

RESEARCH

Open Access



A simple approach for quantifying node centrality in signed and directed social networks

Wei-Chung Liu¹ , Liang-Cheng Huang¹, Chester Wai-Jen Liu² and Ferenc Jordán^{3*}

* Correspondence: jordan.ferenc@gmail.com

³Balaton Limnological Institute, MTA Centre for Ecological Research, Klebelsberg Kuno, Tihany 3, 8237, Hungary

Full list of author information is available at the end of the article

Abstract

The position of a node in a social network, or node centrality, can be quantified in several ways. Traditionally, it can be defined by considering the local connectivity of a node (degree) and some non-local characteristics (distance). Here, we present an approach that can quantify the interaction structure of signed digraphs and we define a node centrality measure for these networks. The basic principle behind our approach is to determine the sign and strength of direct and indirect effects of one node on another along pathways. Such an approach allows us to elucidate how a node is structurally connected to other nodes in the social network, and partition its interaction structure into positive and negative components. Centrality here is quantified in two ways providing complementary information: total effect is the overall effect a node has on all nodes in the same social network; while net effect describes, whether predominately positive or negative, the manner in which a node can exert on the social network. We use Sampson's like-dislike relation network to demonstrate our approach and compare our result to those derived from existing centrality indices. We further demonstrate our approach by using Hungarian school classroom social networks.

Keywords: Centrality, Signed network, Directed network, Interaction structure

Introduction

Sociology is about the study of human social relations and how humans interact. More often than not we engage in intricate webs of social interactions, and making sense of them is by no means a simple task. In the past few decades, sociologists have tried and succeeded in delineating patterns of social interactions by the means of social network analysis. Individual human beings and their social interactions form a network. One important aspect of social network research from the graph perspective is how to quantify the position an individual occupies in a network. This particular development of social network research is quite intuitive since it is reasonable to envisage a linkage between the behavior of individuals, their network position and the structural correlates of group dynamics. Generally speaking, all this boils down to two simple and

related questions: first, how central is a network position; and second, are central positions important from a functional viewpoint?

There already exist several classical measures for quantifying the centrality of an individual in a social network. The simplest one is degree centrality that uses only local information, or the number of immediate neighbor a node has in this case, to quantify nodal centrality. There are also more complicated ones that use less-local information, these include geodesic-based measures such as closeness and betweenness centralities (Bonacich 1987; Wasserman and Faust 1994), and path or walk-based measures like information centrality (Stephenson and Zelen 1989) and Katz centrality (Katz 1953). More recently, given the fact that real social networks are large and are often characterized by the presence of interconnected communities, classical concepts of nodal centrality have been extended to incorporate two components (Cherifi et al. 2019; Ghalmane et al. 2019a, 2019b). One component measures a node's local influence within a community, and the other quantifies its global influence on others in different communities.

With the tool of quantifying nodal centrality at our disposal, the next step is to link centrality with functions. Several studies have demonstrated the positive association between individuals' centrality in a network and their work performances (Ahuja et al. 2003; El-Khatib et al. 2015); while others have shown empirically that central network positions are important and advantageous in various animal social groups (Wey et al. 2008; Jordán 2009; de Silva et al. 2011). The concept of node centrality as a proxy to importance has also been adapted by physicists and biologists and applied in these fields too. For instance, it has been suggested that several physical and biological networks are robust against random errors simply because the majority of connections are centered around a few highly connected nodes called hubs (Albert et al. 2000; Callaway et al. 2000). In ecology, for example, centrality indices have been used to predict the importance of species in ecological communities and the possible effects of their removal on the ecosystem (Dunne et al. 2002; Jordán et al. 2006; Estrada 2007).

Although there are methods for quantifying centrality for directed graphs (e.g. directed betweenness centrality (Borgatti et al. 2002)), and there has been a huge amount of research made on signed graphs (e.g. loop sign and structural balance (Cartwright and Harary 1956; Antal et al. 2006); spread of information and social influence (Li et al. 2015; He et al. 2019); and community detection (Traag and Bruggeman 2009; Esmailian and Jalili 2015)), it is only until recently that methods measuring node centrality for networks of signed interactions start to emerge. For instance, Bonacich and Lloyd (2004) developed a status measure by using eigenvector, where a high status node usually has many positive relations with other members of the same clique and negative relations with members of other cliques. In a very different context, Smith et al. (2014) developed Political Independence Index (PII) to quantify the power of a node by assessing its social dependence on others; and a powerful node here is the one that avoids dependence on its allies and is also less likely to be threatened by its adversaries. Furthermore, Everett and Borgatti (2014) developed a negative centrality measure where a node with fewer negative ties to important nodes is more central; and they further combined this concept with Hubbell's centrality measure for positive ties (Hubbell 1965) and proposed PN centrality for signed digraphs.

In this paper, we present a different approach for quantifying nodal centrality for signed digraphs. Our methodology is based on three intuitive assumptions. First, the likelihood of an individual being affected by one of his/her friends depends on the size of his/her social circle (Dunbar 2018). If an individual has many acquaintances then the probability of he/she being affected by one of them should be small. In contrast, if an individual has only a few acquaintances, say in the extreme case where he/she has only one friend, then the probability of this individual being affected by his/her sole acquaintance should be high. Second, effects such as social influence and sentiments can spread among individuals (Fowler and Christakis 2008; Chan et al. 2018); and following the logic behind our first assumption mentioned above, it is not unreasonable to consider the effect of i on j via k as the product of two probabilities, namely the effect of i on k and the effect of k on j . Third, basing on the notion of transitivity and structural balance, which has been frequently studied on signed social networks (Cartwright and Harary 1956; Du et al. 2016; He et al. 2019), if the effect of i on k and the effect of k on j are of the same signs, then their product, or namely the effect of i on j via k , should be positive; whereas if the effect of i on k and the effect of k on j are of the opposite signs, then the effect of i on j via k should be negative. With those assumptions in mind, the basic idea behind our approach is to consider how positive and negative effects from a node i can propagate through a network. By doing so we can quantify both positive and negative effects of a node i on another node j up to a give number of steps, and then aggregate such information for all j s in order to obtain a new centrality measure for node i . A node i has high centrality if it can affect others (including itself) with strong effects via many different pathways. Our approach is of instrumental nature and has its root in ecology (Müller et al. 1999; Jordán et al. 2003; Liu et al. 2010); and here in this paper we employ and generalize this concept for social network analysis. Since our concept here involves how effects can propagate from a node to others in all direction and pathways, we note that it is also akin to those node centrality measures basing on walks between nodes, this includes for example Katz centrality (Katz 1953) and exponential centrality (Estrada 2012; Benzi et al. 2013). What differs is that, here a walk from nodes i to j constitutes of several edges, each of which has its own probability of passing on the effect; thus the effect of nodes i on j via a walk is the product of probabilities of those constituting edges. And through the multiplication of signs, effect associated with a walk can be positive or negative. We also note that there has been research on how a node can influence another in the context of regulation of cellular pathways (Campbell et al. 2011). Although such a study considers the signs of edges in a pathway, but it considers pathways as paths instead of walks. Thus, effects there, unlike those in our concept, strictly do not travel in all directions. Furthermore, influence of a node on another defined there is based on the number of paths between them (and subtracting the number of negative paths from positive ones); and in contrast, our concept here envisages effect along a pathway as the product of probabilities of constituting edges.

