Centre de Referència en Economia Analítica

Barcelona Economics Working Paper Series

Working Paper nº 176

Optimal Targets in Small and Large Networks Using Game Theory

Antoni Calvó-Armengol and Inés Moreno de Barreda

May, 2005

Optimal Targets in Small and Large Networks, Using Game Theory

Antoni Calvó-Armengol*

Inés Moreno de Barreda†

Abstract. We define a model of peer effects where the intra-group externality is rooted on the network of bilateral influences in the population, rather than consisting on an average effect. Using game theory, we then map the geometric intricacies of this network structure to the distribution of equilibrium outcomes. Nash equilibrium turns out to be well-described by Bonacich network centrality, used in sociology. We then exploit the network variance of peer effects to identify optimal network targets, key groups. Key groups correspond to the highest inter-central groups, a new network measure that subsumes collective optimality concerns. Although intended for small networks, the key group policy coupled with a more standard geometric attack turns out to be optimal for large scale free networks when $2.33 < \beta < 3$. We then apply our model to terrorist networks, and identify the Achille's heel of the 11S cell. (This version: May, 30 2005)

The influence of group outcomes on its members' individual behavior is documented in many ethnographic and empirical studies, and ranges from criminal activity, teen pregnancy and drug use to academic achievements and labor outcomes (Durlauf 2004). The generative mechanisms of peer effects, though, remain a black-box. In fact, most models of peer effects simply assume an average group influence, and derive consequences. But peer effects often arise in small groups, for which averages need not be well defined. Besides, averaging presumes that the intra-group externality is homogeneous across group members, which rules out right away any potential heterogeneity in group exposure. And this closes the door to policy interventions tailored to the pattern, rather than just the aggregate features, of the intra-group influences.

Here, we take the opposite direction. First, we start with a detailed description of the intra-group interaction structure. This is simply the collection of all the bilateral influences in the group. Second, we map the inner intricacies of this interaction structure to the distribution of peer effects across group members. Third, we exploit this variance of peer effects to identify optimal group targets so as to manipulate the overall group outcome. We then apply this general model to terrorist networks, and identify the Achille's heel of the 11S network. Lastly, we extend our model, originally intended for small networks modelled as graphs, to the case of large networks described by statistical objects.

A general model of peer networks Each agent $i = 1, ..., n$ selects an effort $x_i \geq 0$, and gets a payoff $u_i(x_1, ..., x_n)$ strictly concave in own effort x_i. Net of bilateral influences, agents are identical. Bilateral influences $\partial^2 u_i / \partial x_i \partial x_j = \sigma_{ij}$ are a given that can vary across pairs. When $\sigma_{ij} > 0$, an increase in the effort of j creates an incentive for i to increase his effort in turn. We talk of strategic complementarity in efforts. When $\sigma_{ij} < 0$, instead, an extra effort from j triggers a downwards shift in i's effort in response. We say that efforts are strategic substitutes.

Let $\underline{\sigma} = \min\{\sigma_{ij} \mid i \neq j\}$ and $\overline{\sigma} = \max\ \{\sigma_{ij} \mid i \neq j\}$. Assume that $\partial^2 u_i / \partial x_i^2 < \min\{\underline{\sigma}, 0\}$, that is, own marginal returns decrease more steeply with x_i than any cross marginal returns do.

We decompose additively the matrix of cross effects $\mathbf{\Sigma} = [\sigma_{ij}]$ into an idiosyncratic concavity matrix, a global (uniform) substitutability matrix, and a local complementarity matrix.

Let $\lambda = \overline{\sigma} - \underline{\sigma}$. The case $\overline{\sigma} = \underline{\sigma}$ is straightforward, and take $\lambda > 0$. λ measures the dispersion in cross effects.

Let $\gamma = -\min\{\underline{\sigma}, 0\} \geq 0$. If $\underline{\sigma} \geq 0$, then $\gamma = 0$. Otherwise, $\underline{\sigma} < 0$ and $\gamma > 0$.

Let $g_{ij} = (\sigma_{ij} + \gamma)/\lambda$, for $i \neq j$, and $g_{ii} = 0$. This is just a centralization followed by a normalization of the cross effects. By construction, $0 \leq g_{ij} \leq 1$. $\mathbf{G} = [g_{ij}]$ is the adjacency matrix of a network $\mathbf{g}$ that captures the strength of relative payoff comple-

*Corresponding author: ICREA, Universitat Autònoma de Barcelona, Université de Toulouse, and CEPR. Address of correspondence: Department of Economics and Economic History, UAB, Edifici B, 08193 Bellaterra (Barcelona), Spain. Email: antoni.calvo@uab.es. http://selene.uab.es/acalvo

†Universitat Autònoma de Barcelona, Department of Economics and Economic History.

mentarities across all pairs. When $\sigma_{ij} = \sigma_{ji}$, $\mathbf{G}$ is symmetric, and $\mathbf{g}$ is un-directed. When $\sigma_{ij} \in \{\underline{\sigma}, \overline{\sigma}\}$ for all $i \neq j$ with $\underline{\sigma} \leq 0$, $\mathbf{G}$ is a $(0,1)-$matrix, and $\mathbf{g}$ is un-weighted.

