NET Institute*

www.NETinst.org

Working Paper #08-08

September 2008

Who's who in networks. Wanted: the key group

Umed Temurshoev
University of Groningen

* The Networks, Electronic Commerce, and Telecommunications ("NET") Institute, http://www.NETinst.org, is a non-profit institution devoted to research on network industries, electronic commerce, telecommunications, the Internet, "virtual networks" comprised of computers that share the same technical standard or operating system, and on network issues in general.

Who's who in networks. Wanted: the key group

Umed Temurshoev*

September 30, 2008

Abstract

Ballester, Calvó-Armengol, and Zenou (2006, *Econometrica*, 74/5, pp. 1403-17) show that in a network game with local payoff complementarities, together with global uniform payoff substitutability and own concavity effects, the *intercentrality measure* identifies the *key player* - a player who, once removed, leads to the optimal change in overall activity. In this paper we search for the *key group* in such network games, whose members are, in general, *different* from the players with the highest individual intercentralities. Thus the quest for a single target is generalized to a group selection problem targeting an arbitrary number of players, where the key group is identified by a *group intercentrality measure*. We show that the members of a key group are rather nonredundant actors, i.e., they are largely heterogenous in their patterns of ties to the third parties.

Keywords: social networks, centrality measures, intercentrality measures, clusters, policies

JEL Classification Codes: A14, C72, L14

*Department of Economics and SOM Research School, University of Groningen, PO Box 800, 9700 AV Groningen, The Netherlands. Tel.: +31 50 363 8530; fax: +31 50 363 7337. E-mail: u.temurshoev@rug.nl. I gratefully acknowledge financial support from the NET Institute (www.netinst.org) and the Kauffman Foundation.

1 Introduction

One of the important topics in the sociology literature is the problem of identification of a *key actor* in social networks. Different measures of network centralities were proposed for this purpose, such as centralities of degree, closeness, betweenness, and information (see e.g., Sabidussi 1966, Freeman 1977, 1979, Stephenson and Zelen 1989). Other often used centralities include status measure or rank prestige index (Katz 1953), the eigenvector based measure (Bonacich 1972, 1991), and the related centrality in Bonacich (1987).[1] These ideas of finding the "most important" actors in a social network have been applied to a large number of substantive applications across different disciplines.

However, recently Everett and Borgatti (1999, 2005) proposed new measures of a network *group centralities* of degree, closeness, and betweenness to account for the fact that optimal selection of a set of k actors is quite different from selecting the k actors with the largest individual centralities. The inconsistency of the individual and group centralities is termed as the "ensemble issue" in Borgatti (2006), who interprets this by a redundancy principle inherent to a network (see also Burt 1992). That is, for example, two actors with the largest individual centralities cannot be optimal set (target) of two individuals if they "are redundant with respect to their liaising role – they are equivalent in that they connect the same third parties to each other" (p. 24), or these actors are structurally equivalent, meaning that they are connected to the same third parties. Borgatti (2006), in particular, shows that depending on the situation, one needs to use certain measures of centralities. For instance, he distinguishes between the "Key Player Problem/Negative" (KPP-Neg) and "Key Player Problem/Positive" (KPP-Pos). Given a social network, the aim of the KPP-Neg is finding a set of k nodes which, if removed, would maximally disrupt the network, while that of the KPP-Pos is finding a set of k nodes maximally connected to all other nodes.[2] Similar optimization problem, termed KPP-Com, is

[1] A thorough discussion of centrality and many more references can be found in Wasserman and Faust (1994, pp. 169-219).

[2] In practice KPP-Neg, for example, arises whenever it is needed to immunize or quarantine

defined in Puzis et al. (2007), which finds the group with the maximal potential of controlling traffic in communication networks. All in all, the existence of such a large number of "importance" indicators implies that there is *not* a systematic criterion to choose the "right" measure of network centrality for each particular situation.

From an economic point of view the important feature of network games is that actors' payoffs depend on each other through network embeddedness. A set of players each choose a level of some activity in a game, where there are negative global externalities (e.g., competition) and local positive externalities (e.g., learning, collaboration) that come through the network. This system has feedback effects, which are taken into account in the Nash equilibrium activity levels that are dependent on the underlying network topology. Recently, this network game was analyzed by Ballester et al. (2006), who show that its *individual* equilibrium levels of agents are proportional to their Katz-Bonacich centrality measures, hence provide to the status measure of Katz (1953) and the network centrality measure of Bonacich (1987) behavioral foundation "singling [them] out from the vast catalogue of network measures" (p. 1404).[3] Ballester et al. (2006) also propose a new measure of network centrality, named *intercentrality measure*, that finds a *key player* with the maximum influence on *overall* activity. Unlike the Katz-Bonacich (KB) centrality measure, intercentrality measure is derived from the planner's optimality concern, and not from strategic considerations of players, hence internalizes all the network payoff externalities of agents on each other. In particular, it is shown that the key player is not necessarily the player with the highest equilibrium outcome.

In this paper we extend the intercentrality measure from a single player target to a group index. The planner may target certain players by removing them from the network of local interactions, which causes a complete modification of the

a subset of population in order to optimally contain the epidemic, or in a military context, to neutralize a small number of actors in a criminal network in order to maximally disrupt its functioning. KPP-Pos arises, for example, in a public health context, when a health agency wishes to optimally spread information about health promoting practices and attitudes using a small subset of population, or in a military context, when one needs to select an optimal set of actors to quickly diffuse information/misinformation to all criminals.

[3]As will be shown later, the two centrality measures are affine transformations of each other.

distribution of individual outcomes. To characterize the optimal group target, we generalize the individual intercentrality measure to a *group intercentrality measure*. Hence, k (> 1) players with the highest group intercentrality of order k comprise the *key group*, whose elimination from the network results in the maximum impact on overall activity. And similar to the above discussions on the individual and group centralities in the sociology literature, it is true that the *key group* of k players is, in general, quite different from the k players with the highest individual intercentrality measures.[4] Like the individual intercentrality measure, the group intercentrality is obtained from the planner's optimality (collective) concerns. The removal of more than one players from a network has two effects. First, less players contribute to the aggregate equilibrium outcome (direct effect), and, second, the network geometry is modified, implying that the remaining players adopt different actions (indirect effect). These effects are fully taken into account in the group intercentrality measure, which considers not only the individual KB centralities of the group-members, but also their contributions to the KB centralities of players outside the group.

In Section 2 we characterize the optimal target selection task, both the key player and the key group problems. Section 3 applies the measures to an example of covert networks that characterize the organizational structure of the large terrorist organizations. It is shown that the k-top actors ($k > 1$) with the largest intercentrality measures do not necessarily comprise the key group of size k. Further, using the hierarchical agglomerative cluster analysis we show that the members of the key groups are less structurally equivalent (i.e., more heterogenous in their patterns of ties to the third parties) than other sets of players of the same size. Section 4 contains concluding remarks. All proofs are given in the Appendix.

[4]In a simple criminal network model, Ballester et al. (2004) show that the sequential optimal pair does not give the key group of two criminals.

2 Optimal target selection problem

In what follows, we briefly present Ballester et al. (2006, henceforth BCZ) model and their proposed intercentrality measure, and then extend the problem to an optimal search of groups consisting of an arbitrary number of players.

2.1 Key player search

Each player $i = 1, \dots, n$ selects an effort $x_i \geq 0$ and gets the bilinear payoff[5]

$$u_i(\boldsymbol{x}) = \alpha_i x_i + \frac{1}{2}\sigma_{ii} x_i^2 + \sum_{j \neq i} \sigma_{ij} x_i x_j, \tag{1}$$

which is strictly concave in own effort, $\partial^2 u_i / \partial x_i^2 = \sigma_{ii} < 0$, hence marginal utility of player i is decreasing in own action. We set $\alpha_i = \alpha > 0$ and $\sigma_{ii} = \sigma$ for all $i = 1, \dots, n$. The network payoff (relative) complementarities across all pairs of actors are reflected by the cross-derivatives $\partial^2 u_i / \partial x_i \partial x_j = \sigma_{ij}$ for $i \neq j$. That is, marginal utility of actor i is increasing in actor j's effort if $\sigma_{ij} > 0$, implying that i and j efforts are strategic complements. Reciprocally, $\sigma_{ij} < 0$ means that these two efforts are strategic substitutes from i's perspective.

Now the matrix of cross-effects $\boldsymbol{\Sigma} = [\sigma_{ij}]$ is decomposed into an idiosyncratic effect, a global substitutability effect, and a local complementarity effect as follows:[6]

$$\boldsymbol{\Sigma} = -\beta \boldsymbol{I} - \gamma \boldsymbol{U} + \lambda \boldsymbol{G}, \tag{2}$$

where $\beta > 0$, $\gamma \geq 0$, $\lambda > 0$, and $\boldsymbol{I}$ and $\boldsymbol{U}$ denote, respectively, the n-square identity matrix, and the n-square matrix of ones. The matrix $\boldsymbol{G} = [g_{ij}]$ with $g_{ij} = (\sigma_{ij} + \gamma)/\lambda$ for $i \neq j$ and $g_{ii} = 0$ is interpreted as an adjacency matrix of the network $\boldsymbol{g}$ of relative

[5]Adopting usual convention, matrices are given in bold, uppercase letters; vectors in bold, lowercase letters; and scalars in italic lowercase letters. Vectors are columns by definition, and transposition is indicated by a prime.