In a nutshell, our concept of centrality here is of instrumental nature, and given a network, we wish to quantify the effect of one node on another by tracking how such an effect can spread from the source to the target. The sum of these effects reflects the magnitude of its effect on the whole network. Intuitively, a node in an important or central position of a network should easily have its effect spreading out to others. Thus

the total effect a node has on the whole network can be regarded as a measure of the centrality of its network position. In sociology terminology, such effects can be seen as influence of one on others, and larger effect may imply larger influence; this in turn might translate to the power of an actor in influencing others in the same social network, and this consequently should affect the prestige of such actor. Our paper is organized as follows. First, we implement our concept of centrality and describe our methodology for undirected and unsigned graphs; then extend it to signed graphs and then to signed digraphs. Second, we apply our methodology to social network analysis with three examples, specifically on quantifying nodal centrality: a) a toy network for demonstration purpose; b) Sampson's monastery social network dataset (Sampson 1968) for comparing our methodology with those existing centrality measures from the literature; and c) a classroom social network dataset as a case study. Finally, we discuss the sociological implications of our results and the potential use of our methodology in other fields of social network analysis.

Unsigned, undirected graphs

Let us consider a connected graph consisting of N nodes. The direct effect (i.e. one-step effect) of i on j is:

$$a_{ij} = \frac{1}{D_j}, \quad (1)$$

where D_j is the degree of node j and node i is one of its neighbors. For a pathway consisted of more than one steps, we define the indirect effect from the starting node on the ending node via this pathway as the product of the constituent direct effects (i.e. effects are multiplicative). For instance, consider a simple case where there is a 2-step pathway $i-k-j$ and the indirect effect of i on j through k is defined as:

$$a_{ik} \times a_{kj} = \frac{1}{D_k} \times \frac{1}{D_j}. \quad (2)$$

We further assume that effects are additive: if there are only two 2-step pathways linking i and j , one via k and the other via h , then the indirect effect of i on j through the pathway $i-k-j$ is given by (2) while the other effect through the pathway $i-h-j$ is given by:

$$a_{ih} \times a_{hj} = \frac{1}{D_h} \times \frac{1}{D_j}, \quad (3)$$

and the effect of i on j in 2 steps is the sum of these two indirect effects:

$$a_{ik} \times a_{kj} + a_{ih} \times a_{hj}. \quad (4)$$

In general, we define $TE_{ij,n}$ as the total effect of i on j in n steps; and if there are m n -step pathways from i to j , then $TE_{ij,n}$ is the sum of those m indirect effects. We further define an $N \times N$ matrix \mathbf{TE}_n whose ij^{th} element is $TE_{ij,n}$.

Note that direct effects and indirect effects defined above can be interpreted in terms of probability. Direct effect of i on j (i.e. a_{ij}) can be regarded as the probability of j being affected by its neighbor i as far as one step effects are concerned. While the indirect effect of i on j in two steps is equivalent to the probability of j being affected by i if we

consider all pathways of length two ending with j . More specifically, the probability of j being affected by i along one particular pathway of length two is the product of the two constituent direct effects along that pathway (i.e. product of two probabilities); and the probability of j being affected by i in two steps along *any* pathways is the sum of all indirect effects of length two from i to j (i.e. sum of probabilities).

Since there may exist several pathways of various lengths between any two nodes, in order to quantify the effect of one node on another, all pathways of different lengths should be taken into account. Here, we define the cumulative total effect of i on j up to n steps as:

$$CTE_{ij,n} = g(1)TE_{ij,1} + g(2)TE_{ij,2} + g(3)TE_{ij,3} + \dots g(n)TE_{ij,n}, \quad (5)$$

which is simply the sum of total effects each weighted by a function $g(l)$, where l is the corresponding number of steps considered. Here, each TE_{ij} is weighted to account for the possibility that indirect effects of longer pathway lengths should intuitively be weaker than those of shorter pathway lengths. We then construct the \mathbf{CTE}_n matrix where the ij^{th} element is $CTE_{ij,n}$. Lots of useful information are provided by the interaction matrix \mathbf{CTE}_n . First, the i^{th} row encodes the information on how i influences all others including itself, while the i^{th} column records how every node affects node i . Thus, such information can be regarded as the interaction profile of a particular node i . The sum of the i^{th} row is the total effect of i on the whole network, and can be interpreted as a centrality measure for node i up to n steps:

$$\eta_{i,n} = \sum_{j=1}^N CTE_{ij,n}. \quad (6)$$

The sum for the j^{th} column is the total effect received by j up to n steps:

$$\mu_{j,n} = \sum_{i=1}^N CTE_{ij,n}. \quad (7)$$

Signed, undirected graphs, and digraphs

The above methodology can be extended to signed graphs when direct effects are re-defined appropriately. First, we define $i-j^S$ as the link connecting i and j . The superscript S is the sign of interaction: it is either positive (+1) or negative (−1). Let node j has D_j neighbors, and the direct effect of i on j is defined as:

$$c_{i-j^S} = S \times \frac{1}{D_j} = S \times a_{ij}. \quad (8)$$

Since direct effects are now either positive or negative, and indirect effects by definition are the products of direct effects, the indirect effects of i on j can then be partitioned into two components, one positive and one negative. We define $E_{ij,n+}$ as the sum of positive effects of i on j for n steps, and similarly, we define $E_{ij,n-}$ for negative effects. We then construct two interaction matrices \mathbf{E}_{n+} and \mathbf{E}_{n-} whose ij^{th} elements are $E_{ij,n+}$ and $E_{ij,n-}$ respectively. Next, we separately work out the positive cumulative effect and its negative counterpart as:

$$CE_{ij,n+} = g(1)E_{ij,1+} + g(2)E_{ij,2+} + g(3)E_{ij,3+} + \dots g(n)E_{ij,n+}, \quad (9)$$

$$CE_{ij,n-} = g(1)E_{ij,1-} + g(2)E_{ij,2-} + g(3)E_{ij,3-} + \dots g(n)E_{ij,n-}. \quad (10)$$

We can do this for all node pairs and organize positive and negative effects in the \mathbf{CE}_{n+} and \mathbf{CE}_{n-} matrices, respectively. For a given node i , the corresponding rows and columns in \mathbf{CE}_{n+} and \mathbf{CE}_{n-} may be combined as a single vector representing the interaction profile of node i . Note that the cumulative total effect matrix \mathbf{CTE}_n simply equals to:

$$\mathbf{CTE}_n = \mathbf{CE}_{n+} + |\mathbf{CE}_{n-}|. \quad (11)$$

Eq 11 defines the total effects between all node pairs up to n steps regardless of the signs of interactions. A new interaction matrix can be derived from \mathbf{CE}_{n+} and \mathbf{CE}_{n-} by simply adding them up:

$$\mathbf{NE}_n = \mathbf{CE}_{n+} + \mathbf{CE}_{n-}, \quad (12)$$

which is also a square matrix where the ij^{th} element records the net effect of node i on node j . There is a major difference between interaction matrices \mathbf{CTE}_n and \mathbf{NE}_n : while the former has positive entries in it, the latter has entries that may take also zero or negative values.