Finally, let $\sigma_{ii} = -\psi - \gamma$, where $\psi > 0$.

Let $\mathbf{U}$ denote the square matrix of ones. Then:

$$\mathbf{\Sigma} = -\psi\mathbf{I} - \gamma\mathbf{U} + \lambda\mathbf{G}.$$

Bilateral influences boil down into the combination of an idiosyncratic effect, a global interaction, and a local interdependence. The global effect $-\gamma\mathbf{U}$ is uniform across group members. The matrices $-\psi\mathbf{I}$ and $\lambda\mathbf{G}$ are here to compensate for this global effect when required. First, as $\sigma_{ii} < -\gamma$, an extra negative shift is added to the diagonal terms $\partial^2 u_i/\partial x_i^2$ of $-\gamma\mathbf{U}$. This is $-\psi\mathbf{I}$. Second, when $\sigma_{ij} > \underline{\sigma}$, an extra positive shift is added to the out-of-diagonal terms $\partial^2 u_i/\partial x_i \partial x_j$ of $-\gamma\mathbf{U}$. This is $\lambda\mathbf{G}$.

A linear-quadratic approximation of the payoff function is then:

$$u_i(\mathbf{x};\mathbf{\Sigma}) = \alpha x_i - \frac{1}{2}(\psi-\gamma)x_i^2 - \gamma\sum_{j=1}^{n} x_i x_j + \lambda\sum_{j=1}^{n} g_{ij} x_i x_j.$$

We take $\alpha > 0$.

The Bonacich-Nash linkage We use game theory to map the network of cross influences $\mathbf{g}$ to individual outcomes.

Agents choose their efforts simultaneously and independently. Then, likely outcomes correspond to Nash equilibria, where each agent optimizes his choice taken the others' as a given. An interior Nash equilibrium $\mathbf{x}^*(\mathbf{\Sigma}) = (x_1^*, ..., x_n^*)$ is such that $\partial u_i/\partial x_i(\mathbf{x}^*) = 0$ and $x_i^* > 0$, for all i. Let $\mathbf{1}$ be the vector of ones. Existence of interior Nash is obtained whenever the linear system:

$$-\mathbf{\Sigma}\cdot\mathbf{x} = [\psi\mathbf{I} + \gamma\mathbf{U} - \lambda\mathbf{G}]\cdot\mathbf{x} = \alpha\mathbf{1} \qquad (1)$$

has a non-negative solution. Note that this system has a unique solution everywhere, except on a set of Lebesgue measure zero. We provide conditions such that this unique generic solution is non-negative.

Absent any bilateral influence ($\gamma = \lambda = 0$), all agents choose the same equilibrium effort α/ψ. Otherwise, agents' equilibrium efforts depend on the pattern of cross effects ($\mathbf{g}$) and their intensity and sign (γ and λ).

We first comment on the role of γ and λ. Let $n = 2$ and $g_{12} = g_{21} = 1$. When $\lambda < \psi + 2\gamma$, the equilibrium efforts in the dyad are $x_i^* = \alpha/(\psi + 2\gamma - \lambda)$. If $\gamma = 0$ and $\lambda > 0$, agents reap complementarities from their dyad partner, and choose an effort level above the optimal value for an isolated agent, α/ψ. When $\lambda = 0$ and $\gamma > 0$, equilibrium efforts decrease in γ as global substitutability adds to the concavity in own efforts and exhausts marginal returns below the optimal single agent value.

More generally, let $\mathbf{M}(\mathbf{g},\lambda^*) = [\mathbf{I} - \lambda^*\mathbf{G}]^{-1}$, where $\lambda^* = \lambda/\psi$. $\mathbf{M}(\mathbf{g},\lambda^*)$ is well-defined and non-negative if and only if λ^* is strictly smaller than the inverse of the largest eigenvalue of $\mathbf{G}$ (Debreu and Herstein 1953). Then:

$$\mathbf{M}(\mathbf{g},\lambda^*) = [\mathbf{I} - \lambda^*\mathbf{G}]^{-1} = \sum_{\ell=0}^{+\infty} \lambda^{*\ell}\mathbf{G}^{\ell}.$$

This infinite sum converges by the bound on λ^*.

The components $m_{ij}(\mathbf{g},\lambda^*) \geq 0$ of $\mathbf{M}$ count the total number of paths on $\mathbf{g}$ between i and j, with $\ell-$length paths discounted by $\lambda^{*\ell}$. The total number of paths spanning from i in $\mathbf{g}$, $b_i(\mathbf{g},\lambda^*) = \sum_{j=1}^{n} m_{ij}(\mathbf{g},\lambda)$, is the Bonacich network centrality of agent i in the network $\mathbf{g}$ (Bonacich 1987).