[6]See BCZ (2006) for details on derivation of (2). Within the model γ reflects the global substitutability of efforts across all pairs of players, $\lambda > 0$ is the highest possible relative complementarity for all pairs of players, and $\sigma = -\beta - \gamma$.

payoff complementarities across all pairs of players. By construction, $0 \leq g_{ij} \leq 1$. In this subsection, we focus on symmetric matrices such that $\sigma_{ij} = \sigma_{ji} \in \{\underline{\sigma}, \overline{\sigma}\}$ for all $i \neq j$ with $\underline{\sigma} \leq 0$, thus $\boldsymbol{G}$ is a symmetric $(0, 1)$ matrix, and $\boldsymbol{g}$ is an undirected and unweighted network. An idiosyncratic effect, $-\beta \boldsymbol{I}$, specific for each player, reflects (part of) the concavity in own efforts. The global interaction effect, $-\gamma \boldsymbol{U}$, gives a uniform substitutability in efforts across all pairs of players. The local interaction effect, $\lambda \boldsymbol{G}$, reflects a relative complementarity in efforts, which can be heterogenous across different pairs of actors. Further, the strength of local interactions relative to own concavity is denoted by $a = \lambda/\beta$.

Denote the largest eigenvalue of $\boldsymbol{G}$ by $\mu(\boldsymbol{G}) > 0$. Then if $a\mu(\boldsymbol{G}) < 1$, the matrix $\boldsymbol{B}(\boldsymbol{g}, a) = (\boldsymbol{I} - a\boldsymbol{G})^{-1} = \sum_{k=0}^{+\infty} a^k \boldsymbol{G}^k$ is well defined,[7] and its coefficients $b_{ij}(\boldsymbol{g}, a)$ count the number of paths in $\boldsymbol{g}$ starting at i and ending at j, where paths of length k are weighted by a^k. Hence, the parameter a in this interpretation is a decay factor that scales down the weight of longer paths. Let denote the summation vector of ones by $\boldsymbol{\imath}$. The vector of Katz-Bonacich (KB) centralities of parameter a in $\boldsymbol{g}$ is $\boldsymbol{b}(\boldsymbol{g}, a) = \boldsymbol{B}(\boldsymbol{g}, a)\boldsymbol{\imath}$, and its i-th component $b_i(\boldsymbol{g}, a) = \sum_{j=1}^{n} b_{ij}(\boldsymbol{g}, a)$ indicates the *total number of direct and indirect paths* in $\boldsymbol{g}$ that start from position i.[8] Note that, by definition, $b_{ii}(\boldsymbol{g}, a) \geq 1$, hence $b_i(\boldsymbol{g}, a) \geq 1$, with equality holding when player i is an isolate.[9]

From Theorem 1 in BCZ (2006) follows that for $a\mu(\boldsymbol{G}) < 1$, the unique interior Nash equilibrium of the network game is $\boldsymbol{x}^*(\boldsymbol{\Sigma}) = \frac{\alpha}{\beta + \gamma b(\boldsymbol{g}, a)} \boldsymbol{b}(\boldsymbol{g}, a)$, where $b(\boldsymbol{g}, a) = \sum_{i=1}^{n} b_i(\boldsymbol{g}, a)$.[10] This shows that the individual equilibrium outcomes are

[7]This follows from Theorem III* in Debreu and Herstein (1953, p. 601).

[8]In fact, Bonacich (1987) defines the network centrality measure by the vector $\boldsymbol{h}(\boldsymbol{g}, a, b) = b(\boldsymbol{I} - a\boldsymbol{G})^{-1}\boldsymbol{G}\boldsymbol{\imath}$, where the parameter b "affects only the length of the vector $[\boldsymbol{h}(\boldsymbol{g}, a, b)]$" (p. 1173). It is not difficult to show that $\boldsymbol{b}(\boldsymbol{g}, a) = \boldsymbol{\imath} + a\boldsymbol{h}(\boldsymbol{g}, a, 1)$. This measure is directly related to the Katz (1953) network status measure $\boldsymbol{k}(\boldsymbol{g}, a) = a(\boldsymbol{I} - \boldsymbol{G})^{-1}\boldsymbol{G}\boldsymbol{\imath}$, since $\boldsymbol{k}(\boldsymbol{g}, a) = a\boldsymbol{h}(\boldsymbol{g}, a, 1) = \boldsymbol{b}(\boldsymbol{g}, a) - \boldsymbol{\imath}$.

[9]Or when $a = 0$, which is not allowed in this network game.

[10]The condition $a\mu(\boldsymbol{G}) < 1$ for existence and uniqueness requires the payoff complementarity (size and pattern of positive synergies), $\lambda\mu(\boldsymbol{G})$, to be smaller than own concavity, β. This interpretation holds since λ measures the level of positive cross-effects, whereas $\mu(\boldsymbol{G})$ captures the population-wide pattern of these positive cross-effects. Or in terms of the decomposition given in (2), the equilibrium exists, is unique and interior, only when the positive feed-back loops $+\lambda\boldsymbol{G}$ are dampened by own concavity $-\beta\boldsymbol{I}$.

proportional to the KB centrality measures. Thus, the planner may only shift the distribution of individual outcomes by changing the exogenous payoff parameters. But what the planner can also do is to manipulate the network geometry, in which case the distribution of individual outcomes is completely modified. In this sense, the policy relevant issue studied in BCZ (2006) is removing one player, and identifying the network optimal target. Denote by $\boldsymbol{G}^{-i}$ (resp. $\boldsymbol{\Sigma}^{-i}$) the new adjacency matrix (resp. matrix of cross-effects) derived from $\boldsymbol{G}$ (resp. $\boldsymbol{\Sigma}$) by setting to zero all of its i-th row and column elements. The resulting network is $\boldsymbol{g}^{-i}$. Then the planner's problem is picking the appropriate player i from the population, such that its removal from the initial network $\boldsymbol{g}$ gives the highest possible reduction in aggregate equilibrium level. Formally, the problem is $\max\{x^*(\boldsymbol{\Sigma}) - x^*(\boldsymbol{\Sigma}^{-i})|i = 1, \ldots, n\}$, where $x^*(\boldsymbol{\Sigma}) = \boldsymbol{i}'\boldsymbol{x}^*(\boldsymbol{\Sigma})$, which is equivalent to

$$\min\{x^*(\boldsymbol{\Sigma}^{-i})|i = 1, \ldots, n\}. \tag{3}$$

The *key player* i^* is a solution to (3). The *intercentrality* of player i of parameter a in $\boldsymbol{g}$ is defined as[11]

$$c_i(\boldsymbol{g}, a) = b_i(\boldsymbol{g}, a) + \sum_{j \neq i} \left[b_j(\boldsymbol{g}, a) - b_j(\boldsymbol{g}^{-i}, a) \right] = \frac{b_i(\boldsymbol{g}, a)^2}{b_{ii}(\boldsymbol{g}, a)}. \tag{4}$$

While the KB centrality of actor i counts the number of direct and indirect paths in $\boldsymbol{g}$ stemming from i, from (4) it is easy to see that the "intercentrality counts the total number of such paths that hit i; it is the sum of i's Bonacich centrality and i's contribution to every other player's Bonacich centrality" (BCZ, 2006, p. 1411). Theorem 3 in BCZ (2006) shows that the key player i^* has the highest intercentrality, i.e., $c_{i^*}(\boldsymbol{g}, a) \geq c_i(\boldsymbol{g}, a)$ for all $i = 1, \ldots, n$. In their Example 1, the authors show that the most central player (according to the KB centrality measure) is not the key

[11]The last part of (4) follows from Lemma 1 in BCZ (2006). Straight algebra of the middle expression in (4) gives an alternative expression for the intercentrality measure as $c_i(\boldsymbol{g}, a) = b(\boldsymbol{g}, a) - b(\boldsymbol{g}^{-i}, a) + 1$. One is added in the last expression as all the i-th row (and column) elements of $\boldsymbol{B}(\boldsymbol{g}^{-i}, a)$ will be zero except its diagonal elements being $b_{ii}(\boldsymbol{g}^{-i}, a) = 1$, hence $b_i(\boldsymbol{g}^{-i}, a) = 1$.

player for relatively large a. This follows since then indirect effects matter and, as the intercentrality takes into account both a player's centrality and his contribution to the centrality of the others, key player with the highest joint direct and indirect effect on aggregate outcome might be well other than the most central player.

