nodes attached to the same hub node. Then the probability distribution

[15] We abstract from the issue of routing of winning attacks by the Adversary here; for simplicity we assume that winner's resource captures neighboring undefended nodes but does not travel to other defended nodes.

of surviving nodes is given by $\tilde{P}$: probability 4/9 for 0 surviving nodes probability 3/9 for $(n-2)$ surviving nodes and probability 2/9 for $n/2-1)$ surviving nodes.

%beginfigure[tp]

It is easy to verify that $\tilde{P}$ first order stochastically dominates $\bar{P}$. Since network value f is increasing, it follows that $\mathcal{D}$ favors the latter attack strategy. The Designer can enforce his favored distribution by mixing across the allocation of peripheral nodes to hubs. In the face of this mixing, the Adversary is indifferent between mixing and not mixing his attack allocation. This advantage of the Designer in the simultaneous game has a general implication: in equilibrium he must earn (weakly) more in the simultaneous game as compared to the benchmark sequential game analyzed in Section 3. This is because, given a CP-star, it is optimal for the Adversary to target a peripheral nodes. This means that there is a strategy which ensures the Designer expected payoff (3) in the simultaneous game. ■

Proof of Proposition 3: Fix the number of hubs to $k=2$. Then there is an equilibrium (in the set of pure strategies) in which the Designer allocates $d/2$ to each core-node while the Adversary allocates his resources to peripheral nodes (symmetrically across the two hubs). Label the core nodes 1 and 2.

Suppose the Adversary does choose the mimic strategy. Consider a defense allocation $d_1 = d/2 + x$, $d_2 = d/2 - x$. We show that it is optimal for the Designer to set $x = 0$.

Given that $\ell = 1$, it is optimal to allocate no resource to the peripheral nodes. Next, consider allocations on the two core nodes. 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:

$$f(\frac{n}{2}-1)\left[\frac{2da}{(d+2x+a)(d-2x+a)}\right]+f(n-2)\frac{d^2-4x^2}{(d+2x+a)(d-2x+a)} \tag{17}$$

Differentiating with respect to x, we get:

$$f(\frac{n}{2}-1)\left[\frac{-8x}{(d+2x+a)^2(d-2x+a)^2}\right]+f(n-2)\left[\frac{(-8x)((d+2x+a)(d-2x+a)-(d^2-4x^2))}{(d+2x+a)^2(d-2x+a)^2}\right].$$

Simplifying yields

$$\left[\frac{-8x}{(d+2x+a)^2(d-2x+a)^2}\right]\left[-f(\frac{n}{2}-1)+f(n-2)(a^2+2ad)\right].$$

This expression in negative if $\frac{f(n-2)}{f(n/2-1)} > \frac{1}{a^2+2ad}$. So the Designer allocates resources equally to the two core nodes if this inequality is satisfied. This inequality is satisfied for all functions f which satisfy $(A.1)$.

Now consider optimality of the Adversary's strategy in the face of an equal split of defense resources $d/2$ between the two hub nodes. The payoff to an attack strategy $a/2+x, a/2-x$ is given by

$$f(\frac{n}{2}-1)\left[\frac{2da}{(d+2x+a)(d-2x+a)}\right]+f(n-2)\frac{d^2}{(d+2x+a)(d-2x+a)}.$$

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 completes the argument for the case of 2 core nodes.

The argument is now easily generalized to cover $k \geq 2$ nodes. ■

7 References

1. Acemoglu, D., A. Ozdaglar and A. Tahbaz-Salehi (2013), Systemic Risk and Stability in Financial Networks, *NBER Working Paper 18727.*

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

3. Allen, F. and D. Gale (2000), Financial Contagion, *Journal of Political Economy*, 108, 1, 1-33.

4. Alpcan, T. and T. Basar(2011), *Network Security: A Decision and Game Theoretic Approach.* Cambridge University Press. Cambridge, England.

