University of Kent

School of Economics Discussion Papers

THE KEY CLASS IN NETWORKS

Nizar Allouch Jayeeta Bhattacharya

August 2021

KDPE 2110

THE KEY CLASS IN NETWORKS

AUGUST 2021

NIZAR ALLOUCH* JAYEETA BHATTACHARYA†

Abstract. This paper proposes new centrality measures to characterise the 'key class', when agents in a network are sorted into role-equivalent classes, such that its removal results in an optimal change in the network activity. The notion of role-equivalence is defined through the graph-theoretical concept of *equitable partition* of networks, which finds wide empirical and theoretical applicability. Players in the network engage in a non-cooperative game with local payoff complementarities. We establish a link between the generic network and its partitioned or quotient graph, and use it to relate the Nash equilibrium activity of classes with their position within the partitioned network. The result informs two class-based centrality measures that geometrically characterise the key class for an optimal reduction (or increase) in the aggregate and the per-capita network activity, respectively.

JEL classification: C72, D85.

Keywords: Social and economic networks, network games, equitable partition, centrality measures

*School of Economics, University of Kent, Canterbury, United Kingdom. Email: n.allouch@kent.ac.uk
†(Corresponding author) Department of Economics, University of Southampton, Southampton, United Kingdom. Email:j.bhattacharya@soton.ac.uk.

1. Introduction

A central feature of social networks is the prevalence of groups based on similarities in the structural positions and patterns of links or relationships among the individual network agents. Various types of equivalence classes in networks, also called sub-graphs, blocks or communities, have been proposed depending on how the notion of similar pattern of ties amongst the actors is visualised.[1] Grouping actors into classes that are equivalent or comparable relates not only to the structural position they occupy in the network, but also to the dual notion of their network role (see Wasserman and Faust (1994), Lerner (2005) or Baur, Brandes, Lerner, and Wagner (2009), among others, for a survey of equivalence classes and role analysis in networks).

In this paper, we consider networks where agents are sorted into role equivalent classes wherein class refers to an *equitable partition* of the network. As defined in Powers and Sulaiman (1982), equitable partition, which generalises the idea of structural equivalence of actors, requires that all players in a class have the same number of links amongst themselves, and with members of other classes. An equivalent definition is in terms of role assignment in networks[2]. As noted in Lerner (2005), equitable partition is associated with *exact* role assignment, such that individuals that have the same role are equivalent and they must have same number of each of the other roles in their neighbourhood, thus creating a society divided along roles. Furthermore, players in a class have the same value of Bonacich centrality, which is a measure of their network embeddedness. In the context of network games, the Bonacich-Nash linkage obtained in the seminal work of Ballester, Calvó-Armengol, and Zenou (2006) assumes importance: studying network games with pay-off externalities due to interaction among players, they prove that players' Bonacich centrality is directly proportional to their equilibrium strategic behaviour. Hence, an equitable partitioned network envisions a society that is divided along the lines of the network roles of individuals such that members of a class enjoy the same influence in the society and adopt similar actions in equilibrium.

[1]Equivalence classes is often visualised by 'coloration' of nodes in a graph; equivalent nodes have the same coloring.

[2]Role assignment refers to a surjective mapping from the set of nodes (individuals) in the network to a set of network roles.

In this setting of a role-based society, we address the following question: *what is the most important class in a network where players are sorted into role-equivalent classes, when the planner's objective is to minimize (or maximize) the aggregate or per capita network outcome*. Identifying the key class in role-based networks holds both practical and theoretical relevance. Among its empirical applications, equitable partition finds wide applicability for modelling community structure for studying information or epidemics diffusion in networks (see, for instance, Bonaccorsi, Ottaviano, Mugnolo, and Pellegrini (2015) or Ottaviano, De Pellegrini, Bonaccorsi, Mugnolo, and Van Mieghem (2019)). Indeed, a pervasive empirical observation of economic and social networks is the property of 'homophily', meaning that people tend to have ties with individuals similar to themselves, measured broadly by indicators like profession, religion, age, or gender. The resulting 'segregation' in networks is critical in determining agents' behaviour and features like information diffusion in the network, such that it is of interest to identify what is the most important class in such a segregated network.[3] In the context of epidemic outbreaks in networks divided into local communities, such that there is more interaction within a community than across communities, a natural question that arises is which community should be targeted in order to cause maximum disruption in the spread of the disease. This is particularly useful for very large and complex networks, as epidemic diffusion networks typically tend to be: characterising the importance of nodes on an aggregate role-based level, instead of an individual level, may have more appeal from a practically implementable policy perspective. Similarly, there has been a significant interest in studying criminal networks, including in identifying its key players. The presence of community structure in criminal networks is well accepted, as noted by Calderoni, Brunetto, and Piccardi (2017). Modelling criminal networks according to network roles, so as to identify which subgroup to target in order to maximally lower criminal activity, is of interest to the social planner. Moreover, while particular individuals in a criminal network may change, the roles in the criminal network is likely to be more stable, such that it is optimal to target role-based classes than individual players, for reducing overall criminal activity.

[3]See, for instance, Currarini, Jackson, and Pin (2009) or Golub and Jackson (2012) for the role of homophily and segregation in modelling friendship networks, and in examining speed of learning in networks, respectively.

Key class identification in equitable partitioned network can be significant for various theoretical applications as well. Examples include Allouch (2017), who consider segregated group membership-based interaction in studying welfare effects of income redistribution by private provision of public goods in social networks. It can also be useful in studying theoretical properties of networks, as in Rahmani, Ji, Mesbahi, and Egerstedt (2009) who use equitable partition to analyse controllability of multi-agent networks where a set of agents take on leadership roles. Also, it is to be noted that while a significant literature in group-level network analysis concerns with identifying the subgroup or equivalence structure through techniques such as block modeling or role-assignment, this paper takes the network structure as ex-ante given in order to identify the key class. This can, in a backward sense, contribute to the planner's problem of optimal network formation, similar to Belhaj, Bervoets, and Deroıan (2013)'s search for efficient networks, by suggesting which class to target so as to optimally alter the group-based structure for attaining desired network outcome.