From interaction matrix \mathbf{NE}_n , we define the following indices:

$$\eta_{i,n,\text{net}} = \sum_{j=1}^N NE_{ij,n}, \quad (13)$$

$$\mu_{j,n,\text{net}} = \sum_{i=1}^N NE_{ij,n}. \quad (14)$$

(13) is the net effect exerted by a node i up to n steps on the whole network, while (14) is the net effect it received from the whole network. Note that, in contrast to the case for unsigned graph, (13) can be positive, negative or zero, indicating whether node i affects the network positively, negatively or effectively none at all (as the positive effect and negative effect that i exerts cancel out each other). Furthermore, (14) can now take a value from the interval $[-n, n]$: a positive value for node i indicates that the effect it receives from the network is mainly positive, while a negative value means the effect it receives is predominately negative. There might be cases where node i receives a net effect of zero implying that positive effect it receives cancels out negative effect. Finally, for signed digraphs, let $i \rightarrow j^S$ denotes the directed link from node i to node j , indicating i influences j in a positive or negative manner as determined by S . Let D_j^{in} be the number of neighbors influencing j (i.e. we only count the number of links directed towards j , or simply its in-degree), then the magnitude of the unsigned direct effect from i on j is:

$$a_{ij} = \frac{1}{D_j^{\text{in}}}, \quad (15)$$

with the sign of interaction, the signed direct effect from i on j is thus:

$$c_{i \rightarrow j^S} = S \times \frac{1}{D_j^{\text{in}}} = S \times a_{ij}. \quad (16)$$

Importantly, in digraphs it is possible that a node is not reachable from another; thus, when quantifying the indirect effect of node i on node j , we only consider those pathways via which i can reach j . With these differences in mind, methods developed for signed graphs can be applied directly here.

The form of weighting function $g(l)$ and choosing the number of steps n

In (5), (9) and (10), effects of different lengths l are weighted by a function $g(l)$. This is to account for the possibility that indirect effects of longer pathway lengths should intuitively be weaker than those of shorter pathway lengths. Here, we intend to keep our approach general and do not explicitly specify $g(l)$. However, $g(l)$ should be in principle a decreasing function of l , and one example is:

$$g(l) = \frac{1}{l^x}. \quad (17)$$

For simplicity, we have opted for the case where effects have equal weights throughout this paper (i.e. $g(l) = 1$). Furthermore, we also need to choose a value for the maximum number of steps n in order to implement our methodology. Recently Chan et al. (2018) have shown that sentiment and mood can propagate from an individual to another via one intermediary; therefore an appropriate choice for n would be two. However, in this paper (unless stated otherwise), we extend this value slightly to three such that effects travelling along other important pathways can be captured in our analysis. For instance, with $n = 3$, we can also quantify looping-back effect in a triad (i.e. $i \rightarrow j \rightarrow k \rightarrow i$). To summarise, unless stated otherwise, we have used $g(l) = 1$ and $n = 3$ throughout this paper. Note that by doing so, effects from pathway length greater than three will not be included in the calculation.

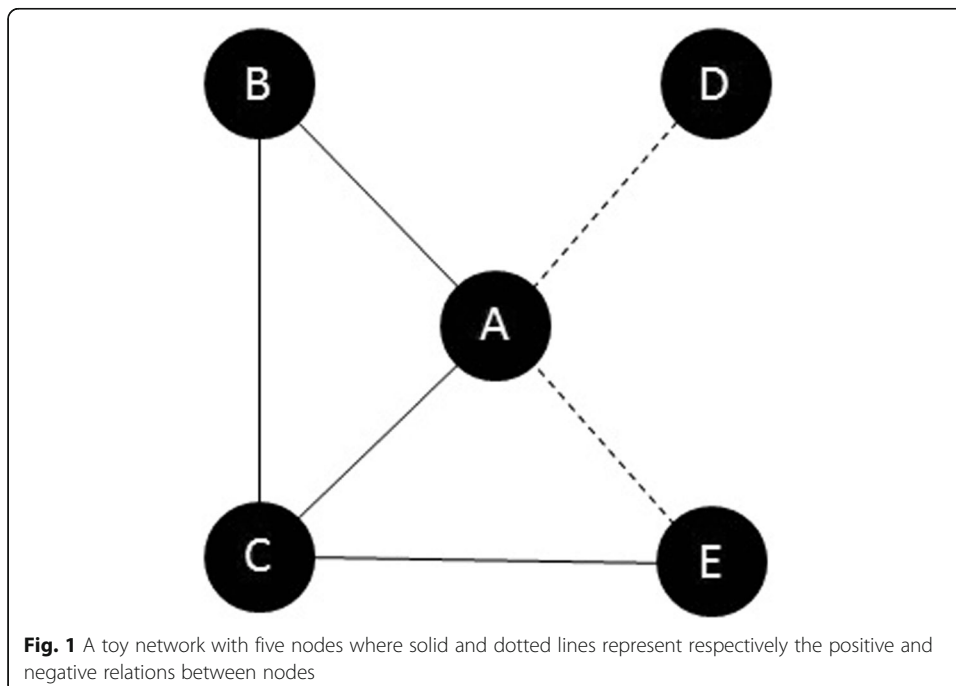


Fig. 1 A toy network with five nodes where solid and dotted lines represent respectively the positive and negative relations between nodes

A toy network

We demonstrate our methodology on a simple toy network. Figure 1 is a toy network with five nodes, A, B, C, D and E. In this network, solid lines represent positive ties while dotted line represents negative ties. For simplicity, we apply our new methodology up to 2 steps with equal weights (i.e. $g(l) = 1$ and $n = 3$). Table 1 summarizes the positive effect of one node on another up to 2 steps, while Table 2 summaries similar information for negative effects. Let's consider the effect of A on C up to two steps. A can affect C in one step as well as in two steps. The one step effect of A on C is positive and is simply $1/3$ (as C has three positive ties and no negative ties). For 2-step effects, A can affect C in two ways. One is via B, and the effect along pathway A-B-C is $1/2 * 1/3 = 1/6$. One is via E, and the effect along pathway A-E-C is $-1/2 * 1/3 = -1/6$. The positive effect of A on C up to 2 steps is $1/3 + 1/6 = 1/2$, and its negative counterpart is $-1/6$. Thus the total effect of A and C up to 2 steps is $1/2 + 1/6 = 2/3$; and the net effect of A on C up to 2 steps is: $1/2 - 1/6 = 1/3$. Total effect exerted by each node is summarized in the last column of Table 3, and the net effect exerted by each node is summarized in the last column of Table 4. We can observe that A has the strongest total effect of 3.58, whereas D has the smallest total effect of 0.83. As for the net effect, C has the largest magnitude of 1.42 and it mainly affects the whole network in a positive manner; whereas D has the smallest magnitude of 0.083 and it does this in a negative manner.