2.2 Key group search

In this section we wish to generalize the key player problem studied in BCZ (2006) to a group target selection problem. Thus, the planner's objective is now optimally reducing the aggregate equilibrium outcome by picking k appropriate players $i_1, i_2, \ldots, i_k$ $(i_s \neq i_r)$ from the population, where $1 \leq k \leq n$. Unlike the previous section, we analyze the more general setting, where the matrix of cross-effects, $\boldsymbol{\Sigma}$, can also be asymmetric. This expands the application of this analysis to the key group search in directed and/or valued graphs as well that are characterized by the network $\boldsymbol{g}$ having an asymmetric adjacency matrix $\boldsymbol{G}$, which is not necessarily a (0,1) matrix. Formally, the planner solves $\max\{x^*(\boldsymbol{\Sigma}) - x^*(\boldsymbol{\Sigma}^{-\{i_1,\ldots,i_k\}})|i_1,\ldots,i_k = 1,\ldots,n; i_r \neq i_s\}$, which is equivalent to

$$\min\{x^*(\boldsymbol{\Sigma}^{-\{i_1,\ldots,i_k\}})|i_1,\ldots,i_k = 1,\ldots,n; i_r \neq i_s\}, \tag{5}$$

where $\boldsymbol{\Sigma}^{-\{i_1,\ldots,i_k\}}$ is the new matrix of cross-effects obtained from $\boldsymbol{\Sigma}$ by setting to zero all its i_1-th, ..., i_k-th rows and columns elements. The resulting network and adjacency matrix are $\boldsymbol{g}^{-\{i_1,\ldots,i_k\}}$ and $\boldsymbol{G}^{-\{i_1,\ldots,i_k\}}$, respectively. Let $\{i_1^*, \ldots, i_k^*\}$ be a solution to (5), which we call the *key group of size k*.

Definition 1. *Consider a network $\boldsymbol{g}$ with adjacency matrix $\boldsymbol{G}$ and a scalar a such that $\boldsymbol{B}(\boldsymbol{g}, a) = [\boldsymbol{I} - a\boldsymbol{G}]^{-1}$ is well defined and nonnegative. The k-th order group intercentrality of players $i_1, \ldots, i_k$ $(i_r \neq i_s)$ of parameter a in $\boldsymbol{g}$ is*

$$c_{\{i_1,\ldots,i_k\}}(\boldsymbol{g}, a) = \boldsymbol{\imath}'\boldsymbol{BE}\left(\boldsymbol{E}'\boldsymbol{BE}\right)^{-1}\boldsymbol{E}'\boldsymbol{b},$$

where $\boldsymbol{E}$ *is the* $n \times k$ *matrix defined as* $\boldsymbol{E} = (\boldsymbol{e}_{i_1}, \ldots, \boldsymbol{e}_{i_k})$ *with* $\boldsymbol{e}_{i_r}$ *being the* i_r*-th column of the identity matrix, and* $1 \leq k \leq n$*.*

The interpretation of the group intercentrality is exactly the same as that of the individual intercentrality measure in (4), but now the target is a group of players.[12] The k-th order group intercentrality of players $i_1, \ldots, i_k$ can be rewritten as (this follows from the proof of Theorem 1)

$$c_{\{i_1,\ldots,i_k\}}(\boldsymbol{g}, a) = \sum_{r=i_1}^{i_k} b_r(\boldsymbol{g}, a) + \sum_{j \neq i_1,\ldots,i_k} \left[b_j(\boldsymbol{g}, a) - b_j(\boldsymbol{g}^{-\{i_1,\ldots,i_k\}}, a) \right],$$

which counts not only the total number of (weighted) paths in $\boldsymbol{g}$ that stem from positions $i_1, \ldots, i_k$ (i.e., the KB centralities of players $i_1, \ldots, i_k$), but also the total number of paths that hit these players. In other words, it is the sum of the KB centralities of all members of the group $\{i_1, \ldots, i_k\}$, and their contributions to every other player's KB centrality.[13]

The following important identity characterizes all the path changes in a network when a group of k nodes is removed.[14]

Lemma 1. *Let* $\boldsymbol{B} = [\boldsymbol{I} - a\boldsymbol{G}]^{-1}$ *be well defined and nonnegative. Let* $\boldsymbol{e}_{i_r}$ *be the* i_r*-th column of the identity matrix, and* $\boldsymbol{B}^{-\{i_1,\ldots,i_k\}} = [\boldsymbol{I} - a\boldsymbol{G}^{-\{i_1,\ldots,i_k\}}]^{-1}$*, where* $1 \leq k \leq n$*. Then the identity* $\boldsymbol{B} - \boldsymbol{B}^{-\{i_1,\ldots,i_k\}} = \boldsymbol{B}\boldsymbol{E}\left(\boldsymbol{E}'\boldsymbol{B}\boldsymbol{E}\right)^{-1}\boldsymbol{E}'\boldsymbol{B} - \boldsymbol{E}\boldsymbol{E}'$ *always holds, where* $\boldsymbol{E}$ *is the* $n \times k$ *matrix defined as* $\boldsymbol{E} = (\boldsymbol{e}_{i_1}, \ldots, \boldsymbol{e}_{i_k})$*.*

Using Lemma 1 we establish the following result that gives the solution to the problem (5) in terms of the k-th order group intercentrality measure.

---

[12]It is important to note that players $i_1, \ldots, i_k$ can be arbitrarily ordered in the matrix $\boldsymbol{E}$. Note that for simplicity, wherever needed, we suppress the expression $(\boldsymbol{g}, a)$.

[13]Given the fact that $b_{i_1}(\boldsymbol{g}^{-\{i_1,\ldots,i_k\}}, a) = \ldots = b_{i_k}(\boldsymbol{g}^{-\{i_1,\ldots,i_k\}}, a) = 1$, straight algebra gives an alternative expression for the k-th order group intercentrality measure as $c_{\{i_1,\ldots,i_k\}}(\boldsymbol{g}, a) = b(\boldsymbol{g}, a) - b(\boldsymbol{g}^{-\{i_1,\ldots,i_k\}}, a) + k > 0$. The positivity holds since the denser pattern of local complementarities always gives strictly higher aggregate activity level (Theorem 2, BCZ, 2006). This also implies that if one sets $a = 0$, then obviously $b(\boldsymbol{g}, 0) = b(\boldsymbol{g}^{-\{i_1,\ldots,i_k\}}, 0) = n$ implying that $c_{\{i_1,\ldots,i_k\}}(\boldsymbol{g}, 0) = k$ for all $k = 1, \ldots, n$.

[14]Lemma 1 in BCZ (2006, p.1411) is a particular case of our Lemma 1 with $k = 1$ and $\boldsymbol{G}$ being a symmetric adjacency matrix.

Theorem 1. *If* $a\mu(\boldsymbol{G}) < 1$*, the key group of size* k $\{i_1^*, \ldots, i_k^*\}$ *that solves the problem* $\min\{x^*(\boldsymbol{\Sigma}^{-\{i_1,\ldots,i_k\}})|i_1,\ldots,i_k = 1,\ldots,n; i_r \neq i_s\}$ *has the highest* k*-th order group intercentrality of parameter* a *in* $\boldsymbol{g}$*, where* $1 \leq k \leq n$*, i.e.,* $c_{\{i_1^*,\ldots,i_k^*\}}(\boldsymbol{g}, a) \geq c_{\{i_1,\ldots,i_k\}}(\boldsymbol{g}, a)$ *for all* $i_1, \ldots, i_k = 1, \ldots, n$ *with* $i_r \neq i_s$*.*

We shall note that Theorem 3 and Remark 5 in BCZ (2006) are particular cases of our Theorem 1 when $k = 1$ and the matrix of cross-effects $\boldsymbol{\Sigma}$ is, respectively, symmetric and asymmetric. This follows since with $k = 1$ the group intercentrality in Definition 1 boils down to

$$c_{\{i\}}(\boldsymbol{g}, a) = \boldsymbol{\imath}'\boldsymbol{B}\boldsymbol{e}_i \left(\boldsymbol{e}_i'\boldsymbol{B}\boldsymbol{e}_i\right)^{-1} \boldsymbol{e}_i'\boldsymbol{b} = \frac{\boldsymbol{\imath}'\boldsymbol{B}\boldsymbol{e}_i \cdot b_i(\boldsymbol{g}, a)}{b_{ii}(\boldsymbol{g}, a)},$$

which is the intercentrality of player i when $\boldsymbol{\Sigma}$ is not symmetric (Remark 5 in BCZ, 2006, p.1412).