This paper brings together the graph-theoretic notion of equitable partition as associated with positional/role analysis in networks, with its game-theoretic analysis to study the equilibrium behaviour of agents, in order to characterise the key class in role-equivalent networks for causing an optimal change in the network outcome. The network game is modelled using linear-quadratic utilities with linear bilateral externalities, as introduced in Ballester et al. (2006), such that there exists strategic complementarity of efforts between pairs of players.[4] We establish and exploit a relationship between the graph representing the overall network with the quotient graph of its equitable partitioning to show that the aggregate equilibrium activity of classes is related to their position within the network. This result is the class analogue of the key Bonacich-Nash linkage established in Ballester et al. (2006) and forms the basis for two class-based centrality measures proposed in this paper. The first is the *class-centrality* index to identify the most important class whose removal results in maximal disruption in the overall network outcome. At first glance, it may seem intuitive to think that this measure would select the class with the most

[4]Linear-quadratic utilities are used to model various social and economic phenomena. See e.g., Calvó-Armengol, Patacchini, and Zenou (2009) who study effect of peer influence on education outcomes in friendship network, Liu, Patacchini, Zenou, and Lee (2012) for criminal networks, or Goyal and Moraga-Gonzalez (2001) for R & D collaboration among Cournot competitors.

members as the key class. However, this is not always the case, since the class-centrality index reflects two kinds of effects that removing a class has on the aggregate network outcome. The first is the direct effect due to lesser contributing members in the resulting network after removing a class. But in addition, there is also the indirect effect due a change in the network architecture which alters the peer influences and their intensity, as the links get altered within and across classes. For instance, if the largest class has few direct links with other classes and most indirect links in the network do not pass through it, then it may not be the key class, especially if the indirect links in the network are strong (high attenuation factor). Moreover, the index is relevant if there are more than one class of the largest size. The second measure is the *per-capita class-centrality* which characterises the class whose removal reduces the per capita network activity by the most. This allows to target the class which may be smaller in size than the class with highest class-centrality index, but results in maximal per capita reduction of overall activity. Hence, this measure provides a cost-sensitive characterisation of removing the optimal class, which can be informative in presence of budget constraints for the policy planner.

The rest of the paper is organized as follows. Section 2 presents an overview of related literature on centrality measures in networks. Section 3 describes the network model and the corresponding equitable partition, while Section 4 carries out the Nash equilibrium analysis for class activity. Section 5 presents the class-based centrality measures, which are illustrated through examples in Section 6. Section 7 concludes the paper. All proofs are presented in Appendix.

2. Related Literature

The problem of developing measures of network centrality to determine which are the most influential, powerful, or important agents in a network has been a key area of focus in network analysis, owing to the ubiquity of social networks and their central role in influencing agents' behaviour.[5] Bonacich (1972)'s eigenvector centrality, which gives more importance to agents that have 'important' neighbours, is a key measure in this regard

[5]Wasserman and Faust (1994) or Jackson (2008) present a comprehensive overview of common centrality amd prestige measures.

and finds wide application in a range of fields.[6] Specifically, Bonacich centrality counts the total number of paths in the network originating from a node, discounted by their length. However, it does not consider the payoff interdependence among agents, which is crucial in analysing the aggregate network outcome in equilibrium. Ballester et al. (2006) propose the *intercentrality* measure which, unlike Bonacich centrality, is derived from the planner's optimality concerns. It characterises the key player whose removal results in maximum disruption to overall network activity in a network game with local complementarities in efforts among agents. The problem of identifying key player has further evolved to determining group-level centrality measures. Everett and Borgatti (1999, 2005) point out that group centrality measures are different from choosing a set of highest individual centralities, since they also depend upon members' connections and the group structure, and propose group version of common centrality measures like degree, closeness and betweenness. Borgatti (2006) distinguishes between 'key player problem-negative' and a 'key player problem-positive' for selecting a set of key nodes, arguing that situation decides which measure of centrality to use. Temurshoev (2008) generalises Ballester et al. (2006)'s intercentrality measure to search for the key collection of players of a particular size whose removal together can result in an optimal change in the overall network activity, without considering their underlying equivalence structure.

The above measures of centrality either focus on node or individual-level properties, or are derived from players' individual considerations without considering the pay-off externalities among agents. To the best of our knowledge, centrality measures that characterise the key class, removing which causes optimal change in network outcome, by considering the equivalence among players in role-based networks do not currently exist.

3. The network model

We consider a network $\mathbf{g}$ of n players. The associated $(0, 1)$-adjacency matrix is denoted by $\mathbf{G} = [g_{ij}]$, where g_{ij} represents unweighted and undirected connection between agents i and j; for $i \neq j$, it takes value of 1 if there is a link between the corresponding two nodes in the network, and 0 otherwise. Further, $g_{ii} = 0$, meaning there are no loops in $\mathbf{g}$, and multiple links between any two nodes are ruled out by construction. Note that

[6] One of its most well-known applications is the "PageRank" algorithm of Google search engine for ranking webpages.

$\mathbf{G}^k$ represents the number of paths of length k between any two nodes in the network; its elements are denoted by $g_{ij}^{[k]}$.

3.1. Equitable partition

Consider an equitable partition of the network $\mathbf{g}$ into $m \leq n$ classes $\{V_1, ..., V_m\}$: for every $i, j \in 1, ..., m$ there is a non-negative integer π_{ij} such that each node in V_i has exactly π_{ij} neighbours in V_j. An equitable partition results in a quotient graph $\boldsymbol{\pi}$ and the corresponding m-square quotient matrix is represented by $\mathbf{\Pi} = [\pi_{ij}]$. Note that unlike $\mathbf{G}$, the quotient matrix $\mathbf{\Pi}$ need not be symmetric. Denote the $(n \times m)$ indicator matrix by $\mathbf{X} = [X_{ij}]$, such that $X_{ij} = 1$ if vertex i is in the class V_j, and 0 otherwise. Let the number of members in a class V_i be denoted by r_i, such that, denoting the $(n \times 1)$ vector of ones by $\mathbf{1}_n$, the vector $\mathbf{r} = \mathbf{X}^T.\mathbf{1}_n$ lists the number of members in each class. The following property holds by definition:

$$\mathbf{GX} = \mathbf{X\Pi} \tag{3.1}$$

Also, the adjacency matrix $\mathbf{G}$ and the quotient matrix of its equitable partition, $\mathbf{\Pi}$, have the same spectral radius.[7] That is, if $\rho(A)$ denotes the largest absolute value of the eigenvalues of square matrix $\mathbf{A}$, then

$$\rho(G) = \rho(\Pi) = \rho.$$

Finally, denote $\mathbf{\Pi}^k = \left[\pi_{ij}^{[k]}\right]$ where $\pi_{ij}^{[k]}$ denotes the total paths of length k for any node in class V_i with its neighbours in class V_j.