Comparison with other centrality measures

In this section we compare our methodology (with pathway length up to three with equal weights, i.e. $g(l) = 1$ and $n = 3$) to three existing centrality measures designed for signed relations, they are the status score, PII and PN indices (Bonacich and Lloyd 2004; Smith et al. 2014; Everett and Borgatti 2014). The dataset used is the like and the dislike relations from Sampson's monastery social network dataset (Sampson 1968).

We follow the work of Bonacich and Lloyd (2004) and present separately the like relation and the dislike relation between 18 monks (Figs. 2 and 3). Basing on the like relation there are visually two cliques with the majority of the links among the clique members. Using the same partition of monks we can observe that most dislike relations are between cliques (note that there is no dislike relation in clique A but there are some in clique B). The status score of Bonacich and Lloyd (2004) shows that members of those two cliques have scores of opposite signs (Table 5): clique A members have positive scores while clique B members have negative scores. Regardless signs, it can be

Table 1 Positive effect of one node on another up to 2 steps for the toy network shown in Fig. 1. Each cell records the positive effect from one node in a particular row on another node in a particular column

	A	B	C	D	E
A	0.583	0.667	0.5	0	0.167
B	0.333	0.292	0.417	0	0.167
C	0.375	0.625	0.417	0	0.5
D	0	0	0	0.25	0.125
E	0.083	0.167	0.333	0.25	0.292

Table 2 Negative effect of one node on another up to 2 steps for the toy network shown in Fig. 1. Each cell records the negative effect from one node in a particular row on another node in a particular column

	A	B	C	D	E
A	0	0	−0.167	−1	−0.5
B	0	0	0	−0.25	−0.125
C	−0.125	0	0	−0.25	−0.125
D	−0.25	−0.125	−0.083	0	0
E	−0.25	−0.125	−0.083	0	0

observed that in clique A, Peter (i.e. PETER) has the highest score because it is not only the most popular individual in clique A (i.e. involving in many like relations with other clique members), but it also has many dislike relations with members of clique B (i.e. having negative ties with members of disvalued or rival clique can promote one's status). In contrast, Bonaventure (i.e. BONAVEN) is as popular as Peter in clique A but his status score is low because he lacks negative ties with members of clique B. Similar reasoning applies to members of clique B. Gregory, Basil, Elias, Simplicius (i.e. GREG, BASIL, ELIAS, SIMP) have the most negative scores as they have many negative ties with members of clique A.

Applying PII index (Smith et al. 2014) to the Sampson like-dislike relation network shows completely different result. PII measures the power of a node which can be defined on how its neighbours are dependent on it. In essence, a node is at a powerful position if all of its interacting neighbours have no connections to other nodes. The effect of negative relation can drastically decreases a node's power because it potentially feels threatened by its adversaries. Here PII index suggests that Bonaventure has by far the largest PII score whereas the remaining monks have relatively very low PII scores (Table 5). Note that all monks with the exception of Bonaventure have negative relation with some others, and this decrease their power. In contrast, Bonaventure has the largest PII scores for two reasons: first is that he has no negative relation with others; and second, his interacting neighbours all have negative relations with some others, which in turn increases Bonaventure's power.

The PN index developed by Everett and Borgatti (2014) extends and generalizes the notion of having tie to a central node increases one's centrality to cover signed digraphs. A node with positive ties to important nodes tends to have higher centrality, whereas negative ties tend to decreases one's centrality. Applying the PN index to the Sampson's like-dislike relation network, one can observe that Bonaventure has the

Table 3 Total effect of one node on another up to 2 steps for the toy network shown in Fig. 1. Each cell records the total effect from one node in a particular row on another node in a particular column. The last column records the total effect each node has on the whole network

	A	B	C	D	E	Total effect
A	0.583	0.667	0.667	1	0.667	3.583
B	0.333	0.292	0.417	0.25	0.292	1.583
C	0.5	0.625	0.417	0.25	0.625	2.417
D	0.25	0.125	0.083	0.25	0.125	0.833
E	0.333	0.292	0.417	0.25	0.292	1.583

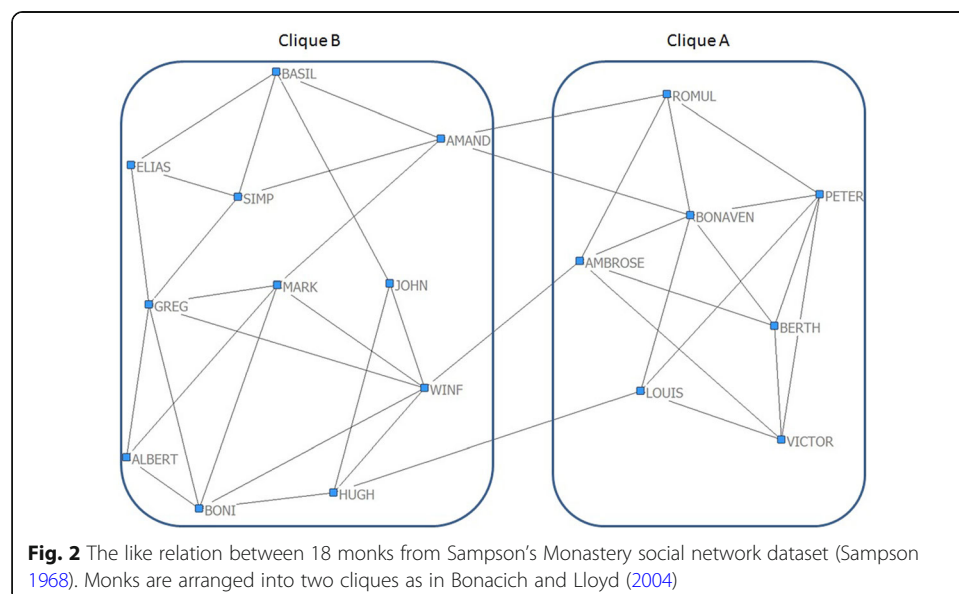
Table 4 Net effect of one node on another up to 2 steps for the toy network shown in Fig. 1. Each cell records the net effect from one node in a particular row on another node in a particular column. The last column records the net effect each node has on the whole network