When the matrix of cross-effects is symmetric, then $b_{kj}(\boldsymbol{g}, a) = b_{jk}(\boldsymbol{g}, a)$ for all k and all j, i.e., $\boldsymbol{B} = \boldsymbol{B}'$. Hence, we have $\boldsymbol{\imath}'\boldsymbol{B} = \boldsymbol{\imath}'\boldsymbol{B}' = (\boldsymbol{B}\boldsymbol{\imath})' = \boldsymbol{b}'$, implying that for a symmetric adjacency matrix $\boldsymbol{G}$ the group intercentrality of players $i_1, \ldots, i_k$ in Definition 1 can be rewritten as

$$c^{\text{sym}}_{\{i_1,\ldots,i_k\}}(\boldsymbol{g}, a) = \boldsymbol{b}'\boldsymbol{E} \left(\boldsymbol{E}'\boldsymbol{B}\boldsymbol{E}\right)^{-1} \boldsymbol{E}'\boldsymbol{b}. \tag{6}$$

Then it immediately follows that with $k = 1$ the above measure is simply the individual intercentrality measure given in (4).[15]

If we set $k = n$ and choose the ordering of players removed such that $\boldsymbol{E} = \boldsymbol{I}$, the group intercentrality in Definition 1 boils down to $c_{\{1,2,\ldots,n\}}(\boldsymbol{g}, a) = \boldsymbol{\imath}'\boldsymbol{B}\left(\boldsymbol{B}\right)^{-1}\boldsymbol{b} = \boldsymbol{\imath}'\boldsymbol{b} = b(\boldsymbol{g}, a)$, which is the sum of the KB centralities of all n players. Different ordering of the n players results in a different permutation matrix $\boldsymbol{E}$ of order n, but it will always give exactly the same outcome.[16] This is not surprising, since when we

[15] Using the analytical formula of the inverse matrix it can be shown, for instance, that the second-order group intercentrality of players i and j ($\neq i$) of parameter a in $\boldsymbol{g}$ with symmetric $\boldsymbol{G}$ is $c_{\{i,j\}}(\boldsymbol{g}, a) = (b_{jj}b_i^2 + b_{ii}b_j^2 - 2b_{ij}b_ib_j)/(b_{ii}b_{jj} - b_{ij}b_{ji})$.

[16] Similarly, using the alternative formulation of the group intercentrality of $c_{\{i_1,\ldots,i_k\}}(\boldsymbol{g}, a) =$

are interested in a group of all players, there are no outside actors left, hence there do not exist any other payoff externalities agents exert on non-members. Recall that namely those externalities are internalized by the intercentrality measure, which makes the last different from the KB centrality measure. But as with $k = n$ there are no other externalities to account for, the group intercentrality is nothing else than the sum of the KB centralities of all players.

The above observation also implies that if the network $\boldsymbol{g}$ consists of two separate (independent) subnetworks, then the group intercentrality of *all* players from one of the subnetworks is just equal to the sum of the KB centralities of players from that group. Mathematically, this can be proved as follows. Let the network $\boldsymbol{g}$ consists of two clusters (I and II), and no player in cluster I has a link to any of the players in cluster II, *and* vice versa, no player in cluster II has a link to any player of cluster I. That is, in terms of partitioned matrices we have

$$\boldsymbol{B} = \left[\boldsymbol{I} - a \begin{pmatrix} \boldsymbol{G}_k & \mathbf{O}_{kt} \\ \mathbf{O}_{tk} & \boldsymbol{G}_t \end{pmatrix}\right]^{-1} = \begin{bmatrix} \boldsymbol{B}_k & \mathbf{O}_{kt} \\ \mathbf{O}_{tk} & \boldsymbol{B}_t \end{bmatrix},$$

where, for example, $\boldsymbol{G}_k$ is the k-square adjacency matrix of all k players in cluster I, $\mathbf{O}_{kt}$ is the $k \times t$ null matrix, $\boldsymbol{B}_k = (\boldsymbol{I}_k - a\boldsymbol{G}_k)^{-1}$, and $k + t = n$. To find the group intercentrality of all players in cluster I, let take $\boldsymbol{E}' = [\boldsymbol{I}_k \ \mathbf{O}_{kt}]$. Note that in this case the vector of KB centralities is equal to

$$\boldsymbol{b} = \boldsymbol{B}\boldsymbol{\imath} = \begin{bmatrix} \boldsymbol{b}_k \\ \boldsymbol{b}_t \end{bmatrix},$$

where, for example, $\boldsymbol{b}_t = \boldsymbol{B}_t\boldsymbol{\imath}_t$ is the vector of KB centralities of all players in cluster II. Then using the group intercentrality formula in Definition 1 and the above partitioned matrix $\boldsymbol{B}$, one can by simple matrix multiplication easily verify that the group intercentrality of all k players from cluster I is $c_{\{i_1,\dots,i_k\}} = \boldsymbol{\imath}_k'\boldsymbol{b}_k$. Similarly, the

$b(\boldsymbol{g}, a) - b(\boldsymbol{g}^{-\{i_1,\dots,i_k\}}, a) + k$ (see fn. 13) it is easy to confirm that for $k = n$ we have $c_{\{i_1,\dots,i_n\}}(\boldsymbol{g}, a) = b(\boldsymbol{g}, a)$, since then $b(\boldsymbol{g}^{-\{1,\dots,n\}}, a) = \boldsymbol{\imath}'(\boldsymbol{I} - a\mathbf{O})^{-1}\boldsymbol{\imath} = n$.

t-th order intercentrality of all t players from cluster II is equal to $c_{\{i_{k+1},\ldots,i_n\}} = \boldsymbol{\imath}_t' \boldsymbol{b}_t$, where now we have to redefine $\boldsymbol{E}$ such that 1's appear in rows corresponding to the players of cluster II only. This result make sense, since when group outsiders do not have any link with insiders, then there also cannot be any kind of payoff externalities that group-members exert on outsiders. Of course, this result does not hold anymore, if the group consists of players from both subnetworks.

Corollary 1. *If* $a\mu(\boldsymbol{G}) < 1$*, the key group of size* k $\{i_1^*, \ldots, i_k^*\}$ *that solves the problem* $\max\{x^*(\boldsymbol{\Sigma}^{-\{i_1,\ldots,i_k\}})|i_1,\ldots,i_k = 1,\ldots,n; i_r \neq i_s\}$ *has the lowest* k*-th order group intercentrality of parameter* a *in* $\boldsymbol{g}$*, where* $1 \leq k \leq n$*, i.e.,* $c_{\{i_1^*,\ldots,i_k^*\}}(\boldsymbol{g}, a) \leq c_{\{i_1,\ldots,i_k\}}(\boldsymbol{g}, a)$ *for all* $i_1,\ldots,i_k = 1,\ldots,n$ *with* $i_r \neq i_s$*.*

3 Application to a convert network example

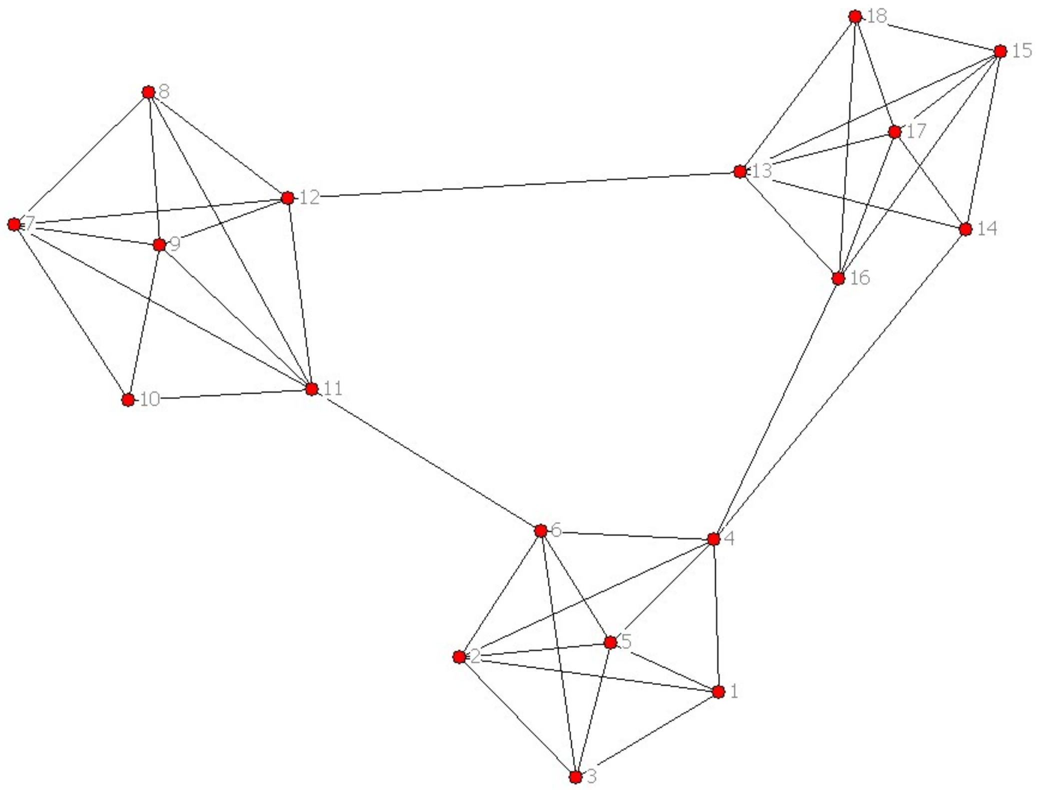
In this section in the example of a covert network we show that the key group of size k (> 1) does not necessarily include all the k players with the highest individual intercentrality measures, and the key groups' members are rather nonredundant actors. The *red team*, or the *covert network* consists of a set of small, but largely interconnected agents with little links between the sets, which mimics the organizational structure of the large terrorist organizations (see e.g., Krebs 2002). An example of (a small part of) such a network is given in Figure 1, which consists of three densely intraconnected groups of six players that are also weakly connected to each other.