3.2. Bonacich centrality

Here, we provide the definition of Bonacich centrality measure which is relevant to our purpose. Bonacich centrality counts the total number of paths starting from node i in the network $\mathbf{g}$, weighted down by their length. The vector of Bonacich centralities, with a decay parameter a, in $\mathbf{g}$ is given by:

$$\mathbf{b}(\mathbf{g}, a) = [\mathbf{I}_n - a\mathbf{G}]^{-1}.\mathbf{1}_n = \sum_{k=0}^{+\infty} a^k \mathbf{G}^k.\mathbf{1}_n \tag{3.2}$$

[7]See Van Mieghem (2010), page 62, **art.** 62.

where $\mathbf{I}_n$ denotes an n-square identity matrix. Note that the above expression is well-defined for small values of a, specifically, if a is less than inverse of the largest absolute eigenvalue of $\mathbf{G}$.

4. Network game: Nash equilibrium class activity

We consider the network game with local payoff complementarities as in Belhaj et al. (2013), which is simplified version of the linear-quadratic utility function of Ballester et al. (2006). Players $\{1, ..., n\}$ in a network engage in a non-cooperative game, where the strategy of each player is to decide the extent of efforts they exert. The utility of player i is given by

$$u_i(x_1, ..., x_n) = x_i - \frac{1}{2}x_i^2 + \lambda \sum_{j=1}^{n} g_{ij} x_i x_j$$

where $x_i \geq 0$ denotes the effort of player i, and $\lambda > 0$ measures the intensity of linear interactions among pairs of players. Hence, the utility function consists of two components: an idiosyncratic component made up of own efforts and an interaction component reflecting strategic complementarities among connected players. Further, the linear-quadratic form implies that utility is strictly concave in one's own efforts.

In this setting of network game with linear-quadratic utility and payoff complementarities, Ballester et al. (2006) establish the proportionality between players' Nash equilibrium outcome and their Bonacich centrality. This is a key result which establishes the intuitive link between players' equilibrium behaviour with their positions within the network. Indeed, it can be shown that player i's unique Nash equilibrium outcome for the game described above, x_i^*, equals their Bonacich centrality $b_i(\mathbf{g}, \lambda)$.

We are interested in equilibrium analysis for determining class activity when the network of relative payoff complementarities has a partition structure as conceptualised by the role-based notion of equitable partition. We first present the following Lemma. Let $\mathbf{A}^T$ denote the transpose of matrix $\mathbf{A}$. Note that in what follows, in the Lemmas and Definitions that pertain to any general network structure and its equitable partition, we use the symbol 'a' to denote the attenuation factor, which plays the role of 'λ' for the results obtained in the context of the network game explained above.

Lemma 1. *Let $0 < a \leq 1/\rho$ such that $[\boldsymbol{I}_n - a\boldsymbol{G}]^{-1}$ and $[\boldsymbol{I}_m - a\boldsymbol{\Pi}]^{-1}$ are well-defined and nonnegative. Then, $[\boldsymbol{I}_n - a\boldsymbol{G}]^{-1}\boldsymbol{X} = \boldsymbol{X}[\boldsymbol{I}_m - a\boldsymbol{\Pi}]^{-1}$.*

Lemma 1 relates the overall network structure with its equitable partition. It enables applying the Nash-Bonacich linkage result of Ballester et al. (2006) to our network game and to express the equilibrium activity of classes in relation to the network's partition structure. For this purpose, define the following matrix,

$$\mathbf{N}(\boldsymbol{\pi}, \lambda) = \left[\mathbf{I}_m - \lambda\mathbf{\Pi}^T\right]^{-1} = \sum_{p=0}^{\infty} \lambda^p (\mathbf{\Pi}^p)^T$$

which is well-defined and nonnegative for $\lambda \leq 1/\rho$. Its elements $N_{ij}(\boldsymbol{\pi}, \lambda) = \sum_{p=0}^{\infty} \lambda^p \pi_{ji}^{[p]}$ count the total number of paths of length p for any node in class V_j with the members in V_i, weighted down by λ^p. Let $\mathbf{y}^*(\boldsymbol{\pi}) = [y_i^*]$ denote the outcome vector for classes at equilibrium, where y_i^* is the sum of equilibrium outcomes of all players of class V_i, $i = 1, \ldots, m$. Also, for a vector $\mathbf{z} \in \mathbb{R}^p$, we denote the sum of its entries as $z = z_1 + \cdots + z_p$.

Theorem 1. *The matrix $\boldsymbol{N}(\boldsymbol{\pi}, \lambda) = \left[\boldsymbol{I}_m - \lambda\boldsymbol{\Pi}^T\right]^{-1}$ is well-defined and nonnegative when $\lambda \leq 1/\rho$. Then, the unique and interior Nash equilibrium class activity for the network game characterised by u_i, $i = 1, \ldots, n$, played over the quotient graph $\boldsymbol{\pi}$, is given by*

$$\boldsymbol{y}^*(\boldsymbol{\pi}) = \boldsymbol{N}(\boldsymbol{\pi}, \lambda).\boldsymbol{r} \equiv \boldsymbol{t}(\boldsymbol{\pi}, \lambda). \tag{4.1}$$

In the above, $\mathbf{N}(\boldsymbol{\pi}, \lambda).\mathbf{r}$ is the vector of sum of Bonacich centralities of members in a class. That the contribution of a class to the overall network activity is proportional to the sum of its members' Bonacich centralities is expected. But more importantly, equation (4.1) links the equilibrium activity of a class with its position in the network, as represented by the equitable partition network structure through the matrix $\mathbf{N}(\boldsymbol{\pi}, \lambda)$. Hence, it can be considered as the class analogue of the Bonacich-Nash linkage of Ballester et al. (2006).

5. The key class: two measures

The above analysis shows that the class outcome at equilibrium is related to its position within the network when there exists payoff externalities among agents. Removing a class alters the network structure of bilinear influences, in addition to reducing the number of agents who contribute to the overall network activity, thus altering the equilibrium network outcome. In this section, we propose two geometric measures to characterise equilibrium outcome, in aggregate and in per-capita terms, upon removing classes. This informs simple criteria for targeting the optimal class if the planner wants to minimise (or maximise) the aggregate or the per-capita network activity, respectively.