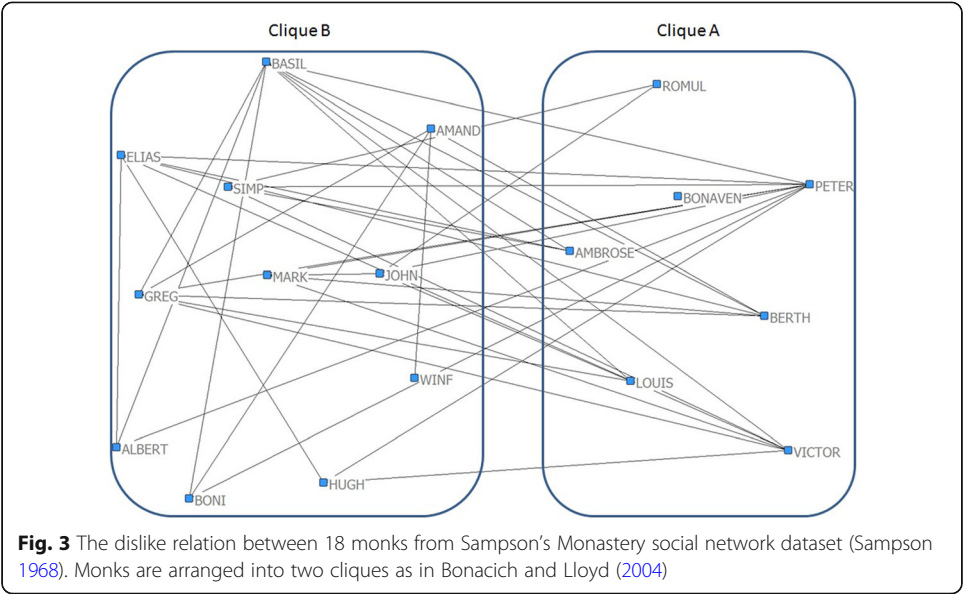
	A	B	C	D	E	Net effect
A	0.583	0.667	0.333	−1	−0.333	0.25
B	0.333	0.292	0.417	−0.25	0.042	0.833
C	0.25	0.625	0.417	−0.25	0.375	1.417
D	−0.25	−0.125	−0.083	0.25	0.125	−0.083
E	−0.167	0.042	0.25	0.25	0.292	0.667

highest PN score (Table 5) as he is involved in the highest number of positive relations while engaged in no negative relations. PN index suggests that the least central monks are Basil, Elias and Simplicius as they have several negative ties with other important individuals (Fig. 3).

Basing on our methodology, as far as total effect exerted by an individual is concerned, Peter, Gregory and Basil are the most important individuals (Table 5). This is because they have many ties with others regardless signs, and this renders them higher influence on others in the same network. However, if we consider the net effects they exerted on the whole network, then Gregory is a positive interactor while Peter and Basil are negative interactors. As far as the net effect is concerned, Bonaventure and Ambrose are the two individuals with the strongest positive effect, whereas Basil, Elias, Simplicius and Peter mainly exert negative effects on the network. Figure 4 is a scatter plot showing clearly the relationship between total effect and net effect of each monk. We suggest that examining the centrality of individuals should take into account both total effect and net effect: total effect provides information on the influence, both positive and negative, one exerts on the whole network; while net effect tells us in what way, predominately positively or negatively, an individual affects the whole network.

Furthermore, we determined the similarity between each pair of centrality indices. Here two measures were used to quantify similarity: one is the well-known Kendall

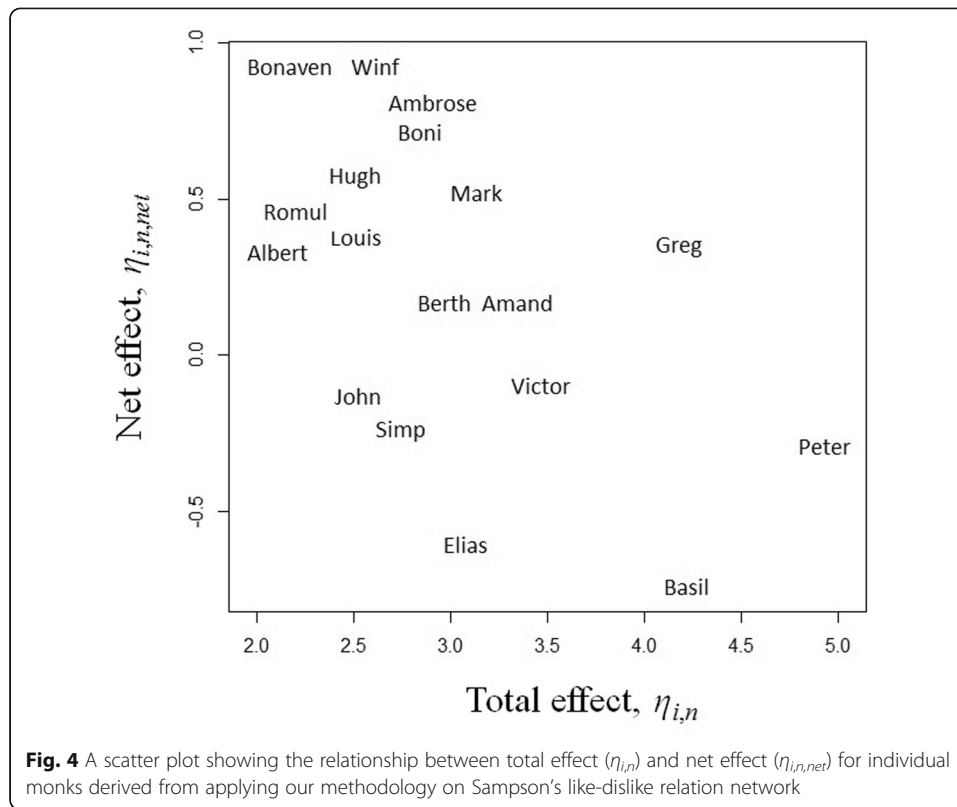




rank correlation coefficient from standard statistics; and the other is the rank-biased overlap approach (RBO) from information science (Webber et al. 2010). Results are summarized in Table 6. We also determined the correlation between those Kendall rank correlation coefficients and those RBO values; and a high rank correlation coefficient ($\tau = 0.911$, $p < 0.01$) between these two sets of similarity measures indicates consistency in two different approaches for assessing the similarity between centrality

Table 5 Different centrality values calculated on the like-dislike relation network from Sampson's Monastery social network dataset. TE and NE stand for total effect and net effect as derived from our new methodology, all other values were calculated using UCINET (Borgatti et al. 2002)

Monk	Status	PII	PN	TE	NE
Romul	−0.174	0.306	1.007	2.175	0.460
Bonaventure	−0.188	2.523	1.159	2.154	0.937
Ambrose	−0.248	0.249	1.008	2.872	0.796
Berthold	−0.319	0.190	0.861	3.076	0.157
Peter	−0.420	−0.605	0.671	5.025	−0.290
Louis	−0.219	0.332	0.963	2.443	0.381
Victor	−0.365	−0.051	0.800	3.433	−0.110
Winfred	0.081	0.937	1.114	2.496	0.897
John	0.142	−0.138	0.870	2.522	−0.134
Gregory	0.292	0.396	0.880	4.174	0.366
Hugh	0.088	0.478	0.985	2.478	0.576
Boniface	0.123	0.424	1.011	2.782	0.727
Mark	0.217	0.256	0.955	3.149	0.508
Albert	0.072	0.610	0.959	1.981	0.328
Amand	0.030	−0.300	0.911	3.222	0.156
Basil	0.254	0.379	0.680	4.196	−0.754
Elias	0.282	−0.210	0.761	3.085	−0.643
Simplicius	0.287	0.774	0.890	2.736	−0.247



indices. In general, the similarity between each pair of centrality indices is not high, however there are two exceptions. First, the PII index and the total effect have moderate negative correlation (-0.438), and their RBO value is the lowest in Table 6 (0.162). And second, the PN index and the net effect show strong and positive correlation (0.791), and their RBO value is the largest in Table 6 (0.890). A node will have a high PII score if its interacting neighbours have no other relations with others (i.e. those neighbours are all dependent entirely on the focal node), but such a node should have low total effect here as it has fewer channels to influence others in the same network. Thus this explains the negative correlation between PII score and total effect. The strong positive correlation between PN index and net effect here suggests that both methodologies might provide similar information on node centrality despite the fact that they respective are derived in different contexts. According to our methodology, a node can influence others in a profound and positive manner if it has positive ties with important nodes; whereas negative ties to important nodes can be efficient channels through which one can exert negative influence to others. Thus, all this amount to similar centrality property that PN index was intended to portray, and therefore explains the high correlation between those two measures.