Table 1 gives the KB centrality, individual and group intercentrality measures for $a = 0.1$.[17] Since the graph of the network is undirected, we use the symmetric intercentrality measure given in (6).[18] Although players 4, 11 and 13 have the highest number of direct links (i.e., six direct contacts each), player 13 is the most central player (it has the highest KB centrality), while player 4 is the key player (it has the

[17]The largest eigenvalue of the network in Figure 1 is equal to 4.894, hence the values of $a \in (0, 0.204)$ results in a well-defined and nonnegative matrix $\boldsymbol{B}$.

[18]Our MATLAB program to compute the group intercentrality measures is given in the Appendix.

Figure 1: The red team, or the covert network



highest intercentrality). This outcome was already shown in a different example in BCZ (2006, Table 1), which implies that the most central player is not necessarily an optimal target for the social planner who seeks the key player - a player with the highest joint direct and indirect impact on aggregate equilibrium outcome.

Turning our attention to the key group problem, Table 1 clearly demonstrates that the key group of size 2 consists of actors 4 and 11, and does not include one of the 2-top players (i.e, actor 13) in the ranking of the individual intercentrality measures. Note that both players 4 and 13 are also the most central players indicated by their respective KB centralities. In this example, the key group of size 3 includes all three players with the largest intercentrality (and KB centrality) measures. However, again in case of a group of size 4 not all the players with the highest intercentralities (and centralities) comprise the key group. Together with the players from the key group of size 3, the place of actor 16 is taken by players 2, 3, and 5, respectively, in three key groups of size 4 with equal intercentrality measures. Moreover, the set of

Table 1: Centrality and intercentrality measures

Rank	Player	b_i	Player	$c^{\text{sym}}_{\{i\}}$	Group of size 2	$c^{\text{sym}}_{\{i_1,i_2\}}$
1 (key)	13	2.161	4	4.282	{4,11}	8.307
2	4	2.156	13	4.269	{4,13}	8.297
3	11	2.130	11	4.152	{11,13}	8.284
4	16	2.009	16	3.748	{4,12}	7.954
	Group of size 3					$c^{\text{sym}}_{\{i_1,i_2,i_3\}}$
1 (key)	{4,11,13}					12.196
2	{2,11,13}, {5,11,13}					11.716
3	{4,11,15}, {4,11,17}					11.679
4	{4,7,13}, {4,9,13}					11.671
	Group of size 4					$c^{\text{sym}}_{\{i_1,i_2,i_3,i_4\}}$
1 (key)	{2,4,11,13}, {3,4,11,13}, {4,5,11,13}					14.685
2	{4,7,11,13}, {4,9,11,13}, {4,11,13,15}, {4,11,13,17}					14.575
3	{1,4,11,13}, {4,6,11,13}, {4,6,7,13}, {4,6,9,13}					14.320
4	{2,11,13,16}, {5,11,13,16}					14.311

Note that the intercentralities of all possible groups of size $k \in [1,4]$ were computed, which mathematically amount to the combinations of $n = 18$ players taken k at a time, $C^n_k = n!/(k!(n-k)!)$. Hence, all 18, 153, 816, and 3060 groups of size $k = 1, \ldots, 4$ were considered, respectively.

four players with the highest intercentrality measures appears only in the fifth rank with $c^{\text{sym}}_{\{4,11,13,16\}} = 14.254$ (which is not shown in Table 1).

The lack of coincidence between the composition of the key group and the ranking based on the key player problem is due to the *redundancy principle* inherent to the majority real life networks, a topic largely studied in the sociology literature. Arguing that the information and control benefits of a large and *diverse* network are more than those of a small and homogeneous network, Burt (1992, p.17), for example, states: "What matters is the number of nonredundant contacts. Contacts are redundant to the extent that they lead to the same people, and so provide the same information benefits." In general, redundancy of players in a network may be with respect to adjacency, distance, and bridging (see e.g., Borgatti 2006). One of the measures of redundancy is the notion of structural equivalence of nodes that reflects their similarity in terms of linkages to the third parties.

To compare our results of the key group problem with the notion of structural equivalence of players, we use the *hierarchical agglomerative cluster analysis* to identify groups of players that are similar in their patterns of ties to all other players (see e.g., Lattin et al. 2003, Chapter 8). Cluster analysis partitions actors to subgroups of perfectly or approximately structurally equivalent members. Each actor is initially considered as a singleton cluster, and then clusters are successively joined until all players merge into a single cluster. The process starts with constructing a so-called similarity matrix of players. We measure similarity of a pair of players by counting the proportion of their matches to all other alters.[19] The resulting similarity matrix is given in Table 2. The number 0.94 in the cell (2,1), for example, means that actors 1 and 2 have the same tie (present or absent) to other actors 94% of the time. Hence, the higher this score the more similar are the players in a particular pair.

Table 2: Similarity matrix of the red team in Figure 1

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1	1.00																	
2	0.94	1.00																
3	0.88	0.94	1.00															
4	0.75	0.81	0.88	1.00														
5	0.94	1.00	0.94	0.81	1.00													
6	0.94	0.88	0.81	0.69	0.88	1.00												
7	0.44	0.38	0.44	0.31	0.38	0.50	1.00											
8	0.50	0.44	0.50	0.38	0.44	0.56	0.94	1.00										
9	0.44	0.38	0.44	0.31	0.38	0.50	1.00	0.94	1.00									
10	0.56	0.50	0.56	0.44	0.50	0.63	0.88	0.94	0.88	1.00								
11	0.38	0.44	0.50	0.38	0.44	0.44	0.94	0.88	0.94	0.81	1.00							
12	0.44	0.38	0.44	0.31	0.38	0.50	0.88	0.94	0.88	0.88	0.81	1.00						
13	0.38	0.31	0.38	0.50	0.31	0.31	0.44	0.50	0.44	0.44	0.38	0.44	1.00					
14	0.63	0.56	0.50	0.50	0.56	0.56	0.44	0.50	0.44	0.56	0.38	0.56	0.75	1.00				
15	0.44	0.38	0.44	0.56	0.38	0.38	0.38	0.44	0.38	0.50	0.31	0.50	0.94	0.81	1.00			
16	0.56	0.50	0.44	0.44	0.50	0.50	0.38	0.44	0.38	0.50	0.31	0.50	0.81	0.94	0.88	1.00		
17	0.44	0.38	0.44	0.56	0.38	0.38	0.38	0.44	0.38	0.50	0.31	0.50	0.94	0.81	1.00	0.88	1.00	
18	0.50	0.44	0.50	0.50	0.44	0.44	0.44	0.50	0.44	0.56	0.38	0.56	0.88	0.88	0.94	0.94	0.94	1.00

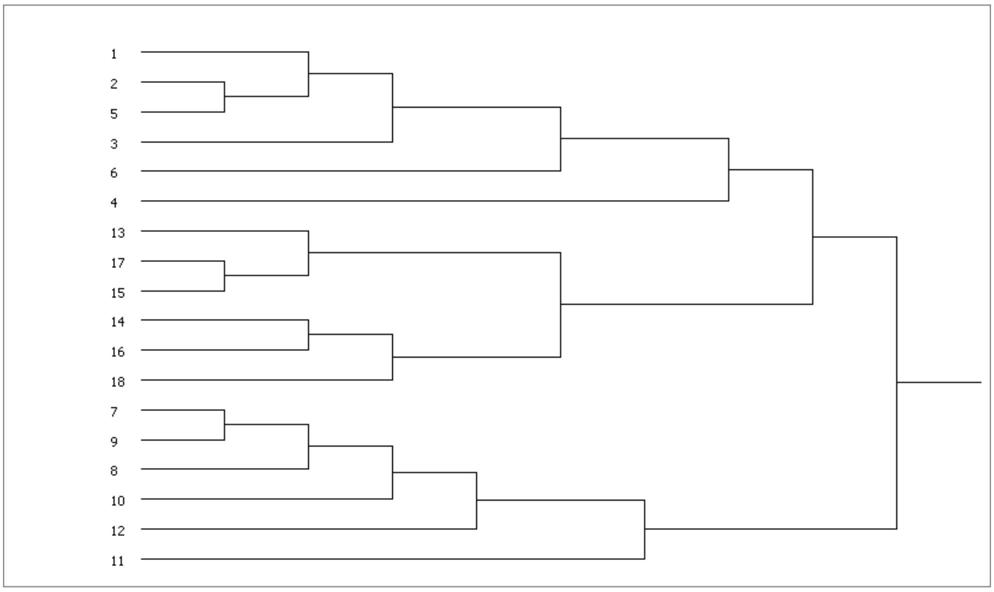
Next the cluster analysis starts from n clusters of size 1 and at each stage of the process finds the two "closest" (most homogenous) clusters and join them together. This process continues until one cluster of size n remains. This hierarchical sequence of merging clusters is visually depicted by a tree diagram, also called a *dendrogram*. We have used the *average link* criteria for forming clusters, which computes the similarity of the average scores in the newly formed cluster to all other clusters.[20]

[19]We also used a matrix of Euclidian distances to measure the "distance" or "dissimilarity" between the tie profiles of each pair of actors. The outcome of the cluster analysis totally coincides with that using proportion of matches in a similarity matrix.