Under the condition on λ^*, the Nash equilibrium is uniquely defined and proportional to the vector of Bonacich centralities (Ballester *et al.* 2005a):

$$\mathbf{x}^*(\mathbf{\Sigma}) = \frac{\alpha}{\psi + \gamma b(\mathbf{g},\lambda^*)}\mathbf{b}(\mathbf{g},\lambda^*), \qquad (2)$$

where $b(\mathbf{g},\lambda^*) = \sum_{i=1}^{n} b_i(\mathbf{g},\lambda^*)$.

Let $x^* = \sum_{i=1}^{n} x_i^*$. This is the aggregate equilibrium effort, increasing in α and λ^*, but decreasing in γ and ψ. More generally, x^* increases in every σ_{ij}. Strengthening a single cross effect spills over to a higher overall group outcome (Ballester *et al.* 2005a). In particular, when $\mathbf{g}$ is un-weighted, $\mathbf{g} \subset \mathbf{g}'$ implies $x^*(\mathbf{g}') > x^*(\mathbf{g})$.

Now, the intra-group externality is not homogeneous across agents, but varies with network location:

$$x_i^*(\mathbf{\Sigma}) = \frac{b_i(\mathbf{g},\lambda^*)}{b(\mathbf{g},\lambda^*)} x^*(\mathbf{\Sigma}).$$

Bonacich centrality captures the variance in peer effects, and measures individual exposure to the group influence.

Network inter-centrality and key player The Bonacich-Nash linkage has implications for public policy. A standard policy intervention fine-tunes the exogenous payoff parameters (α, ψ, γ and λ) to achieve the desired objective on, say, group outcome. Instead, the planner can decide to manipulate the network of cross effects $\mathbf{g}$, for instance by removing some agents from the group. By doing so, the pattern of cross effects changes, and so does the aggregate group outcome.

We identify the optimal network targets, the key groups. Key groups are such that, once removed, the planner achieves the highest overall reduction in group outcome, compared to any alternative subgroup removal of same size.

For concreteness, let the planner remove a single agent k. $\mathbf{\Sigma}^{-k}$ is the matrix obtained from $\mathbf{\Sigma}$ by setting $\sigma_{ik} = \sigma_{ki} = 0$, for all i. The key player k^* solves:

$$k^* \in \arg\max\{x^*(\mathbf{\Sigma}) - x^*(\mathbf{\Sigma}^{-k}) \mid k = 1, ..., n\}. \quad (3)$$

This is a finite optimization problem, with at least one solution. From now on, take $\mathbf{g}$ un-directed.

Define the inter-centrality of node i by:

$$c_i(\mathbf{g}, \lambda^*) = b_i(\mathbf{g}, \lambda^*) + \sum_{j \neq i} [b_j(\mathbf{g}, \lambda^*) - b_j(\mathbf{g}^{-i}, \lambda^*)].$$

This is the sum of i's own centrality plus i's contribution to every other agent centrality. It turns out that the key player solving (3) is the agent with highest inter-centrality (Ballester *et al.* 2005a):

$$k^* \in \arg\max\{c_k(\mathbf{g}, \lambda^*) \mid k = 1, ..., n\}.$$

Bonacich centrality measures ego-centered influences, and captures individual strategic behavior. Inter-centrality also accounts for cross-contributions in ego-centered influences, and captures group optimal concerns. In fact:

$$c_k(\mathbf{g}, \lambda^*) = \frac{b_k^2(\mathbf{g}, \lambda^*)}{m_{kk}(\mathbf{g}, \lambda^*)}.$$

$m_{kk}(\mathbf{g}, \lambda^*)$ counts the total number of self-loops in $\mathbf{g}$ that start and end at k. Holding $b_k(\mathbf{g}, \lambda^*)$ constant, k's inter-centrality decreases with the share in Bonacich centrality due to self-loops, $b_k(\mathbf{g}, \lambda^*)/m_{kk}(\mathbf{g}, \lambda^*)$.

More generally, the group inter-centrality of $S = \{i_1, ..., i_s\}$ with cardinality $\#S = s$ is:

$$c_S(\mathbf{g}, \lambda^*) = c_{i_1}(\mathbf{g}, \lambda^*) + ... + c_{i_s}(\mathbf{g}^{-i_1 - ... - i_{s-1}}, \lambda^*),$$

for every labeling of members in S. The key group S^* of size s then has highest group inter-centrality:

$$S^* \in \arg\max\{c_S(\mathbf{g}, \lambda^*) \mid \#S = s\}.$$

In practice, finding the key group is computationally demanding, an NP-hard problem. A simple greedy algorithm that iteratively chooses an optimal vertex from the network provides an approximated solution. This algorithm is in polynomial time, with an approximation error bounded from above by $1/e \approx 36.79\%$ (Ballester *et al.* 2005b).

Terrorist networks We now apply the general peer networks model to terrorist networks. We start with a network payoff function that captures the main concerns of a terrorist network organization. Then, the Bonacich-Nash linkage associates a level of activity to each terrorist as a function of his location in the conspiracy network. Finally, the key group policy identifies optimal group targets in the terrorist network.