A case study using a Hungarian high school social network dataset

In this section we apply our methodology to a high school social network dataset (with pathway length up to three and equal weights, i.e. $g(l) = 1$ and $n = 3$), and examine whether boys and girls are different in terms of how they affect others, as well as how

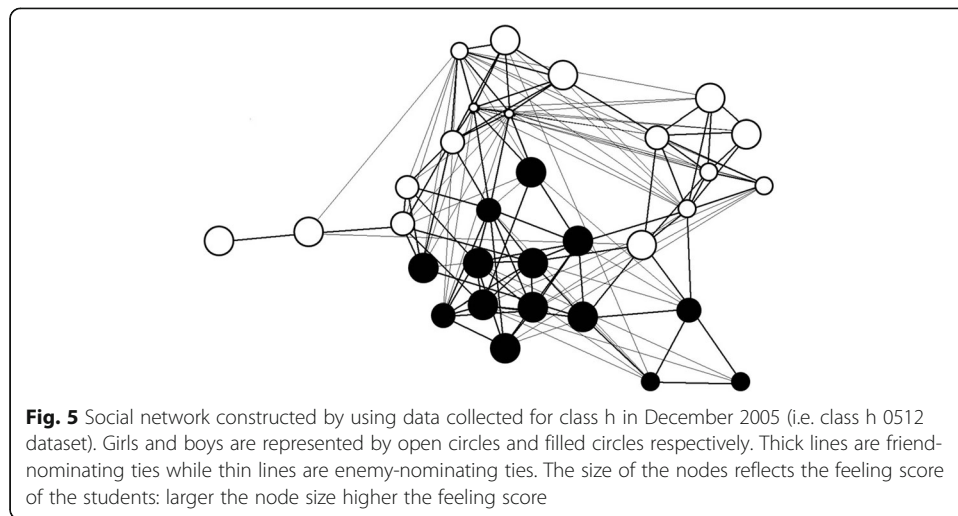
Table 6 Measuring similarity between pairs of centrality indices from Table 5. In each cell, the first value is Kendall rank correlation coefficient and the second value is RBO similarity value. TE and NE stand for total effect and net effect as derived from our new methodology

	PII	PN	TE	NE
Status	0.216, 0.522	−0.072, 0.310	0.059, 0.407	−0.098, 0.400
PII		0.477, 0.678	−0.438, 0.162	0.399, 0.682
PN			−0.542, 0.148	0.791, 0.890
TE				−0.386, 0.238

they are being affected by others in the same class. Furthermore, we also examine whether various effects derived from our methodology are related to their mood in the corresponding class. We used data from eight Hungarian classes sampled over a 3-year period. The classes belonged to Budapest secondary schools, kids were either aged 12–15 or 15–18 years old during the survey. On each sampling date, we recorded the gender and the feeling score for all students and asked them to nominate their friends and enemies. The feeling score is a rough measurement of how they feel in the school ranging on an integer scale from very bad (1) to very good (5). After the data were collected, all students' names were replaced with unique ID numbers such that the identity of individual students can never be revealed. The full dataset and a brief summary are provided in the [supplementary information](#), and Fig. 5 provides an example of a classroom social network derived from our dataset.

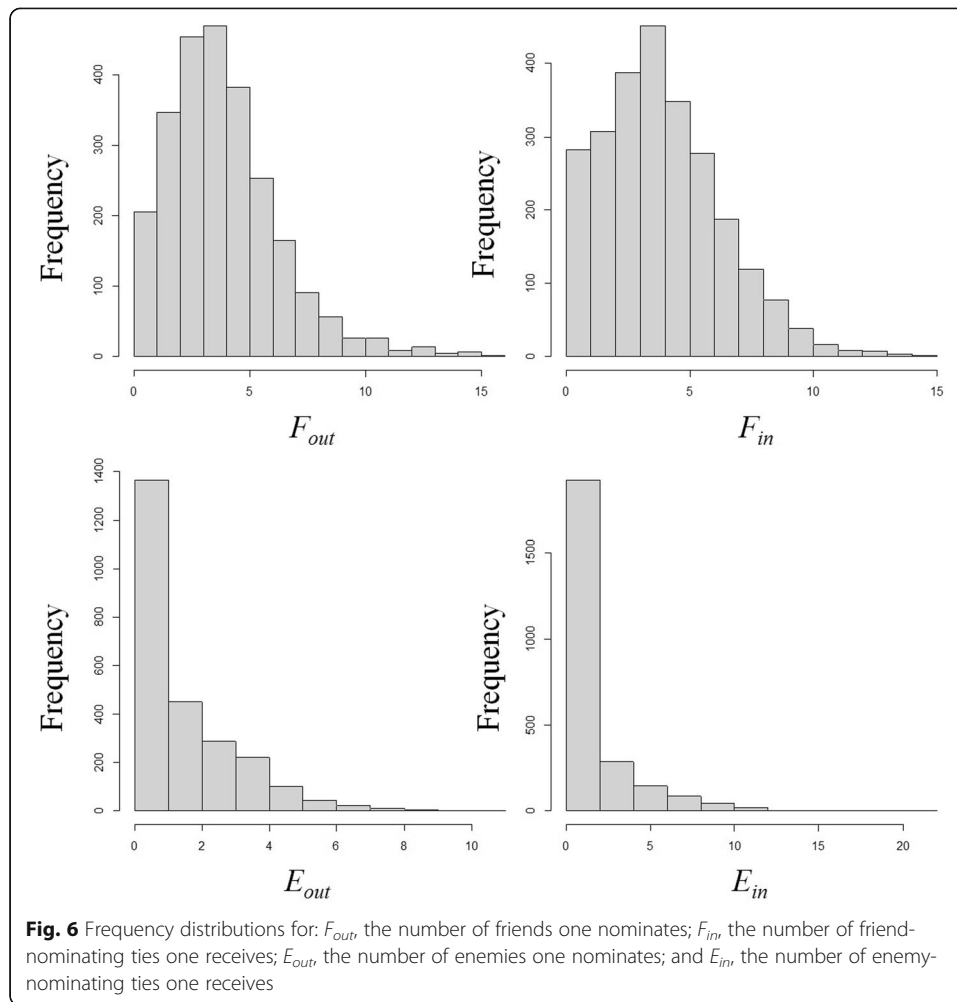
For each student in a dataset, we count the number of friend-nominating ties and the number of enemy-nominating ties he/she makes, and define those as F_{out} and E_{out} respectively. Furthermore, we determine the number of friend-nominating ties and the number of enemy-nominating ties he/she receives, and let these be F_{in} and E_{in} respectively. Note that F_{in} can be regarded as the popularity of a student, F_{out} can be seen as a measure how unpopular a student is (i.e. public enemy). Figure 6 shows the distributions of those four quantities for the whole dataset. It can be seen that F_{out} and F_{in} have bell-shaped distributions slightly skewed to the right, with an average around four. In contrast, E_{out} and E_{in} have highly skewed distributions to the right where the majority of the students make or receive no or very few enemy ties. There is no correlation between the number of friends (i.e. F_{out}) and the number of enemies (i.e. E_{out}) a student nominates ($r < 0.01$, $p = 0.982$). In contrast, there is some negative correlation between the number of friend-nominating ties (i.e. F_{in}) and the number of enemy-nominating ties (i.e. E_{in}) a student receives ($r = -0.30$, $p < 0.01$), indicating that popular individuals have fewer enemies. There is also some correlation between the numbers of friend-nominating ties one makes and receives ($r = 0.41$, $p < 0.01$), showing that outgoing individuals tend to be popular in the class. And lastly, there is weak correlation between the numbers of enemy-nominating ties one makes and receives ($r = 0.12$, $p < 0.01$).