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

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

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

8. Aspnes, J., K. Chang, A. Yampolskiy (2006), Inoculation strategies for victims of viruses and the sum-of-squares partition problem, *Journal of Computer and System Sciences*,

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

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

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

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

13. 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.

14. Blume, L., D. Easley, J. Kleinberg, R. Kleinberg, E. Tardos (2011), Network Formation in the Presence of Contagious Risk. *Proc. 12th ACM Conference on Electronic Commerce.*

15. Bollobas, B (1998), *Modern Graph Theory*, Springer-Verlag. New York.

16. Cabrales, A., P. Gottardi and F. Vega-Redondo (2013), Risk-sharing and contagion in networks. Mimeo, EUI Florence.

17. Cunningham, W. (1985), Optimal Attack and Reinforcement in a Network, *Journal of the ACM*, 32, 3, 549-61.

18. Elliott, M., B. Golub and M.O. Jackson (2012), Financial Networks and Contagion. *Mimeo.* Caltech and Stanford.

19. Farrell, F. and G. Saloner (1986), Installed base and compatibility: Innovation, product preannouncements, and predation *American Economic Review*, 76, 940-955.

20. Garfinkel, M. and Skaperdas, S. (2012), *The Oxford Handbook of the Economics of Peace and Conflict.* Oxford University Press.

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

22. Geoffard, P-Y., and Philipson, T. (1997), Disease eradication: private versus public vaccination, *American Economic Review*, 87(1):222-230.

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

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

25. Gueye, A. and V. Marbukh (2012), Toward a network of communication network vulnerability to attack: a game theoretic approach. *mimeo*, National Institute of Standards and Technology.

26. Gueye, A., Walrand, J., and V. Anantharam (2011), A network topology design game: How to choose communication links in an adversarial environment, *Proceedings of the 2nd International ICTS Conference on Game Theory for Networks.*

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

28. Hirshleifer, D. (1995), Theorizing about conflict. *Mimeo*, UCLA.

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

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

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

32. Katz, M. and C. Shapiro, (1985), Network Externalities, Competition and Compatibility, *American Economic Review*, 75, 3, 424-440.

33. Konrad, K. (2009), *Strategy and Dynamics in Contests.* Oxford University Press.

34. Kovenock, D. and B. Roberson (2012), Conflicts with multiple battle fields, in Garfinkel, M. and Skaperdas, S. (eds), *The Oxford Handbook of the Economics of Peace and Conflict.* Oxford University Press.

35. Kremer, M. (1996), Integrating Behavioral Choice into Epidemiological Models of AIDS, *Quarterly Journal of Economics*, 111, 2, 549-73.

36. Laszka, A. and Szeszlér, D. and Buttyán, L. (2012), Linear Loss Function for the Network Blocking Game: An Efficient Model for Measuring Network Robustness and Link Criticality, *Proceedings of the 3rd International Conference on Decision and Game Theory for Security.*

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

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

39. Newman, M. (2010), *Networks: an introduction.* Oxford University Press.

40. Pongou, R., and R. Serrano (2009) A Dynamic Theory of Fidelity Networks with an Application to the Spread of HIV/AIDS, *Working Paper 2009-02*, Department of Economics, Brown University.

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

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

43. Roy, S., C. Ellis, S. Shiva, D. Dasgupta, V. Shandilya, Q. Wu, (2010), A Survey of Game Theory as Applied to Network Security, *Proceedings of the 43nd Hawaii International Conference on System Sciences (HICSS-43)*, 10 pages, IEEE Computer Society.

44. Saia, J., A.Fiat, S.Gribble, A.R.Karlin, and S.Saroiu (2002), Dynamically fault-tolerant content addressable networks, in *Proceedings of the 1st International Workshop on Peer-to-Peer Systems (IPTPS '02).*

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

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

47. Staniford, S., V. Paxson and N. Weaver (2002), How to own the Internet in your spare time, *Proceedings of the 11th USENIX Security Symposium*, 149-167.