The network disruption policy we propose tackles purposefully and directly the terrorist organization's activities, and causes the major possible harm to it. The geometric disruption of the terrorist network is a means to destabilize the group activities, rather than an objective *per se*.

Title 22, Section 2656f(d) of the United States Code defines terrorism as a "premeditated, politically motivated violence perpetrated against noncombatant targets by sub-national groups or clandestine agents, usually intended to influence an audience." Thus, terrorist organizations need to reconcile two antagonistic interests. On one hand, they mobilize and coordinate resources to perpetrate agreed-upon and purposeful targeted actions. On the other hand, they face the permanent external threat of law enforcement and thus seek to remain clandestine. In substance, terrorist and conspiracy organizations face an inherent trade-off between internal coordination and external vulnerability (McCormick and Owen 2000, Baccara and Bar-Isaac 2005).

Network forms of organization solve this trade-off adequately. First, network links allow coordination across directly linked members, while the network overlap across such sub-entities induces coordination for the whole organization. Coordination both before and during the attacks is, indeed, a key pre-requisite for the kind of high-scale lethal actions of 11S in New-York or 11M in Madrid. Second, decision-making in a networked organization is, by essence, highly decentralized. Each sub-entity can react on its own and carry on when others' are destroyed. As such, networks constitute more flexible, adaptive and resilient structures than classical hierarchies, where a breach in the chain of command jeopardizes the whole group operations. In fact, the late unraveling of the clandestine groups responsible for the 11S and 11M attacks found out sparse and decentralized network organizations (Krebs 2002, Rodríguez 2004).

A terrorist organization is modelled by an un-directed and un-weighted network $\mathbf{g}$ involving n terrorists, with $g_{ij} \in \{0,1\}$ and $g_{ij} = g_{ji}$. Network links correspond to coordination channels. x_i is the level of operations by terrorist i, and $x = \sum_{i=1}^{n} x_i$ the overall group activity.

The terrorist clandestine organization faces a trade-off between coordination and vulnerability. This is reflected in the following objective function, at the group level:

$$u(x_1, .., x_n; \mathbf{g}) = \alpha x + \lambda \sum_{i=1}^{n} \sum_{j=1}^{n} g_{ij} x_i x_j - \gamma x^2. \quad (4)$$

The vulnerability cost is $-\gamma x^2$. Vulnerability increases with the level of activity in proportion with the current volume of operations. The group outcome is $\alpha x + \lambda \sum_{i=1}^{n} \sum_{j=1}^{n} g_{ij} x_i x_j$, and increases with the collaborative links in the organization network. The vulnerability and coordination parameters $(\gamma, \alpha, \lambda)$ depend on the environment where the organization operates, its communication possibilities, and level of internal conflict. A terrorist organization is fully described by $(\mathbf{g}, \alpha, \lambda, \gamma)$.

Suppose that all terrorists have the same number of links in $\mathbf{g}$. Let $g_i = \sum_{j=1}^{n} g_{ij} = g/n$. When $\gamma n > \lambda g$, the optimal level of operations maximizing (4) is then $\alpha n / 2(\gamma n - \lambda g)$. It increases with the value of each collaborative agreement (λ) and their number (g), but decreases with the vulnerability cost γ.

In a hierarchical organization, the chain of command and authority dictates the behavior of each member in the hierarchy. In a networked organization, instead, decision-making is decentralized at the level of each individual. The equilibrium that emerges from these decentralized decisions then acts as an implicit self-enforcing contract, or focal point, for the actions of the organization members. The network of collaborative agreements $\mathbf{g}$ becomes a key organization variable to determine this focal point.

Consider a terrorist organization $(\mathbf{g}, \alpha, \lambda, \gamma)$. The individual terrorists' objective functions are:

$$u_i(x_1, .., x_n; \mathbf{g}) = \alpha x_i + \lambda \sum_{j} g_{ij} x_i x_j - \gamma x_i x.$$

Individual objectives aggregate into the organization goal, $u(\mathbf{x}; \mathbf{g}) = \sum_{i=1}^{n} u_i(\mathbf{x}; \mathbf{g})$. The terrorists' objectives fit into our general peer network model:

$$\left[\frac{\partial^2 u_i}{\partial x_i \partial x_j}(x_1, ..., x_n; \mathbf{g})\right] = -\gamma \mathbf{I} - \gamma \mathbf{U} + \lambda \mathbf{G}.$$

Suppose that $\gamma > \lambda \sqrt{g + n - 1}$. Using a crude upper bound on highest eigenvalues (Cvetković *et al.* 1997), we conclude that terrorists's outcomes are:

$$\mathbf{x}^*(\mathbf{g}, \alpha, \lambda, \gamma) = \frac{\alpha}{\gamma} \frac{\mathbf{b}(\mathbf{g}, \lambda/\gamma)}{1 + b(\mathbf{g}, \lambda/\gamma)}.$$

The organization overall activity x^* is increasing and concave in the total number of paths $b(\mathbf{g}, \lambda/\gamma)$ in $\mathbf{g}$. In particular, when $g_i = g/n$ for all i, x^* is increasing and concave in the network connectivity g. Recall that, in this case, the optimal level of operations is $\alpha n / 2(\gamma n - \lambda g)$, an increasing and convex function of g. One can then easily compute, for every given $(\alpha, \lambda, \gamma)$, a uniquely defined value for g that decentralizes the optimal value of operations. More generally, the organizational design problem (that we do not explicitly tackle here) consists on characterizing the whole network geometry (not only its connectivity) whose corresponding focal point is optimal for (4) (Corbo *et al.* 2005, Jackson 2005).