Consider the game of Section 4 being played over the network $\mathbf{g}$ with symmetric square adjacency matrix $\mathbf{G} = [g_{ij}]$, where $g_{ij} \in \{0, 1\}$ for $i \neq j$ and g_{ii} is set to 0; its corresponding quotient network is $\boldsymbol{\pi}$ with quotient matrix $\mathbf{\Pi} = [\pi_{ij}]$. Let a class j be removed from the network. The corresponding partition matrix is denoted by $\mathbf{\Pi}^{-j}$, by setting the jth row and jth column of $\mathbf{\Pi}$ to zero. Also, $\mathbf{r}^{-j}$ is the class size vector associated with removing class j by setting j-th coordinate of $\mathbf{r}$ to 0. The overall network activity is the sum of the activities due to all remaining classes $y^*(\boldsymbol{\pi}^{-j}) = \sum_{i=1,\ i\neq j}^{m} y_i^*(\boldsymbol{\pi}^{-j})$. The derivation of the class-based centrality measure makes use of the following Lemma, which characterises changes in the structure of links among players of remaining class when a class is removed.

Lemma 2. *Let* $0 \leq a \leq 1/\rho$ *such that* $\boldsymbol{N}(\boldsymbol{\pi}, a) = \left[\boldsymbol{I}_m - a\mathbf{\Pi}^T\right]^{-1}$ *is well-defined and non-negative. Let* $\boldsymbol{N}(\boldsymbol{\pi}^{-j}, a) = [\boldsymbol{I}_m - a(\mathbf{\Pi}^{-j})^T]^{-1}$*. Then:*

$$N_{ik}(\boldsymbol{\pi}, a) - N_{ik}(\boldsymbol{\pi}^{-j}, a) = \frac{N_{ij}(\boldsymbol{\pi}, a).N_{jk}(\boldsymbol{\pi}, a)}{N_{jj}(\boldsymbol{\pi}, a)}. \tag{5.1}$$

5.1. Class-centrality

The class-centrality index is concerned with identifying the class removing which results in an optimal reduction in the aggregate network outcome.[8] Formally, the planner's objective is to:

$$\min\left\{y^*(\boldsymbol{\pi}^{-j})\right\} \quad \text{or} \quad \max\left\{y^*(\boldsymbol{\pi}) - y^*(\boldsymbol{\pi}^{-j})\right\}, \quad j = 1, \ldots, m. \tag{5.2}$$

Definition 1. Let there be a quotient network $\boldsymbol{\pi}$ that divides the network $\mathbf{g}$ into m classes, with the associated partition matrix $\mathbf{\Pi}$ and a decay factor $a > 0$ such that $[\mathbf{I}_m - a\mathbf{\Pi}]^{-1}$ is well-defined and non-negative. The class-centrality measure of class j is given by:

$$e_j(\boldsymbol{\pi}, a) = \frac{t_j(\boldsymbol{\pi}, a).s_j(\boldsymbol{\pi}, a)}{N_{jj}(\boldsymbol{\pi}, a)}, \tag{5.3}$$

[8] The planner may, equivalently, wish to increase overall network activity.

where $\mathbf{N}(\boldsymbol{\pi}, a) = \left[\mathbf{I}_m - a\boldsymbol{\Pi}^T\right]^{-1}$, $\mathbf{s}(\boldsymbol{\pi}, a) = \mathbf{1}_m^T.\mathbf{N}(\boldsymbol{\pi}, a)$, and $\mathbf{t}(\boldsymbol{\pi}, a) = \mathbf{N}(\boldsymbol{\pi}, a).\mathbf{r}$.

The above index informs a simple criterion to characterise the key class j^* to optimally reduce (or increase) network outcome, as presented in the following Theorem.

Theorem 2. *If $\lambda \leq 1/\rho$, the class that solves $\max\{y^*(\boldsymbol{\pi}) - y^*(\boldsymbol{\pi}^{-j})\}$ is the j^* for which the class-centrality measure is the highest, that is, $e_{j^*}(\boldsymbol{\pi}, \lambda) \geq e_j(\boldsymbol{\pi}, \lambda)$ for all $j = 1, ..., m$.*

Note that removing a class has a direct and an indirect effect on network activity. Direct effect is by virtue of a reduction in the number of players who contribute to network activity. Indirect effect is due to the fact that removing a class alters the network structure such that the remaining classes adopt different equilibrium actions, thereby again altering the aggregate (or per-capita) network activity. Hence, the class with the most players need not be the key class for reducing the aggregate network activity.

5.2. Per-capita class-centrality

Other than bilinear influences, the size of a class plays a significant, and sometimes, indeed, the deciding role in determining the key class using the class-centrality index. Targeting large classes can prove to be restrictive in presence of budget constraints for the planner. The per-capita class centrality index addresses this limitation by providing a geometric measure for the class removing which results in maximum per-capita reduction (or increase) in network activity. The planner's objective, therefore, is:

$$\min\left\{\frac{y^*(\boldsymbol{\pi}^{-j})}{n - r_j}\right\} \quad \text{or} \quad \max\left\{\frac{y^*(\boldsymbol{\pi})}{n} - \frac{y^*(\boldsymbol{\pi}^{-j})}{n - r_j}\right\}, \quad j = 1, \ldots, m. \tag{5.4}$$

Definition 2. For the quotient network and decay factor a as specified in Definition 1, the per-capita class-centrality measure of class j is given by:

$$h_j(\boldsymbol{\pi}, a) = \frac{n.\left(t_j(\boldsymbol{\pi}, a)/N_{jj}(\boldsymbol{\pi}, a)\right).s_j(\boldsymbol{\pi}, a) - r_j.t(\boldsymbol{\pi}, a)}{n(n - r_j)}, \tag{5.5}$$

where $\mathbf{N}(\boldsymbol{\pi}, a) = \left[\mathbf{I}_m - a\boldsymbol{\Pi}^T\right]^{-1}$, $\mathbf{s}(\boldsymbol{\pi}, a) = \mathbf{1}_m^T.\mathbf{N}(\boldsymbol{\pi}, a)$, $\mathbf{t}(\boldsymbol{\pi}, a) = \mathbf{N}(\boldsymbol{\pi}, a).\mathbf{r}$, and $t(\boldsymbol{\pi}, a)$ denotes the sum of the coordinates of $\mathbf{t}(\boldsymbol{\pi}, a)$.