For each class, we constructed an influence network from a friendship network as follows. First, we assume a friend-nominating tie from student i to student j meaning that i believes j is his/her friend; therefore we assume that j can influence i positively. Second, for the enemy-nominating tie, if i nominates j , then we assume j can influence i negatively. We calculated the total effect and the net effects (both exerted and received) for individual students in their respective networks by using our methodology up to three steps (i.e., $n = 3$), and investigated how they correlate with the numbers of friend



and enemy-nominating ties (Table 7). We found that the numbers of friend and enemy-nominating ties one receives tend to correlate with the total effect he/she exerts on the whole social network (i.e. F_{in} and TE, E_{in} and TE). This translates to a setting where popular individuals or public enemies tend to affect the whole social network more than their lesser counterparts. However, total effect alone cannot tell us in what way an individual affect others. Once the sign and the direction of social interactions are taken into account, we can observe that the number of friend-nominating ties one receives correlate strongly with the net effect he/she exerts on the whole social network (i.e. F_{in} and NE exerted). This implies that popular individuals affect the whole social network in a more positive way. In contrast, the number of enemy-nominating ties one receives shows strong negative correlation with the net effect he/she exerts on others (i.e. E_{in} and NE exerted), implying that public enemies tend to affect the whole social network in a more negative manner. One other interesting observation from our analysis is that, the net effect received by an individual tends to correlate positively with the number of nominated friends (i.e. F_{out} and NE received), but negatively with the number of nominated enemies (i.e. E_{out} and NE received); and the magnitude of the correlation for the latter relationship is almost twice that of the former. This suggests that, while making friends might have some positive effect on an individual, but making enemies can certainly affect him/her more and in a very negative manner.

More interesting information can be gained from our analysis of this particular dataset using our methodology. First, there is some positive correlation between an individual's exerted net effect and received net effect ($r = 0.34$, $p < 0.01$); this demonstrates that if one influences others more positively then he/she will also receive more positive effects from others (Fig. 7). Furthermore, it appears that the majority of nodes are positive-positive interactors (i.e. the top-right quadrant of Fig. 7) and a substantial amount of nodes are either positive-negative or negative-positive interactors (the top-left and bottom-right quadrants of Fig. 7), while only very few nodes are negative-negative interactors (bottom-left quadrant of Fig. 7). Second, taking into account information on gender, we partitioned all observations into male and female groups; and calculated the mean total effect and the mean net effects exerted and received for each sex, and determined their associated 95% confidence intervals by using bootstrap. We



found that boys (mean total effect = 3.05, 95% CIs = 2.95–3.16) tend to exert significantly stronger total effect on the whole class than girls do (mean total effect = 2.75, 95% CIs = 2.67–2.83). Moreover, we found that boys (mean net effect exerted = 1.05, 95% CIs = 0.96–1.14) marginally tend to influence the network more positively than girls (mean net effect exerted = 0.97, 95% CIs = 0.89–1.03). Also, boys tend to receive significantly more positive net effects (mean net effect received = 1.20, 95% CIs = 1.13–1.25) than girls do (mean net effect received = 0.83, 95% CIs = 0.78–0.90).

Lastly, we also tested how various effects calculated using our methodology are associated with an individual's feeling score. Since feeling was scored discretely using a scale

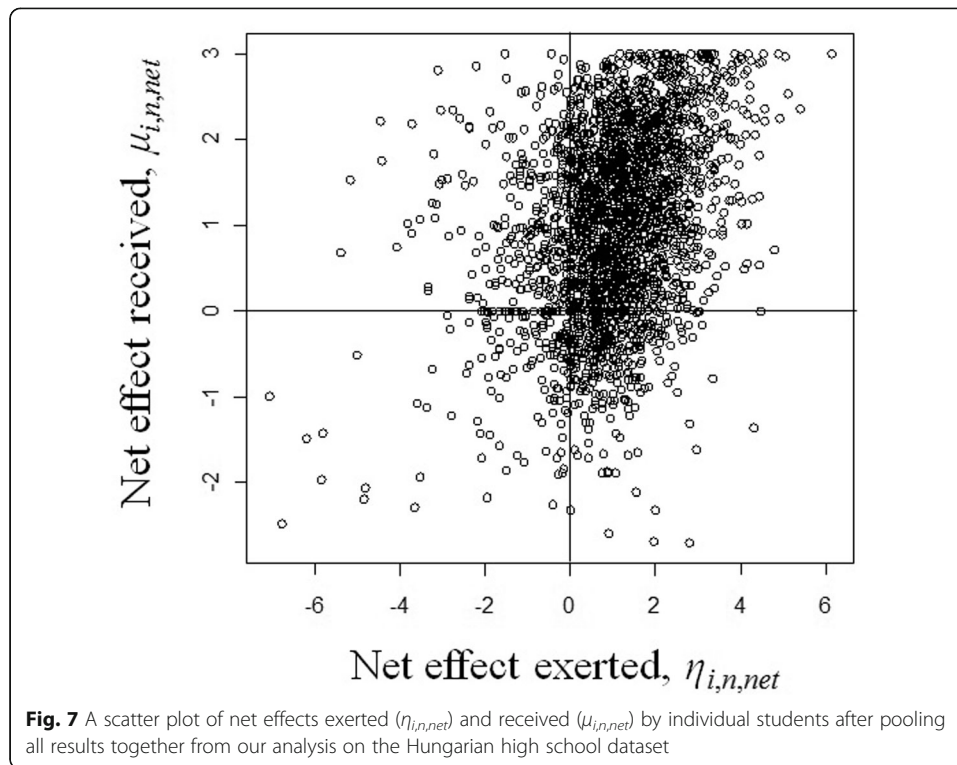
Table 7 Correlation coefficient between numbers of friend and enemy-nominating ties and the effects calculated from our methodology. Numbers in parenthesis are p -values for the significance of the corresponding correlation analysis