[20]There are two other basic criteria for forming clusters: *single link* and *complete link*. The single (complete) method computes similarities on the base of the similarity of the member of the new cluster that is most (least) similar to each other clusters. While the single link approach is too myopic, the complete link method tends to give highly separated diagrams with tightly bound

The tree diagram of the cluster analysis of the above-illustrated covert network is given in Figure 2.

Figure 2: Hierarchical dendrogram of the covert network in Figure 1



As can be seen from Figure 2, nonoverlapping clusters are a product of the hierarchical agglomerative cluster analysis, i.e., the smaller clusters are subsumed within successively larger clusters at higher levels of agglomeration. It is clear that higher values of agglomeration indicate lower structural equivalence, less similarity, or greater within-cluster "distance". However, for our purposes we are not interested in choosing the level of agglomeration that provides "best" representation of the number of structurally equivalent positions in the network. Instead we aim at confirming or rejecting our conjecture that the key group includes less structurally equivalent (rather heterogenous) players.

The dendrogram in Figure 2 identifies two clusters at relatively high agglomeration level: $\{7, 8, 9, 10, 11, 12\}$ and the rest subsuming the second cluster. Note that the two actors with the highest individual intercentrality measures (i.e., players 4 and 13) are both members of the second cluster, hence are more homogenous than a

clusters. Hence, the average linkage is a sort of compromise between the single and the complete linkages. "Some author prefer [average approach] because it comes closest to fitting a tree that satisfies a least squares minimization criterion" (Lattin et al. 2003, p. 282).

pair of players from the two different clusters. As we expected the key group of size 2 consists of players 4 and 11 that are part of the two different clusters, thus being less redundant with respect to each other than the pair $\{4, 13\}$. Moreover, within these two clusters, respectively, actors 4 and 11 are less similar to all other members, which is shown by the fact that they join their clusters only at the highest level of agglomeration. Similarly, the relatively higher level of similarity produces three clusters: $A = \{1, 2, 3, 4, 5, 6\}$, $B = \{7, 8, 9, 10, 11, 12\}$, and $C = \{13, 14, 15, 16, 17, 18\}$. Within these clusters the (relatively) less structural equivalent actors (shown by the length of the line until it joins some cluster) are, respectively, players 4, 11, and 13.[21] And as Table 1 shows namely these actors comprise the key group of size 3. Still higher similarity level (lower agglomeration level) disaggregate cluster A into two sets of relatively homogenous players by identifying actor 4 being one cluster and the rest comprising the second, while remaining clusters B and C unchanged. Hence, our key group of size 4 besides actors 4, 11, 13 picks one player from the remaining part of cluster A, i.e., from $\{1, 2, 3, 5, 6\}$. This coincidence of the key group problem and cluster analysis can be also shown for larger size of the key group and lower level of agglomeration, which confirm our expectation that the key groups' members are rather nonredundant with respect to each other in terms of their linking patterns in a network.

However, we should note that the key group selection problem is *not* identical to a sequential key player problem, although in our case it happened that all the members of the key group of size k are included in the key group of size $k + 1$. This is a mere coincidence (hence we compared the results to those from the cluster analysis) and does not hold in general. Thus, in general, comparing the key group

[21] Note that according to the dendrogram in Figure 2 in cluster C the least structural equivalent actor is player 18, however, player 13 is taken as a key group member instead. This shows that using only cluster analysis for finding key groups would be misleading, since the right candidate for the key group should have not only diverse *direct* linking structure, but also diverse *indirect* impact on the rest of the system. Another reason why cluster analysis cannot exactly determine the key group members is that different criteria for forming clusters may very well give different outcomes. Despite this inconsistency, however, the key group problem and the cluster analysis are related in a sense that the key group members are less structurally equivalent, hence individually should be a part of different clusters.

outcomes to those from the hierarchical agglomerative cluster analysis makes little sense, since in the last case once actors are part of a group they will never leave it and only additional different actors join the group at the higher level of agglomeration.

4 Conclusion

In this paper we show that in a network game with local payoff complementarities together with global uniform payoff substitutability and own concavity effects, the *group intercentrality measure* identifies the *key group* - a set of player which, once removed, has the optimal impact on the overall activity level. This generalizes the key player problem in Ballester et al. (2006) from a search of a single player to a group selection problem targeting an arbitrary number of players. In the application of intercentrality measures to an example of a covert network it is shown that players with the highest individual intercentrality (and centrality) measures do not necessarily comprise the key group. This is because in a network players may be redundant with respect to each other, i.e., some actors might be quite similar to each other in terms of their linking structure. Therefore, it is expected the the key group members consists of relatively nonredundant players, which is confirmed by comparing the results of the key group identification problem and the hierarchical agglomerative cluster analysis.

The possible empirical applications of the group intercentrality measure, depending on the research question and the network content, will aim at finding a group of players of certain size with biggest (or smallest) influence over the overall activity level. Thus the measure is particularly useful for addressing such kind of issues in economics literature, since the notions of competition and complementarity due to the network embeddedness are explicitly taken into account. Just a few examples include the analysis of crime networks (Ballester et al. 2004, Calvó-Armengol and Zenou 2004), conformism and social norms (Bernheim 1994, Akerlof 1997), firms' collaboration networks (Goyal and Moraga-González 2001, Goyal and Joshi 2003), networks of interlocking directorates (Dooley 1969, Mizruchi 1996, Heemskerk and Schnyder 2008), and coauthor networks (Goyal et al. 2006).[22]

The final remark is that this analysis is not restricted to linear-quadratic utilities that incorporates linearly the players' actions externalities. In a general utility function that captures nonlinear externalities, a decomposition similar to (2) can be made. This in turn implies that the first-order approximation of the levels of players' actions will correspond to the Katz-Bonacich centrality measures. Also the entire analysis was made for a given network. Endogenizing the network decision is possible in a two-stage game, where in the first stage players decide whether to stay in the network or leave it for some outside option. In the second stage the network game is played by the remaining actors. This is particularly useful for the analysis of the effect of different policies in addressing the same issue. Such study was undertaken by Ballester et al. (2004), who show that the policy of increasing wages raises the effectiveness of another policy (the key player policy) in reducing crime.

[22]See also Temurshoev (2008), who applies the idea of the key group search to the framework of input-output analysis, where the key group of sectors are found on the base of a hypothetical extraction method.

Appendix

Proof of Lemma 1. From the monotonicity of the largest eigenvalue with the coefficients of the matrix it follows that $\mu(\boldsymbol{G}) \geq \mu(\boldsymbol{G}^{-\{i_1,\ldots,i_k\}})$.[23] Thus, when $\boldsymbol{B}$ is well defined and nonnegative, so is $\boldsymbol{B}^{-\{i_1,\ldots,i_k\}}$ for all $i_1,\ldots,i_k = 1,\ldots,n$ ($i_r \neq i_s$). For simplicity sake, define $\boldsymbol{Q}_k \equiv \boldsymbol{I} - \sum_{s=1}^{k} \boldsymbol{e}_{i_s}\boldsymbol{e}'_{i_s}$ and $\boldsymbol{A} \equiv a\boldsymbol{G}$. It is easy to verify that $\boldsymbol{A}^{-\{i_1,\ldots,i_k\}} = a\boldsymbol{G}^{-\{i_1,\ldots,i_k\}} = \boldsymbol{Q}_k\boldsymbol{A}\boldsymbol{Q}_k$, and $\boldsymbol{Q}_k\boldsymbol{Q}_k = \boldsymbol{Q}_k$, i.e., $\boldsymbol{Q}_k$ is an idempotent matrix. Further, we will use the well-known formula of the inverse of a sum of matrices (see e.g., Henderson and Searle 1981):

$$(\boldsymbol{X} - \boldsymbol{U}\boldsymbol{D}^{-1}\boldsymbol{Z})^{-1} = \boldsymbol{X}^{-1} + \boldsymbol{X}^{-1}\boldsymbol{U}(\boldsymbol{D} - \boldsymbol{Z}\boldsymbol{X}^{-1}\boldsymbol{U})^{-1}\boldsymbol{Z}\boldsymbol{X}^{-1}. \tag{A1}$$

Employing (A1) yields[24]

$$\boldsymbol{B}^{-\{i_1,\ldots,i_k\}} = (\boldsymbol{I} - \boldsymbol{Q}_k\boldsymbol{A}\boldsymbol{Q}_k)^{-1} = \boldsymbol{I} + \boldsymbol{Q}_k(\boldsymbol{A}^{-1} - \boldsymbol{Q}_k)^{-1}\boldsymbol{Q}_k. \tag{A2}$$