Per-capita class-centrality $h_j(\boldsymbol{\pi}, a)$ characterises the per-capita network activity upon removing class j, in terms of the position that its players occupy within the partitioned network. This informs a simple criterion for selecting the key class for optimally lowering per-capita network activity from the planner's perspective, as given by the following Theorem.

Theorem 3. *If* $\lambda \leq 1/\rho$, *the class that solves* $\max \left\{ \frac{y^*(\boldsymbol{\pi})}{n} - \frac{y^*(\boldsymbol{\pi}^{-j})}{n-r_j} \right\}$ *is the* j^* *for which the per-capita class-centrality measure is the highest, that is,* $h_{j^*}(\boldsymbol{\pi}, \lambda) \geq h_j(\boldsymbol{\pi}, \lambda)$ *for all* $j = 1, ..., m$.

Similar to class-centrality, the per-capita measure also reflects the dual effects of lesser contributing members as well changes in the network structure of peer-effects, in determining the network activity of the resultant network. Note that the key class in the above Theorem refers to one whose removal reduces per capita network activity by the most. This is not the same as the class which contributes the most in per capita terms. The reason is as previously explained: when a class j is removed, the equilibrium activity of others without class j will no longer be the same if j had been there, due to alterations in the network structure of bilateral influences. The same argument applies to the class-centrality measure as well, such that removing the corresponding key class optimally reduces the aggregate network activity but does not have the interpretation of being the class that contributes the most in terms of its share in overall network activity.

Remark 1. Both the class-based centrality measures are generic indices applicable to any network structure. At the extreme case when a network does not have an equitable partition structure, 'class' simply refers to individual players. In that case, it is straightforward to notice for the class-centrality index that the planner's problem (5.2) translates to the key player problem of Ballester et al. (2006); the class-centrality, then, is the same as their intercentrality measure.

Remark 2. The notion of equitable partition for specifying classes is also applicable to nested split graph structures in networks which postulates neighbourhoods of agents of lower degree to be contained in the neighbourhoods of higher degree agents.[9] In presence of such a hierarchical structure, the nested split graph is the same as the equitable partition. However, the latter is general enough to include other types of grouping structure for other role-based networks, whether hierarchical or not.

[9]See, for instance, König, Tessone, and Zenou (2014) for discussion on nestedness in networks.

6. Examples and discussion

In this section, we illustrate the two proposed class-based centrality measures on some example networks, and compare them with other common centrality measures. Three example networks are considered.

6.1. Example 1: class-based centrality vs intercentrality

Figure 1 considers the 11-player network $\mathbf{g}$ with three classes, as used in Ballester et al. (2006), and compares the proposed centrality measures with their intercentrality index, another centrality metric from planner's optimality concerns to identify the key player type. Table 1 computes centralities for two values of the decay factor a.[10] We find that

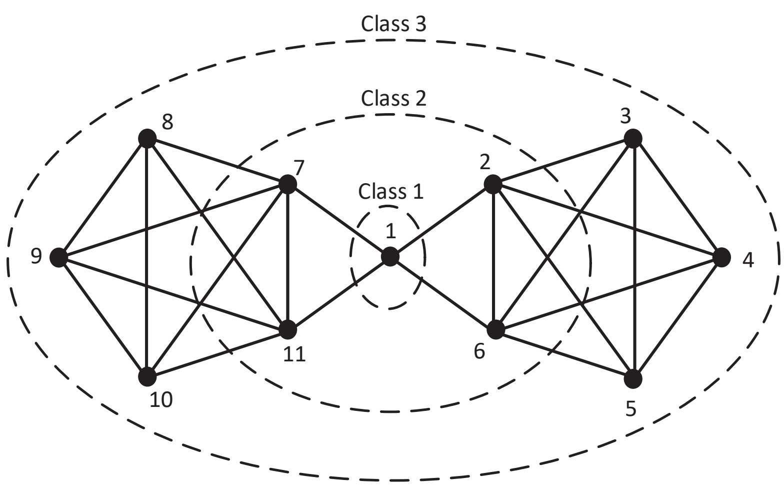

Figure 1. Example network: class-based centrality vs intercentrality

Table 1. Class-based centrality vs intercentrality

	$a = 0.1$			$a = 0.2$		
Class type	e_i	h_i	c_i	e_i	h_i	c_i
1	2.92	0.11	2.92	41.67	3.33	41.67∗
2	11.09	0.57∗	3.28∗	80.67	6.76∗	40.33
3	12.96∗	0.46	2.79	81.67∗	6.33	32.67

e_i and h_i denote class-centrality and per-capita class-centrality, respectively. c_i denotes intercentrality measure of Ballester, Calvó-Armengol, and Zenou (2006). The highest values are indicated by '∗'.

[10]Here, the maximum value of a in line with our centrality definitions is 0.227.

the largest class (class 3) is also the key class for reducing overall equilibrium activity, for both values of a. This is because along with having most members, this class is also quite well-connected. It has direct links with class 2 (which, by being the link between the other two classes, is the most central class - it's players have the highest Bonacich centrality), and indirect links with class 1. Hence, removing class 3 alters the network structure in a way to cause maximal disruption in equilibrium contribution by remaining players. However, in terms of per-capita network activity, class 2 becomes the most important one since it is smaller than class 3 but has direct links with both classes 1 and 3, removing which causes most damage to the network activity of the altered network, measured in per-capita terms.

We also note that the key class, both for total and per-capita outcome reduction, mostly differs from the player type with the highest intercentrality value. This is expected as intercentrality depends on an individual level analysis of peer-effects between pairs of players for characterising their importance, while class-based centrality internalises the group-level dynamics among the members within a class as well, in addition to studying the peer-effects across members of different classes. For the class with only one member (class 1), there is no such intra-group dynamics per se, and the intercentrality as well as class-centrality are the same, as also noted in Remark 1.

6.2. Example 2: class-based centrality vs Bonacich centrality

In the above example, the class which was most central was also the key class for optimally reducing per-capita activity. It is, however, not necessary that removing the most central class in terms of position alone, that is, whose players have the highest Bonacich centrality, will result in an optimal change in the structure of bilinear influences so as to minimise the per-capita network activity. This is evident in the example considered in Figure 2, borrowed from Allouch (2017) who considers segregation in social networks.

The class-based centrality values for the three classes, along with the Bonacich centrality for players in those classes is reported in Table 2, for two different values of a.[11] For this simple network where two of the classes are of same size, the key class for total and per-capita activity reductions comes out to be the same (class 2). Note that while class 1 is most centrally located, since its players, who have the highest Bonacich centrality,

[11]The largest value for a compatible with our definitions is 0.427.