The 11S terrorist cell We illustrate the key group policy for the case of the Al Qaeda cell that perpetrated the 11S attack.

We restrict our analysis to the 19 terrorists that were travelling on-board during the attacks. The network of communication links connecting them is reconstructed in Krebs (2002) with publicly released information on communication and friendship ties among the 19 terrorists. See Figure 1.

The network consists in four different and intertwinned sub-entities with four to five terrorists each. Each sub-entity constitutes a commando that was flying in the same plane. Members of a same commando need not know each other directly previous to the attack. Terrorists in the same plane are, sometimes, two to three steps away from in each other in the network. This network is sparse. The most connected terrorist knows directly only seven out of nineteen terrorists. The average path length is 4.75, and four intermediate brokers are needed, on average, to convey information between two distinct and distant terrorists. Low connectivity, high sparsity and high path length guarantee the security of the terrorist organization and its resilience to local dismantlement. One week before the 11S, though, the terrorists strengthened their network links to the extent that average path length decreased by as much as 40%, thus facing higher risks. The new and transitory contacts added to the dormant permanent network created network shortcuts useful for last-minute coordination. Maximal connectivity stays below seven all along.

Previous works single out optimal network targets by ranking terrorists according to some arbitrary centrality measure such as connectivity, closeness or betweenness. Connectivity just counts direct links, closeness measures distance to others, and betweenness counts shortest paths crossing through a given individual. All these measures are purely geometric in nature. Because they lead to different ranking of network nodes, they identify different optimal group targets.

Our approach is different.

First, we view individuals as pursuing objectives that reflect the coordination-vulnerability trade-off characteristic of a terrorist organization. These objectives depend on the pattern of collaborative links. Second, using game theory, we compute the organization's focal point for the decentralized decision-making process in the network. This focal point yields to a Bonacich-Nash linkage that maps network structure into group behavior. Third, we set up and solve the problem of optimally reducing the network level of operations by eliminating some suitably selected targets in the network. The key group optimal target is the one with highest inter-centrality. This novel network measure captures the inter-dependent pattern of behavior induced at equilibrium by the direct and indirect network links.

The inter-centrality ranking does not coincide with the connectivity, closeness or betweenness rankings. See Figure 2. They thus result in different policy prescriptions. Interestingly, for a target group of size four, and contrarily to alternative prescriptions, the key group policy removes scattered terrorists in the network and disrupts every existing plane commando. Note that the key group prescription relies on the inter-centrality measure which is completely agnostic about the internal network organization into four distinct commandos. Yet, the key player prescription dismantles the four commandos. Contrarily to other network measures, it is able to identify these four distinct sub-entities whose overlap generates the whole network.

To pursue the comparison across policy prescriptions based on different network measures, we compute the overall reduction in the organization output when the highest ranked terrorists are progressively removed, up to a maximum of five individual target removals (one quarter of the organization size). See Figure 3. We express the activity reduction in percentage terms.

Large networks Our game theoretical characterization of optimal network targets is well-suited (and intended) for small networks. In what follows, we extrapolate the key player policy to large scale networks described by statistical objects. We show that, even with arbitrary large populations, a key player attack is a first-best policy compared to a geometric attack in all scale free networks with connectivity parameter $7/3 < \beta < 3$.

Consider a large network described by its connectivity distribution. A network is scale free when this distribution follows a power law. Because of their high polarization in connectivity values, scale free networks are very fragile against geometric attacks directed to their network hubs (Albert *et al.* 2000, Bollobás and Riordan 2003).

In fact, geometric attacks are only a second-best disruption policy for a large class of power law distributions. The reason is the following.

Take two hubs of identical connectivity in a scale free network. A geometric attack selects network targets based solely on their connectivity. The two hubs are thus interchangeable targets for this policy. Although they have exactly the same number of direct connections, the two hubs may well differ by the structure of, say, their two-links away contacts, in terms of size or connectivity. More generally, the two hubs may differ by their larger circle of friends. Because of this, they need not contribute equally to the network functioning. A key player policy allows precisely to discriminate among these two hubs and singles out the one whose removal causes the highest harm to the network operations.

We identify exactly the class of scale free networks for which a geometric attack coupled with key player targeting constitutes a first-best disruption policy.

Capital letters correspond to random variables, and small letters to realizations.