Equation (A1) also implies that $\boldsymbol{B} = (\boldsymbol{I} - \boldsymbol{A})^{-1} = \boldsymbol{I} + (\boldsymbol{A}^{-1} - \boldsymbol{I})^{-1}$, hence $(\boldsymbol{A}^{-1} - \boldsymbol{I})^{-1} = \boldsymbol{B} - \boldsymbol{I}$. Thus, again using (A1) with $\boldsymbol{X} = \boldsymbol{A}^{-1} - \boldsymbol{I}$ and $\boldsymbol{U} = -\sum_{s=1}^{k} \boldsymbol{e}_{i_s}\boldsymbol{e}'_{i_s} \equiv -\boldsymbol{\Sigma}$, the inverse in the rhs of (A2) can be rewritten as $\left[\boldsymbol{A}^{-1} - \boldsymbol{I} + \boldsymbol{\Sigma}\right]^{-1} = \boldsymbol{B} - \boldsymbol{I} - (\boldsymbol{B} - \boldsymbol{I})\boldsymbol{\Sigma}\left[\boldsymbol{I} + (\boldsymbol{B} - \boldsymbol{I})\boldsymbol{\Sigma}\right]^{-1}(\boldsymbol{B} - \boldsymbol{I})$. Plugging this back in (A2) yields

$$\boldsymbol{B}^{-\{i_1,\ldots,i_k\}} = \boldsymbol{I} + \boldsymbol{Q}_k(\boldsymbol{B} - \boldsymbol{I})\boldsymbol{Q}_k - \boldsymbol{Q}_k(\boldsymbol{B} - \boldsymbol{I})\boldsymbol{\Sigma}\left[\boldsymbol{I} + (\boldsymbol{B} - \boldsymbol{I})\boldsymbol{\Sigma}\right]^{-1}(\boldsymbol{B} - \boldsymbol{I})\boldsymbol{Q}_k. \tag{A3}$$

Next, without loss of generality, we partition the matrices $\boldsymbol{B}$ and $\boldsymbol{B}^{-\{i_1,\ldots,i_k\}}$ is such a way that the k removed players constitute their upper left submatrices. Then from the theory of partitioned matrices it follows that the matrix $[\boldsymbol{I} + (\boldsymbol{B} - \boldsymbol{I})\boldsymbol{\Sigma}]^{-1}$ in the rhs of (A3) is equal to

$$\left[\boldsymbol{I} + \begin{pmatrix} \boldsymbol{B}_{kk} - \boldsymbol{I}_k & \mathbf{O}_{kt} \\ \boldsymbol{B}_{tk} & \mathbf{O}_{tt} \end{pmatrix}\right]^{-1} = \begin{bmatrix} \boldsymbol{B}_{kk} & \mathbf{O}_{kt} \\ \boldsymbol{B}_{tk} & \boldsymbol{I}_t \end{bmatrix}^{-1} = \begin{bmatrix} \boldsymbol{B}_{kk}^{-1} & \mathbf{O}_{kt} \\ -\boldsymbol{B}_{tk}\boldsymbol{B}_{kk}^{-1} & \boldsymbol{I}_t \end{bmatrix}, \tag{A4}$$

where, for example, $\boldsymbol{I}_t$ is the t-dimensional identity matrix, $\mathbf{O}_{kt}$ is the $k \times t$ null matrix, and $k + t = n$.

The other terms in (A3) can be written as follows.

$$(\boldsymbol{B} - \boldsymbol{I})\boldsymbol{Q}_k = \begin{bmatrix} \boldsymbol{B}_{kk} - \boldsymbol{I}_k & \boldsymbol{B}_{kt} \\ \boldsymbol{B}_{tk} & \boldsymbol{B}_{tt} - \boldsymbol{I}_t \end{bmatrix} \begin{bmatrix} \mathbf{O}_{kk} & \mathbf{O}_{kt} \\ \mathbf{O}_{tk} & \boldsymbol{I}_t \end{bmatrix} = \begin{bmatrix} \mathbf{O}_{kk} & \boldsymbol{B}_{kt} \\ \mathbf{O}_{tk} & \boldsymbol{B}_{tt} - \boldsymbol{I}_t \end{bmatrix}, \tag{A5}$$

$$\boldsymbol{Q}_k(\boldsymbol{B} - \boldsymbol{I})\boldsymbol{Q}_k = \begin{bmatrix} \mathbf{O}_{kk} & \mathbf{O}_{kt} \\ \mathbf{O}_{tk} & \boldsymbol{I}_t \end{bmatrix} \begin{bmatrix} \mathbf{O}_{kk} & \boldsymbol{B}_{kt} \\ \mathbf{O}_{tk} & \boldsymbol{B}_{tt} - \boldsymbol{I}_t \end{bmatrix} = \begin{bmatrix} \mathbf{O}_{kk} & \mathbf{O}_{kt} \\ \mathbf{O}_{tk} & \boldsymbol{B}_{tt} - \boldsymbol{I}_t \end{bmatrix}, \tag{A6}$$

$$\boldsymbol{Q}_k(\boldsymbol{B} - \boldsymbol{I})\boldsymbol{\Sigma} = \begin{bmatrix} \mathbf{O}_{kk} & \mathbf{O}_{kt} \\ \boldsymbol{B}_{tk} & \boldsymbol{B}_{tt} - \boldsymbol{I}_t \end{bmatrix} \begin{bmatrix} \boldsymbol{I}_k & \mathbf{O}_{kt} \\ \mathbf{O}_{tk} & \mathbf{O}_{tt} \end{bmatrix} = \begin{bmatrix} \mathbf{O}_{kk} & \mathbf{O}_{kt} \\ \boldsymbol{B}_{tk} & \mathbf{O}_{tt} \end{bmatrix}. \tag{A7}$$

Now plugging (A4)-(A7) in (A3) after simple algebra gives

$$\boldsymbol{B}^{-\{i_1,\ldots,i_k\}} = \begin{bmatrix} \boldsymbol{I}_k & \mathbf{O}_{kt} \\ \mathbf{O}_{tk} & \boldsymbol{B}_{tt} - \boldsymbol{B}_{tk}\boldsymbol{B}_{kk}^{-1}\boldsymbol{B}_{kt} \end{bmatrix},$$

[23] This follows from Theorem I* in Debreu and Herstein (1953, p. 600).

[24] We should note that the inverse matrix $\boldsymbol{A}^{-1}$ might not exist, when, for example, some actor is an isolate. But it is used just as an analytical tool to derive our main equation (A3), which does not include this inverse matrix.

which implies that

$$\boldsymbol{B} - \boldsymbol{B}^{-\{i_1,\ldots,i_k\}} = \begin{bmatrix} \boldsymbol{B}_{kk} - \boldsymbol{I}_k & \boldsymbol{B}_{kt} \\ \boldsymbol{B}_{tk} & \boldsymbol{B}_{tk}\boldsymbol{B}_{kk}^{-1}\boldsymbol{B}_{kt} \end{bmatrix} = \begin{bmatrix} \boldsymbol{B}_{kk}\boldsymbol{B}_{kk}^{-1}\boldsymbol{B}_{kk} - \boldsymbol{I}_k & \boldsymbol{B}_{kk}\boldsymbol{B}_{kk}^{-1}\boldsymbol{B}_{kt} \\ \boldsymbol{B}_{tk}\boldsymbol{B}_{kk}^{-1}\boldsymbol{B}_{kk} & \boldsymbol{B}_{tk}\boldsymbol{B}_{kk}^{-1}\boldsymbol{B}_{kt} \end{bmatrix}. \tag{A8}$$

The final result in (A8) shows that the partition of having the k eliminated players in the upper left block diagonal matrix is quite arbitrary, hence the result holds for any non-ordered matrix $\boldsymbol{B}$. Moreover, the set of players $i_1, \ldots, i_k$ can be arbitrarily ordered in the matrix $\boldsymbol{B}_{kk}$ as well. So, (A8) proves that for all h and l we have

$$b_{hl} - b_{hl}^{-\{i_1,\ldots,i_k\}} = \boldsymbol{b}'_{h\bullet}\boldsymbol{B}_{kk}^{-1}\boldsymbol{b}_{\bullet l} - 1 \cdot I(h = l \in \{i_1, \ldots, i_k\}), \tag{A9}$$

where $\boldsymbol{b}'_{h\bullet} = (b_{hi_1}, \ldots, b_{hi_k})$, and $I(h = l \in \{i_1, \ldots, i_k\})$ is an indicator function that takes value one if $h = l \in \{i_1, \ldots, i_k\}$ and zero otherwise. The compact matrix form of (A9) is given in Lemma 1. □