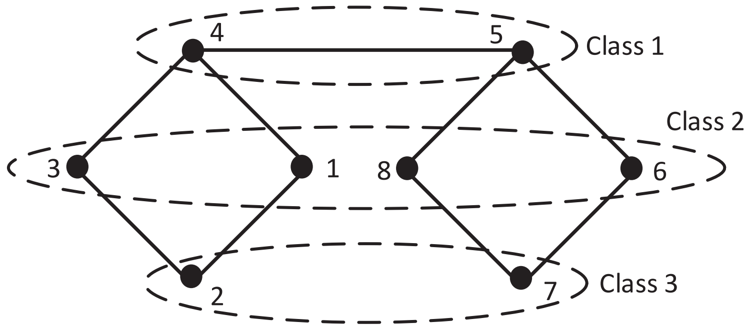

FIGURE 2. Example network: class-based centrality vs Bonacich centrality

TABLE 2. Class-based centrality vs Bonacich centrality

	$a = 0.1$			$a = 0.2$		
Class type	e_i	h_i	b_i	e_i	h_i	b_i
1	3.41	0.14	1.39∗	6.50	0.47	2.13∗
2	6.13∗	0.24∗	1.26	10.26∗	0.72∗	1.77
3	3.08	0.08	1.25	5.31	0.27	1.71

e_i, h_i and b_i denote class-centrality, per-capita class-centrality and Bonacich centrality, respectively. The highest values are indicated by '∗'.

form a bridge through which the other players are connected, it is not the key class, for either optimally reducing total or per-capita network activity. This is because for network activity, how removing a class alters the peer-effects within and across classes matter. Taking this into account makes class 2 the key class.

6.3. Example 3: Class-centrality need not be highest for the largest class

In the above two examples, we find that the key class for inducing maximal disruption in total network activity is the one that has most members. While this was true for the simplistic network structures considered in Figures 1-2, it will not, in general, be the case. We consider the example in Figure 3, from Bonaccorsi et al. (2015)'s study of epidemic outbreaks in networks with equitable partitions. Unlike the previous examples, this network is more complex and displays asymmetry in the indirect links between members of various classes (features which are likely to be present in realistic networks) - it can be seen that even though all players in class 4 have same number of links amongst themselves and with class 2, the indirect links for players 8 and 11 are different from others in their class.

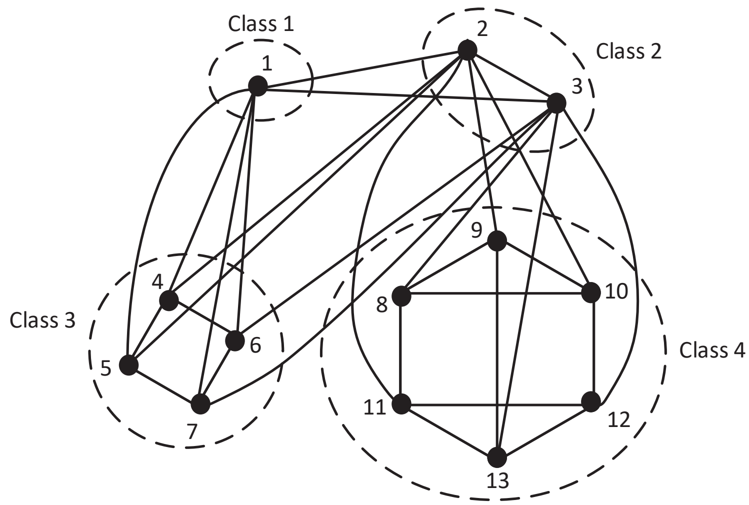

FIGURE 3. Example network: analysing class-centrality

Table 3 reports the centrality values for the aggregate and per-capita indices, for $a = 0.1$ and 0.2.[12]

TABLE 3. Analysing class-centrality

	$a = 0.1$		$a = 0.2$	
Class type	e_i	h_i	e_i	h_i
1	4.40	0.21	514.35	39.08
2	8.81	0.45∗	561.62∗	42.79∗
3	9.59	0.22	540	39.80
4	12.40∗	0.14	535.71	37.85

We focus on the class-centrality e_i which is informative for our purpose. Notice that for the lower value of a, the largest class (class 4) is also the key class, while with $a = 0.2$, class 2 (which is much smaller in size than class 4) becomes key for optimally decreasing overall network activity. This is because, with smaller value of a, the direct effect due to class size is the dominant factor in determining the key class. But when the indirect links become stronger, removing class 2, through which most of the indirect links are formed,

[12]For this example, the maximum permissible value of a to satisfy our centrality definitions is 0.204.

has the highest combined direct and indirect effects in determining the aggregate network activity.

7. Concluding remarks

This paper considers equitable partitioning of networks where players engage in a network game with local payoff complementarities, and proposes centrality measures to geometrically characterise the key class for the social planner who wishes to minimize (or maximize) the aggregate or the per-capita network activity. The measures derive from establishing a relationship between the Nash equilibrium activity of classes and their position within the network, in a result that can be considered the Bonacich-Nash equivalent of Ballester et al. (2006) for class-based networks. The work assumes importance in light of numerous examples of role-based stratification or hierarchy in networks, so that studying the direct and indirect effect of classes to the overall or per-capita economic activity in the network is critical for policy decisions.

While class, here, has been defined through the notion of equitable partition, an interesting and challenging future work would be to consider any general partitioning of networks, as defined in Van Mieghem (2010), in order to find the key class for any general grouping structure in networks.