Consider a random graph $\mathbf{G}$ on a population of n agents. This is a joint distribution on the set of $n(n-1)$ possible links. We assume that each link is assigned a positive probability. Let $\mathbf{g}$ be a realization of this process. This is just a collection of un-weighted links $g_{ij} \in \{0, 1\}$. Agent i has exactly $g_i = \sum_{j=1}^{n} g_{ij}$ direct contacts in $\mathbf{g}$, i's degree. We associate a probability distribution $P_{\mathbf{G}}$ to the random graph $\mathbf{G}$ given by $P_{\mathbf{G}}(z) = \sum_{i=1}^{n} \Pr_{\mathbf{G}}\{g_i = z\}/n$, the probability with which a randomly chosen node in the average network has degree z.

From now on, we take $\gamma = 0$. To spare on notations, let $\alpha = \beta = 1$.

Consider a realization $\mathbf{g}$ of $\mathbf{G}$. Using (2), the equilibrium aggregate outcome is:

$$x^*(\mathbf{g}) = n + \lambda \sum_i g_i + \lambda^2 \sum_i g_i^2 + o(\lambda^2).$$

Remove an agent k from the terrorist group. The new network is $\mathbf{g}^{-k}$, and the group outcome changes accordingly:

$$x^*(\mathbf{g}) - x^*(\mathbf{g}^{-k}) = 1 + 2\lambda g_k - \lambda^2 g_k(1-g_k) + 2\lambda^2 \sum_{j:g_{kj}=1} g_j + o(\lambda^2).$$

Removing an agent from the group has both a direct and an indirect effect. The direct effect results from the reduction in group size and network connectivity. This is $1 + 2\lambda g_k - \lambda g_k(1-g_k)$. The indirect effect accounts for the response of those that stay in the network to the agent removal. At a second order in λ, only the agents directly connected to k in $\mathbf{g}$ are concerned. This is:

$$\eta(\mathbf{g}, \lambda; k) = 2\lambda^2 \sum_{j:g_{kj}=1} g_j. \tag{5}$$

The value of $\eta(\mathbf{g}, \lambda; k)$ depends on the connectivity of k's direct contacts, and on λ. Given a network $\mathbf{g}$ and a target k, it is maximal when λ hits its upper bound $\lambda_{\max}$, equal to the inverse of the highest eigenvalue of $\mathbf{g}$.

Suppose that a geometric attack prescription identifies two candidates k and k' with, $g_k = g_{k'}$. The planner is indifferent between removing any of them, and picks the actual target randomly from this pool.

The direct effect of removing either k or k' is indeed identical. Yet, k and k' may have different friendship structures and, for instance, $\eta(\mathbf{g}, \lambda; k) > \eta(\mathbf{g}, \lambda; k')$. Then, removing either k or k' has a different sizeable indirect effect on group outcome reduction. Here, the optimal policy should single out k unambiguously as the actual target. The key player policy precisely operates this optimal selection as, by definition, the key player maximizes $x^*(\mathbf{g}) - x^*(\mathbf{g}^{-k})$.

Key players in large networks We bound the relative gains in outcome reduction when we amend the geometric attack prescription with a key player selection device.

For each realization $\mathbf{g}$ of $\mathbf{G}$, the geometric attack policy identifies a pool $\{k \mid g_k = a\}$ of target candidates with identical connectivity a. Under geometric attack alone, the target is selected (say, uniformly) among this pool. When geometric attack is coupled with a key player policy, the selected target k^* is such that:

$$k^* \in \arg\max\{\eta(\mathbf{g}, \lambda; k) \mid g_k = a\}.$$

Conditional upon removing hubs of connectivity a, the random graph $\mathbf{G}$ induces a distribution $\eta(\mathbf{G}, \lambda \mid a)$ over indirect gains. The gains from a key player selection within a pool of hubs with degree a is of the order of the standard deviation of $\eta(\mathbf{G}, \lambda \mid a)$. We compute the variance of this random variable.

Consider the degree distribution $P_{\mathbf{G}}(z)$ associated to the random graph $\mathbf{G}$. Fix some agent k with degree $g_k = a$, and let j such that $g_{kj} = 1$, j is a direct contact of k. First, note that the degree distribution of j is proportional to $zP_{\mathbf{G}}(z)$. Next, assume that the degrees of k's direct contacts are independent from each other. The variance of the joint connectivity $\sum_{j:g_{kj}=1} g_j$ of k's direct contacts is then proportional to:

$$a\left[\frac{\langle z^3\rangle}{\langle z\rangle} - \left(\frac{\langle z^2\rangle}{\langle z\rangle}\right)^2\right], \tag{6}$$

where $\langle\cdot\rangle$ is the standard notation for the expected value under $P_{\mathbf{G}}$.