Proof of Theorem 1. When $\alpha > 0$, $\frac{\partial x^*(\boldsymbol{\Sigma}^{-\{i_1,\ldots,i_k\}})}{\partial b(\boldsymbol{g}^{-\{i_1,\ldots,i_k\}},a)} = \frac{\alpha\beta}{(\beta+\gamma b(\boldsymbol{g}^{-\{i_1,\ldots,i_k\}},a))^2} > 0$, thus (5) is equivalent to $\min\{b(\boldsymbol{g}^{-\{i_1,\ldots,i_k\}}, a) | i_1, \ldots, i_k = 1, \ldots, n; i_r \neq i_s\}$, which have the same solution as $\max\{b(\boldsymbol{g}, a) - b(\boldsymbol{g}^{-\{i_1,\ldots,i_k\}}, a) | i_1, \ldots, i_k = 1, \ldots, n; i_r \neq i_s\}$. Using the definition of the KB centrality, Lemma 1, and the fact that $\boldsymbol{\imath}'\boldsymbol{E}\boldsymbol{E}'\boldsymbol{\imath} = k$, we have

$$\begin{aligned} b(\boldsymbol{g}, a) - b(\boldsymbol{g}^{-\{i_1,\ldots,i_k\}}, a) &= \boldsymbol{\imath}' \left[\boldsymbol{B} - \boldsymbol{B}^{-\{i_1,\ldots,i_k\}}\right] \boldsymbol{\imath} \\ &= \boldsymbol{\imath}' \left[\boldsymbol{B}\boldsymbol{E}\left(\boldsymbol{E}'\boldsymbol{B}\boldsymbol{E}\right)^{-1}\boldsymbol{E}'\boldsymbol{B} - \boldsymbol{E}\boldsymbol{E}'\right]\boldsymbol{\imath} = \boldsymbol{\imath}'\boldsymbol{B}\boldsymbol{E}\left(\boldsymbol{E}'\boldsymbol{B}\boldsymbol{E}\right)^{-1}\boldsymbol{E}'\boldsymbol{B}\boldsymbol{\imath} - k. \end{aligned} \tag{A10}$$

Players $i_1, \ldots, i_k$ that maximize (A10), also maximize the first expression in its right hand side, which is $c_{\{i_1,\ldots,i_k\}}(\boldsymbol{g}, a)$ in Definition 1. □

Computation of the group intercentrality measure

In computing the group intercentrality measures, we used the following MATLAB program.

```
G = load('D:\My Documents\MatlabKeyGroup\RedTeam.txt');
n = length(G);
mu = max(eig(G))   %the largest eigenvalue of G
I = eye(n);        %the identity matrix
amax = 1/mu
a = 0.1            %choose the value of parameter a
B = inv(I-a*G);
b = B*ones(n,1);   %Katz-Bonacich centralities

%%Key player problem (i.e., the key group of size 1)

c1 = inv(diag(diag(B)))*b.^2;
[C,IX] = sort(c1,'descend');
c1 = [C,IX]        %intercentrality with the identity of the player

%% Searching for the key group of size 2

t = nchoosek(n,2);  %number of combinations of n players taken 2 at a time
R = zeros(t,3);
m = 1;
for i=1:n
    for j=i+1:n
        E = [I(:,i),I(:,j)];
        c = b'*E*inv(E'*B*E)*E'*b;
        R(m,:) = [c,i,j];
        m = m+1;
    end
end
c2 = sortrows(R,1)    %2nd order group intercentrality with its members (in ascending order)

%% Searching for the key group of size 3

t = nchoosek(n,3);    %number of combinations of n players taken 3 at a time
R = zeros(t,4);
m = 1;
for i=1:n
    for j=i+1:n
        for k=j+1:n
            E = [I(:,i),I(:,j),I(:,k)];
            c = b'*E*inv(E'*B*E)*E'*b;
            R(m,:) = [c,i,j,k];
            m = m+1;
        end
    end
end
c3 = sortrows(R,1)   %3rd order group intercentrality with its members (in ascending order)

%% Searching for the key group of size 4

t = nchoosek(n,4);   %number of combinations of n players taken 4 at a time
R = zeros(t,5);
m = 1;
for i=1:n
    for j=i+1:n
        for k=j+1:n
            for h=k+1:n
                E = [I(:,i),I(:,j),I(:,k),I(:,h)];
                c = b'*E*inv(E'*B*E)*E'*b;
                R(m,:) = [c,i,j,k,h];
                m = m+1;
            end
        end
    end
end
c4 = sortrows(R,1)    %4th order group intercentrality with its members (in ascending order)
```

The program computes the group intercentrality measures for the symmetric networks. Asymmetric relations in a network can be easily introduced. For example, in computing the 3-rd order group intercentrality, one only needs to make a small change to `c = b'*E*inv(E'*B*E)*E'*b` (which is equation (6)), namely write instead `c = ones(n,1)*B*E*inv(E'*B*E)*E'*b`. For computing higher order group intercentrality measures one just needs to add extra loops similarly.

References

Akerlof, G. A.: 1997, 'Social Distance and Social Decisions'. *Econometrica* **65**, 1005–1028.

Ballester, C., A. Calvó-Armengol, and Y. Zenou: 2004, 'Who's who in crime networks. Wanted: the key player'. *Discussion Paper* **4421**. CERP.

Ballester, C., A. Calvó-Armengol, and Y. Zenou: 2006, 'Who's who in networks. Wanted: the key player'. *Econometrica* **74**, 1403–1417.

Bernheim, B. D.: 1994, 'A Theory of Conformity'. *Journal of Political Economy* **102**, 841–877.

Bonacich, P.: 1972, 'Factoring and weighting approaches to clique identification'. *Jounal of Mathematical Sociology* **2**, 113–120.

Bonacich, P.: 1987, 'Power and centrality: a family of measures'. *American Jounal of Sociology* **92**, 1170–1182.

Bonacich, P.: 1991, 'Simultaneous group and individual centralities'. *Social Networks* **13**, 155–168.

Borgatti, S. P.: 2006, 'Identifying sets of key players in a social network'. *Computational and Mathematical Organization Theory* **12**, 21–34.

Burt, R. S.: 1992, *Structural Holes. The Social Structure of Competition.* Cambridge: Hardvard University Press.

Calvó-Armengol, A. and Y. Zenou: 2004, 'Social networks and crime decisions: the role of social structure in facilitating delinquent behavior'. *International Economic Review* **45**, 939–958.

Debreu, G. and I. N. Herstein: 1953, 'Nonnegative square matrices'. *Econometrica* **21**, 597–607.

Dooley, P. C.: 1969, 'The interlocking directorate'. *American Economic Review* **59**, 314–323.

Everett, M. G. and S. P. Borgatti: 1999, 'The centrality of groups and classes'. *Journal of Mathematical Sociology* **23**, 181–201.

Everett, M. G. and S. P. Borgatti: 2005, 'Extending centrality'. In: S. J. Carrington, Peter J. and W. Stanley (eds.): *Models and Methods in Social Network Analysis.* Cambridge: Cambridge University Press, pp. 57–76.

Freeman, L. C.: 1977, 'A set of measures of centrality based on betweenness'. *Sociometry* **40**, 35–41.

Freeman, L. C.: 1979, 'Centrality in social networks: conceptual clarification'. *Social Networks* **1**, 215–239.

Goyal, S. and S. Joshi: 2003, 'Networks of collaboration in oligopoly'. *Games and Economic Behavior* **43**, 57–85.

Goyal, S. and J. L. Moraga-González: 2001, 'R&D networks'. *RAND Journal of Economics* **32**, 686–707.

Goyal, S., M. J. van der Leij, and J. L. Moraga-González: 2006, 'Economics: an emerging small world'. *Journal of Political Economy* **114**, 403–412.

Heemskerk, E. M. and G. Schnyder: 2008, 'Small states, international pressures, and interlocking directorates: the cases of Switzerland and the Netherlands'. *European Management Review* **5**, 41–54.

Henderson, H. V. and S. R. Searle: 1981, 'On deriving the inverse of a sum of matrices'. *SIAM Review* **23**, 53–60.

Katz, L.: 1953, 'A new status index derived from sociometric analysis'. *Psychometrica* **18**, 39–43.

Krebs, V. E.: 2002, 'Mapping networks of terrorist cells'. *Connections* **24**, 43–52.

Lattin, J., J. D. Carrol, and P. E. Green: 2003, *Analyzing Multivariate Data.* London: Thomson Learning, Inc.

Mizruchi, M. S.: 1996, 'What do interlocks do? An analysis, critique, and assessment of research on interlocking directorates'. *Annual Review of Sociology* **22**, 271–298.

Puzis, R., Y. Elovici, and S. Dolev: 2007, 'Finding the most prominent group in complex networks'. *AI Communications* **20**, 287–296.

Sabidussi, G.: 1966, 'The centrality index of a graph'. *Psychometrica* **31**, 581–603.

Stephenson, K. and M. Zelen: 1989, 'Rethinking centrality: methods and examples'. *Social Networks* **11**, 1–37.

Temurshoev, U.: 2008, 'Hypothetical extraction and fields of influence approaches: integration and policy implications'. *Mimeo.* University of Groningen.

Wasserman, S. and K. Faust: 1994, *Social Network Analysis: Methods and Applications.* Cambridge: Cambridge University Press.