Appendix: Proof section

Proof of Lemma 1. Both the inverse matrices are well-defined and non-negative for $0 \leq a \leq 1/\rho$. Then, since from (3.1) $\mathbf{G}^k\mathbf{X} = \mathbf{X}\mathbf{\Pi}^k$, we have

$$[\mathbf{I}_n - a\mathbf{G}]^{-1}\mathbf{X} = \left[\sum_{k=0}^{\infty} a^k \mathbf{G}^k\right]\mathbf{X} = \sum_{k=0}^{\infty} a^k \mathbf{X}\mathbf{\Pi}^k = \mathbf{X}\left[\sum_{k=0}^{\infty} a^k \mathbf{\Pi}^k\right]$$

which proves the Lemma. ■

Proof of Theorem 1. The pure Nash equilibrium strategies $\mathbf{x}^* \in \mathbb{R}_+^n$ for the network game in Section 4 solves $\partial u_i/\partial x_i(\mathbf{x}^*) = 0$, such that it satisfies the first order conditions:

$$[\mathbf{I}_n - \lambda\mathbf{G}]\,\mathbf{x}^* = \mathbf{1}_n$$

As shown in Ballester et al. (2006), the Nash equilibrium exists and is unique if the inverse $[\mathbf{I}_n - \lambda\mathbf{G}]^{-1}$ exists, that is, when $\lambda \leq 1/\rho$. Then, from definition of $\mathbf{b}(\mathbf{g}, \lambda)$ in (3.2),

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

Hence, from Lemma 1, we have

$$\mathbf{y}^*(\boldsymbol{\pi}) = \mathbf{X}^T.\mathbf{x}^* = \left[\mathbf{I}_m - a\mathbf{\Pi}^T\right]^{-1}.\mathbf{X}^T.\mathbf{1}_n$$

Noting that $\mathbf{X}^T.\mathbf{1}_n = \mathbf{r}$ then proves the Theorem. ■

Proof of Lemma 2. Recall that the elements of $\mathbf{\Pi}^p$, $\pi_{ik}^{[p]}$, denotes the total paths of length p for any v in class V_i with its neighbours in V_j. Let $\pi_{i(j^0)k}^{[p]}$ denote the total number of such paths not containing the class j. Similarly, $\pi_{i(j)k}^{[p]}$ denotes only such p-length paths that contain class j. Then, denoting the ik-th element of $(\mathbf{\Pi}^p)^T$ as $\pi_{ik}^{[p,T]}$ and setting $\pi_{jj}^{[0]} = 1$, for $0 \leq a \leq 1/\rho$, we have

$$N_{ik}(\boldsymbol{\pi}, a) - N_{ik}(\boldsymbol{\pi}^{-j}, a) = \sum_{p=1}^{\infty} a^p(\pi_{ik}^{[p,T]} - \pi_{i(j^0)k}^{[p,T]})$$

Noting that,

$$\pi_{ik}^{[p,T]} - \pi_{i(j^0)k}^{[p,T]} = \pi_{i(j)k}^{[p,T]} = \pi_{i(j)k}^{[p,T]}.\pi_{jj}^{[0,T]} = \sum_{\substack{r'+s'=p \\ r'\geq 1,\ s'\geq 1}} \pi_{ij}^{[r',T]}.\pi_{jk}^{[s',T]} - \sum_{\substack{r+s=p \\ r\geq 2,\ s\geq 1}} \pi_{i(j)k}^{[r,T]}.\pi_{jj}^{[s,T]},$$

we have

$$a^p \sum_{\substack{r+s=p \\ r\geq 2,\ s\geq 0}} \pi_{i(j)k}^{[r,T]}.\pi_{jj}^{[s,T]} = a^p \sum_{\substack{r'+s'=p \\ r'\geq 1,\ s'\geq 1}} \pi_{ij}^{[r',T]}.\pi_{jk}^{[s',T]}.$$

This equates to $[N_{ik}(\boldsymbol{\pi}, a) - N_{ik}(\boldsymbol{\pi}^{-j}, a)].N_{jj}(\boldsymbol{\pi}, a) = N_{ij}(\boldsymbol{\pi}, a).N_{jk}(\boldsymbol{\pi}, a)$ which proves the Lemma. ■

Proof of Theorem 2. Note that from Theorem 1, $y^*(\boldsymbol{\pi})$ and $y^*(\boldsymbol{\pi}^{-j})$ are increasing in $t(\boldsymbol{\pi}, \lambda)$ and $t(\boldsymbol{\pi}^{-j}, \lambda)$, respectively. Hence, the planner's objective function (5.2) can be re-written as follows:

$$\sum_{i=1,\ i\neq j}^{m} (t_i(\boldsymbol{\pi}, \lambda) - t_i(\boldsymbol{\pi}^{-j}, \lambda)) + t_j(\boldsymbol{\pi}, \lambda).$$

In what follows, we drop arguments in function for simplicity of notation wherever convenient, and write ik-th element of $\mathbf{N}(\boldsymbol{\pi}^{-j}, \lambda)$ as N_{ik}^{-j}. Since $\mathbf{t}(\boldsymbol{\pi}, \lambda) = \mathbf{N}(\boldsymbol{\pi}, \lambda).\mathbf{r}$, we re-write the above expression as:

$$\sum_{i=1,\ i\neq j}^{m} \left[\sum_{k=1}^{m} N_{ik} r_k - \sum_{k=1,\ k\neq j}^{m} N_{ik}^{-j} r_k\right] + \sum_{k=1}^{m} N_{jk} r_k$$
$$= \sum_{i=1,\ i\neq j}^{m} \left[N_{ij} r_j + \sum_{k=1,\ k\neq j}^{m} \left\{(N_{ik} - N_{ik}^{-j}) r_k\right\}\right] + \sum_{k=1}^{m} N_{jk} r_k.$$

Using Lemma 2, this becomes

$$\sum_{i=1,\ i\neq j}^{m} \left[N_{ij} r_j + \sum_{k=1,\ k\neq j}^{m} \left\{\frac{N_{ij}.N_{jk}}{N_{jj}} r_k\right\}\right] + \sum_{k=1}^{m} N_{jk} r_k$$
$$= \sum_{i=1,\ i\neq j}^{m} \left[\sum_{k=1}^{m} \frac{N_{ij}.N_{jk}}{N_{jj}} r_k\right] + \sum_{k=1}^{m} N_{jk} r_k = \sum_{i=1,\ i\neq j}^{m} \left[\frac{N_{ij}}{N_{jj}} t_j\right] + t_j \frac{N_{jj}}{N_{jj}} = \frac{t_j}{N_{jj}} \sum_{i=1}^{m} N_{ij}$$

where the last line uses the equality $t_j = \sum_{k=1}^{m} N_{jk} r_k$. Noting that $\sum_{i=1}^{m} N_{ij} = s_j$ proves the Theorem. ■

Proof of Theorem 3. As in proof for Theorem 2, the problem statement translates to:

$$\max \left\{ \frac{\sum_{i=1}^{m}(t_i(\boldsymbol{\pi},\lambda)}{n} - \frac{\sum\limits_{i=1,\ i\neq j}^{m} t_i(\boldsymbol{\pi}^{-j},\lambda))}{n-r_j} \equiv h_j(\boldsymbol{\pi},\lambda) \right\},\ j=1,\ldots,m.$$