Let m be the highest degree in the support of $P_{\mathbf{G}}(z)$, and d the average degree, $d = \sum_{z\in\mathbb{N}} P_{\mathbf{G}}(z)/n$. Then:

$$\max\{Var\left[\eta(\mathbf{G}, \lambda \mid a)\right] \mid \lambda, m\} = Var\left[\eta(\mathbf{G}, \lambda_{\max} \mid m)\right].$$

Using (5) and (6), we get:

$$Var\left[\eta(\mathbf{G}, \lambda_{\max} \mid m)\right] = 4\lambda_{\max}^4 m \left[\frac{\langle z^3\rangle}{\langle z\rangle} - \left(\frac{\langle z^2\rangle}{\langle z\rangle}\right)^2\right].$$

Let $\mathbf{G}$ be a scale free network with a power law degree distribution $P_{\mathbf{G}}(z) \propto z^{-\beta}$. Using the spectrum of a power law distribution (Chung *et al.* 2003), we get:

$$Var\left[\eta(\mathbf{G}, \lambda_{\max} \mid m)\right] \propto \begin{cases} d\left(\frac{m}{d}\right)^{-1}, \text{ if } 4 < \beta \\ d\left(\frac{m}{d}\right)^{3-\beta}, \text{ if } 2.5 < \beta < 4 \\ \frac{1}{d}\left(\frac{m}{d}\right)^{3\beta-7}, \text{ if } \beta < 2.5 \end{cases}.$$

Note that the variance exhibits a phase transition at $\beta = 2.5$, that mimics the phase transition for the highest eigenvalue of a scale free network in Chung *et al.* (2003).

Let now the maximum connectivity increase without bound, $m \uparrow +\infty$. Given the scale free nature of the connectivity distribution, the ratio of the maximal to the average connectivity m/d also increases without bound whenever $\beta > 2$. Then:

$$\lim_{m/d\uparrow+\infty} Var[\eta(\mathbf{G}, \lambda_{\max} \mid m)] = \begin{cases} \infty, \text{ if } 7/3 < \beta < 3 \\ 0, \text{ otherwise} \end{cases}.$$

The highest key player policy gains are proportional to the standard deviation of $\eta(\mathbf{G}, \lambda_{\max} \mid m)$. These gains are bounded away from zero as the population size increases if and only if β takes values between 2.33 and 3. This range of values is frequently encountered in real-life large-scale networks (Albert and Barábasi 2002, Jackson and Rogers 2005). Therefore, beyond small networks, a coupled geometric attack-key player policy is also optimal in most practical applications for large networks.

References

[1] Albert, Jeong and Barabási (2000): "Error and Attack Tolerance of Complex Networks," *Nature*, 406, 378-382.

[2] Albert and Barábasi (2002) "Statistical Mechanics of Complex Networks," *Review of Modern Physics* 74, 47-97.

[3] Ballester, Calvó-Armengol and Zenou (2005a) "Who's Who in Networks. Wanted: the Key Player," mimeo UAB.

[4] Ballester, Calvó-Armengol and Zenou (2005b) "Optimal Group Targets in Peer Networks," mimeo UAB.

[5] Baccara and Bar-Isaac (2005) "Crime, Punishment and Organizational Structure," mimeo NYU.

[6] Bollobás and Riordan (2003) "Robustness and Vulnerability of Scale-Free Random Graphs," *Internet Mathematics*, 1, 1-35.

[7] Corbo, Calvó-Armengol and Parkes (2005) "A Study of Nash Equilibrium via Bonacich Network Centrality in Contribution Games for Peer-to-Peer Networks," mimeo Harvard and UAB.

[8] Cvetković, Rowlinson and Simić (1997) *Eigenspaces of Graphs*. Cambridge: Cambridge University Press.

[9] Chung, Lu and Vu (2003) "The Spectra of Random Graphs with Given Expected Degrees," *Proceeding of National Academy of Sciences of the U.S.A.* 100, 11, 6313-6318.

[10] Debreu and Herstein (1953) "Nonnegative Square Matrices," *Econometrica*, 21, 597-607.

[11] Durlauf (2004) "Neighborhood Effects," in *Handbook of Regional and Urban Economics Vol. 4*, ed. by J.V. Henderson and J-F. Thisse. Amsterdam: Elsevier Science, 2173-2242.

[12] Jackson (2005) "A Survey of Models of Network Formation: Stability and Efficiency," in *Group Formation in Economics: Networks, Clubs and Coalitions*, ed. by G. Demange and M. Wooders. Cambridge: Cambridge University Press.

[13] Jackson and Rogers (2005) "Search in the Formation of Large Networks: How Random are Socially Generated Networks," mimeo California Institute of Technology.

[14] Krebs (2002) "Uncloaking Terrorist Networks," *First Monday* 7, 4.

[15] McCormick and Owen (2000) "Security and Coordination in a Clandestine Organization," *Mathematical and Computer Modelling* 31, 175-192.

[16] Rodríguez (2004) "La Red Terrorista del 11M," *Revista Española de Investigaciones Sociológicas* 107, 155-179.