48. Suto, K., H. Nishiyama, X. Shen, N. Kato (2012), Designing P2P Networks Tolerant to Attacks and Faults Based on Bimodal Degree Distribution, *Journal of Communications* 7, 8, 587-595,

49. Tambe, M. (2011), *Security and Game Theory.* Cambridge University Press.

50. 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.

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

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

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

Attack, Defence and Contagion in Networks
On-line Appendix

A model with continuous allocation of defence and attack resources

Our basic model assumes integer allocation of defence and attack resources. The aim of this appendix is to drop this assumption and develop a simple version of our model with continuous allocation of defence and attack resources.

The **Designer** now has $d \in \mathbb{R}$ resource units to allocate across the nodes, where $d \geq 1$. Let $\underline{d} = (d_1, d_2, ..., d_n)$ denote this allocation, where $d_i \in \mathbb{R}$ and $\sum_{i \in N} d_i \leq d$. Say that a path from node i to node j is *weak* if all nodes on that path and node j itself all have strictly less than one unit of defence resources. Similarly, the **Adversary** now has $a \in \mathbb{R}$ resource units to allocate across the nodes. The vector $\underline{a} = (a_1, a_2, .., a_n)$, where $a_i \in \mathbb{R}$ and $\sum_{i \in N} a_i \leq a$, denotes his initial allocation of attack resources.

The Tullock contest function (equation (2) in our paper) supposes that an unprotected node attacked with $\varepsilon_i > 0$ resource units is captured with probability 1. We assume instead that the probability of successful attack is low when the amount of attack resources allocated are small. Formally, we set

$$\mathsf{probability\ of\ successful\ attack} = \min\{a_i, \frac{a_i^\gamma}{a_i^\gamma + d_i^\gamma}\} \tag{1}$$

Finally, we simplify the dynamics of attack by assuming that capture of a node, i say, spreads to node j if and only if there exists a weak path from node i to node j.

Theorem 1 *Assume that (A.1) holds, $\ell = 1$, and $n > a + 1$. Let $\epsilon > 0$ and consider the class of connected networks. There exists n_0 such that, for all $n > n_0$, the CP-star is ϵ-optimal.*

Proof: We first show that, for n large, then given a CP-star the Adversary's best response consists in allocating all resources to the central node. The following intermediary claim is useful.

Claim: Given a CP-star, then in the Adversary's best response at most one periphery node has $0 < a_i < 1$.

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 the adversary can do better by transferring a small amount of resources from

i_1 to i_2.

In the event that the central node is captured then so are i_1 and i_2, by contagion. Shifting attack resources from i_1 to i_2 thus makes no difference in this case. Hence suppose the central node survives, and let m denote the total number of surviving nodes other than i_1 and i_2. The payoff of the designer, conditional on this event, is (observe that any surviving nodes are moreover connected)

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

Denoting $\overline{a}=a_{i_1}+a_{i_2}$, the former expression becomes

$$(1-\overline{a})f(m+2)+\overline{a}f(m+1)+a_{i_1}(\bar{a}-a_{i_1})\big[f(m+2)-2f(m+1)+f(m)\big].$$

Now, by convexity of f we have $f(m+2)-2f(m+1)+f(m)>0$. So the expression is increasing in a_{i_1}. This establishes the claim, since in the above m was arbitrary.

Using the Claim allows us to 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 (the treatment of the case where one periphery node is allocated strictly less than one unit of attack resources is similar, and therefore omitted). The payoff of the designer under such an attack is $\frac{d}{c+d}f(n-a+c)$. Since $\ell=1$, for n large this last expression is decreasing in c. The Adversary's best response to a CP-star thus consists in allocating all resources to the central node, yielding

$$\overline{\Pi}^e(g^s,\underline{d}^s)=\frac{d^\gamma}{d^\gamma+a^\gamma}f(n). \tag{2}$$

The remaining arguments of the proof are identical to those developed in the proof of Theorem 2 in our paper. For any defended network $(g,\underline{d})\neq(g^s,\underline{d}^s)$ notice that the Adversary can always approximate a mimic attack strategy by allocating $a_i=\frac{a}{d}d_i$.

■