Dropping arguments in function for simplicity of notation and denoting the ik-th element of $\mathbf{N}(\boldsymbol{\pi}^{-j},\lambda)$ as N_{ik}^{-j}, from $\mathbf{t}(\boldsymbol{\pi},\lambda)=\mathbf{N}(\boldsymbol{\pi},\lambda).\mathbf{r}$ such that $t_i=\sum\limits_{k=1}^{m} N_{ik}r_k$, we have

$$\begin{aligned} h_j &= \sum_{i=1,\ i\neq j}^{m} \left\{ \frac{(n-r_j)\sum_{k=1}^{m} N_{ik}r_k - n\sum_{k=1,\ k\neq j}^{m} N_{ik}^{-j}r_k}{n(n-r_j)} \right\} + \frac{\sum_{k=1}^{m} N_{jk}r_k}{n} \\ &= \sum_{i=1,\ i\neq j}^{m} \left\{ \frac{nN_{ij}r_j - r_j\sum_{k=1}^{m} N_{ik}r_k + n\sum_{k=1,\ k\neq j}^{m}\left(N_{ik}-N_{ik}^{-j}\right)r_k}{n(n-r_j)} \right\} + \frac{\sum_{k=1}^{m} N_{jk}r_k}{n} \end{aligned}$$

Using Lemma 2, this becomes

$$\begin{aligned} h_j &= \sum_{i=1,\ i\neq j}^{m} \left\{ \frac{nN_{ij}r_j - r_jt_i + n\sum_{k=1,\ k\neq j}^{m}\left(\frac{N_{ij}.N_{jk}}{N_{jj}}\right)r_k}{n(n-r_j)} \right\} + \frac{t_j}{n} \\ &= \sum_{i=1,\ i\neq j}^{m} \left\{ \frac{nN_{ij}r_j - r_jt_i + n\left\{\sum_{k=1}^{m}\left(\frac{N_{ij}.N_{jk}}{N_{jj}}\right)r_k - N_{ij}r_j\right\}}{n(n-r_j)} \right\} + \frac{t_j}{n} \\ &= \sum_{i=1,\ i\neq j}^{m} \left\{ \frac{n(N_{ij}/N_{jj})t_j - r_jt_i}{n(n-r_j)} \right\} + \frac{t_j}{n} = \sum_{i=1}^{m} \left\{ \frac{n(N_{ij}/N_{jj})t_j - r_jt_i}{n(n-r_j)} \right\}. \end{aligned}$$

Noting that $\sum\limits_{i=1}^{m} N_{ij} = s_j$, then, proves the Theorem. ■

References

Allouch, N. (2017). The cost of segregation in (social) networks. *Games and Economic Behavior*, *106*, 329–342.

Ballester, C., Calvó-Armengol, A., & Zenou, Y. (2006). Who's who in networks. Wanted: The key player. *Econometrica*, *74*(5), 1403–1417.

Baur, M., Brandes, U., Lerner, J., & Wagner, D. (2009). Group-level analysis and visualization of social networks. In *Algorithmics of large and complex networks* (pp. 330–358). Springer.

Belhaj, M., Bervoets, S., & Deroıan, F. (2013). Efficient networks in games with local complementarities. *Theoretical Economics*.

Bonaccorsi, S., Ottaviano, S., Mugnolo, D., & Pellegrini, F. D. (2015). Epidemic outbreaks in networks with equitable or almost-equitable partitions. *SIAM Journal on Applied Mathematics*, *75*(6), 2421–2443.

Bonacich, P. (1972). Factoring and weighting approaches to status scores and clique identification. *Journal of Mathematical Sociology*, *2*(1), 113–120.

Borgatti, S. P. (2006). Identifying sets of key players in a social network. *Computational & Mathematical Organization Theory*, *12*(1), 21–34.

Calderoni, F., Brunetto, D., & Piccardi, C. (2017). Communities in criminal networks: A case study. *Social Networks*, *48*, 116–125.

Calvó-Armengol, A., Patacchini, E., & Zenou, Y. (2009). Peer effects and social networks in education. *The Review of Economic Studies*, *76*(4), 1239–1267.

Currarini, S., Jackson, M. O., & Pin, P. (2009). An economic model of friendship: Homophily, minorities, and segregation. *Econometrica*, *77*(4), 1003–1045.

Everett, M. G., & Borgatti, S. P. (1999). The centrality of groups and classes. *The Journal of Mathematical Sociology*, *23*(3), 181–201.

Everett, M. G., & Borgatti, S. P. (2005). Extending centrality. In *Models and methods in social network analysis* (pp. 57–76). Cambridge University Press New York.

Golub, B., & Jackson, M. O. (2012). How homophily affects the speed of learning and best-response dynamics. *The Quarterly Journal of Economics*, *127*(3), 1287–1338.

Goyal, S., & Moraga-Gonzalez, J. L. (2001). R& D networks. *Rand Journal of Economics*, 686–707.

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

König, M. D., Tessone, C. J., & Zenou, Y. (2014). Nestedness in networks: A theoretical model and some applications. *Theoretical Economics*, *9*(3), 695–752.

Lerner, J. (2005). Role assignments. In *Network analysis* (pp. 216–252). Springer.

Liu, X., Patacchini, E., Zenou, Y., & Lee, L.-F. (2012). Criminal networks: Who is the key player? *FEEM Working Paper*.

Ottaviano, S., De Pellegrini, F., Bonaccorsi, S., Mugnolo, D., & Van Mieghem, P. (2019). Community networks with equitable partitions. In *Multilevel strategic interaction game models for complex networks* (pp. 111–129). Springer.

Powers, D. L., & Sulaiman, M. M. (1982). The walk partition and colorations of a graph. *Linear Algebra and its Applications*, *48*, 145–159.

Rahmani, A., Ji, M., Mesbahi, M., & Egerstedt, M. (2009). Controllability of multi-agent systems from a graph-theoretic perspective. *SIAM Journal on Control and Optimization*, *48*(1), 162–186.

Temurshoev, U. (2008). Who's who in networks-wanted: The key group. *Available at SSRN 1285752*.

Van Mieghem, P. (2010). *Graph spectra for complex networks*. Cambridge University Press.

Wasserman, S., & Faust, K. (1994). *Social network analysis: Methods and applications*. Cambridge University Press.