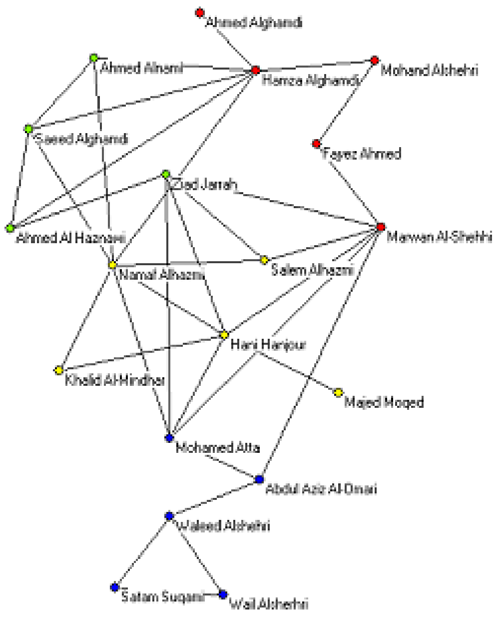

Figure 1 – The 11S on-board terrorist network from Krebs (2002).

Ranking	Inter-centrality	Closeness	Betweenness
1	Nawaf Alhazmi	Mohamed Atta	Nawaf Alhazmi
2	Hani Hanjour	Nawaf Alhazmi	Abdul Aziz Al-Omari
3	Mohamed Atta	Marwan Al-Shehhi	Mohamed Atta
4	Marwan Al-Shehhi	Hani Hanjour	Marwan Al-Shehhi
5	Ziad Jarrad	Ziad Jarrad	Waleed Alshehri
6	Salem Alhazmi	Hamza Alghamdi	Hamza Alghamdi
7	Hamza Alghamdi	Salem Alhazmi	Hani Hanjour
8	Saeed Alghamdi	Abdul Aziz Al-Omari	Ziad Jarrad
9	Ahmed Alnami	Saeed Alghamdi	Fayez Ahmed
10	Ahmed Al Haznawi	Ahmed Al Haznawi	Mohand Alshehri
11	Khalid Al-Mihdhar	Fayez Ahmed	Ahmed Al Haznawi
12	Abdul Aziz Al-Omari	Ahmed Alnami	Salem Alhazmi
13	Fayez Ahmed	Khalid Al-Mihdhar	Saeed Alghamdi
14	Majed Moqed	Mohand Alshehri	Majed Moqed
15	Mohand Alshehri	Majed Moqed	Khalid Al-Mihdhar
16	Ahmed Alghamdi	Waleed Alshehri	Ahmed Alnami
17	Waleed Alshehri	Ahmed Alghamdi	Ahmed Alghamdi
18	Wail Alshehri	Wail Alshehri	Wail Alshehri
19	Satam Suqami	Satam Suqami	Satam Suqami

Figure 2 – The 11S on-board terrorist ranking for the inter-centrality, closeness and betweenness network measures.

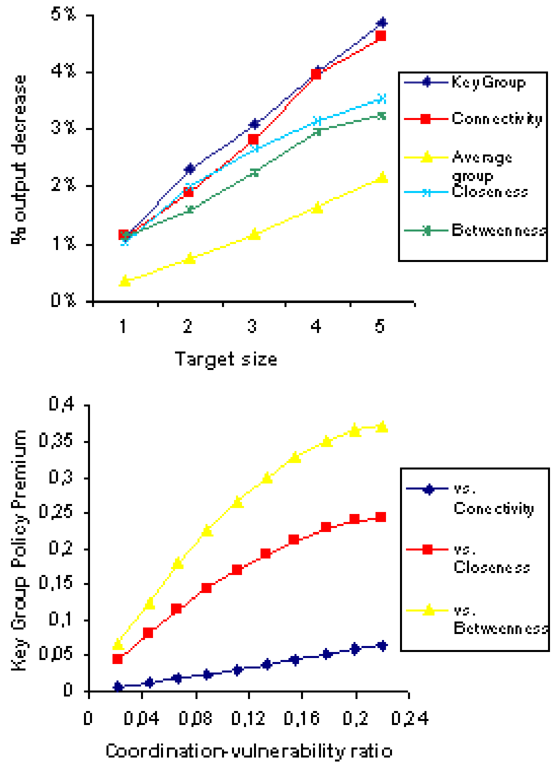

Figure 3 – The graph in the top shows the percentage decrease in organization output when we remove the highest ranked individuals according to inter-centrality, closeness, betweenness and connectivity. Average corresponds to a uniform random selection of the targets. We compute the inter-centrality for $\lambda^* = 0.22133$, just below its upper bound, equal to the inverse of the largest eigenvalue of g. In case of ties, we compute the average output decrease. The graph in the bottom shows the ratio (minus 1) of the percentage decrease in output for the key player inter-centrality prescription relative to the percentage decrease for alternative prescriptions (closeness, betweenness and connectivity), and for different values of λ^*. Note that inter-centrality increases with λ^*, whereas connectivity, closeness and betweenness are parameter-free. Also, when $\lambda^* \to 0$, $b_i(g, \lambda^*) = 1 + \lambda^* g_i + o(\lambda^*)$ and inter-centrality and connectivity prescriptions converge towards each other as $\lambda^* \to 0$.