	F_{out}	F_{in}	E_{out}	E_{in}
TE	0.012 (0.540)	0.396 (< 0.01)	− 0.029 (0.150)	0.552 (< 0.01)
NE (exerted)	0.280 (< 0.01)	0.685 (< 0.01)	− 0.143 (< 0.01)	− 0.716 (< 0.01)
NE (received)	0.355 (< 0.01)	0.188 (< 0.01)	− 0.714 (< 0.01)	− 0.205 (< 0.01)

from one to five, for each effect type we calculated the mean effect for each feeling score and determined its 95% confidence intervals by bootstrap. Figure 8 (top figure) shows that there is no apparent pattern showing the relationship between total effect and feeling scores. Again, it is only once the direction and sign of social interactions are being taken into account, we then start to see interesting patterns. Figure 8 (middle figure) also shows that, for low to mid feeling scores, there is little difference between them in terms of their exerted net effects; however, for higher feeling scores, they tend to have much more positive exerted net effects. The relationship between effect size and feeling score is even more apparent for received net effects as shown in Fig. 8 (bottom figure): increasing feeling scores meets an increase in received net effect in a more positive manner. All in all, higher feeling scores are associated with strong and positive net effects (both exerted and received). We also investigated how other centrality indices for signed graphs are associated with feeling scores. We calculated status score, PII and PN centralities for this dataset by using UCINET (Borgatti et al. 2002). For each centrality index we calculated its mean for each feeling score and determined its 95% confidence intervals by bootstrap. Figure 9 shows that higher feel scores are also associated with higher centrality values. However, there are some overlaps in the 95% confidence intervals, therefore the trend here is not as pronounced as the trend for net effects.

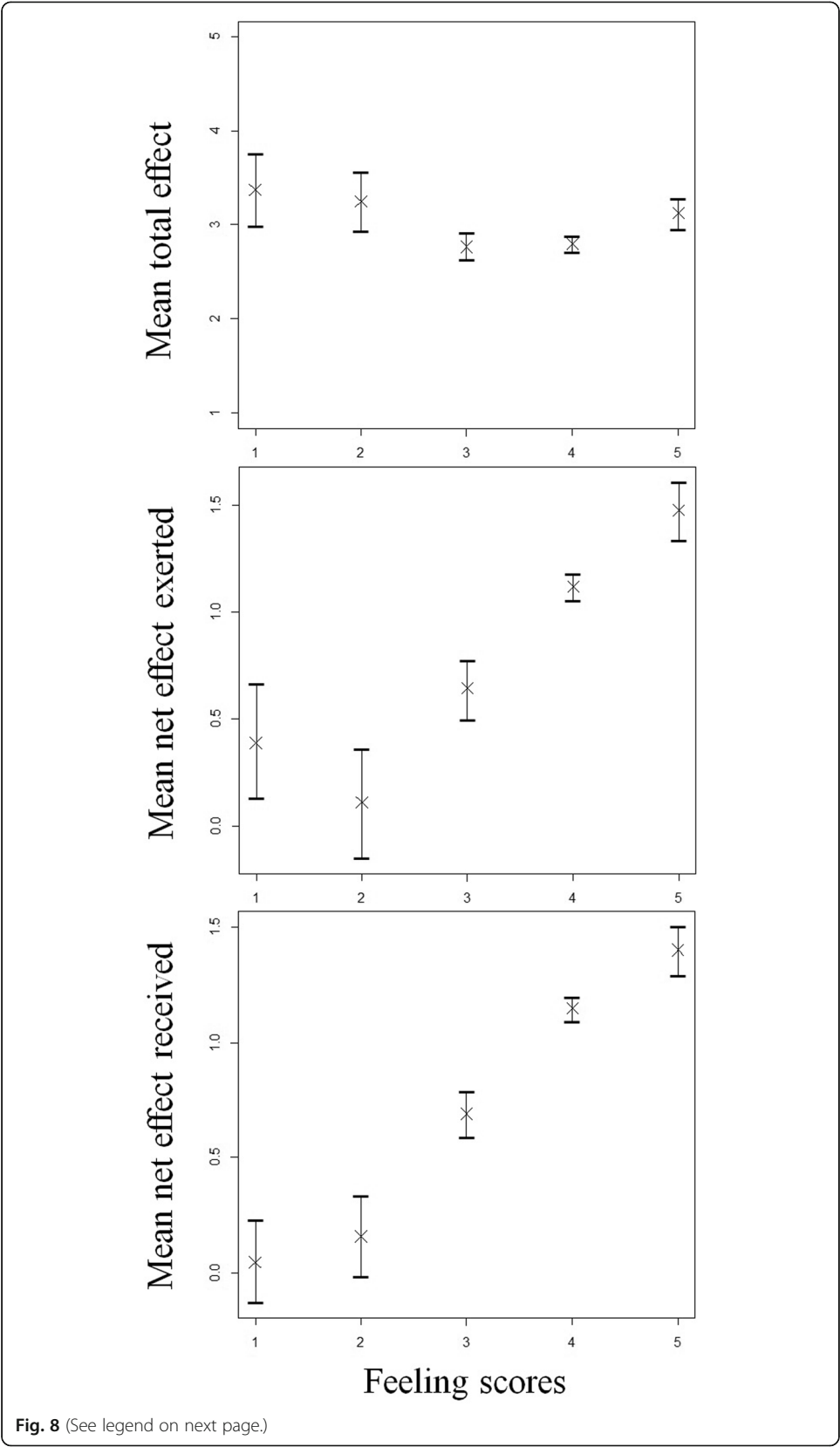
Testing the effect of step length on total effects and net effects

In this study, a step length of two was used in our demonstrative example using the toy network, and a step length of three was used in our analysis of Sampson's monastery dataset and Hungarian school dataset. In this section we examine how step length n can affect total effect and net effect. Again, we analysed Sampson's monastery dataset and quantified for each monk his total effect and net effect up to n steps with n varies from one to ten. Kendall rank correlation coefficient between total effect up to n steps and that up to $n-1$ steps was then calculated. Figure 10 shows that correlation coefficient (i.e. thin solid line) quickly increases and attains the value of one after three steps. This suggests that centrality ranking of monks as measured by the total effect stabilizes and remains the same after only three steps. As for net effect, correlation coefficient also quickly increases and stabilizes after four steps (thin dotted line in Fig. 10). In this case, correlation coefficient approaching a value very close to one suggests that rankings of monks are very similar for n beyond four. Centrality rankings of those monks, as measured by total effect and net effect, stabilise so quickly might be due to the small size of the social network such that effects can spread out quickly. In order to investigate the effect of social network size on this, we also performed the same analysis on a larger classroom social network shown in Fig. 5. In this case, centrality ranking of students (thick solid line and thick dotted line in Fig. 10) stabilizes at a slower rate than that for Sampson's monk social network. Thus, in order to use our methodology, one might suggest that a cut-off value for n should be chosen basing on when centrality ranking stabilizes. Although this might be a good cut-off for n , but we suggest that the value of n chosen should also be guided by empirical studies on how far social influence can spread from one individual, and this warrants future research.



Comparison with centrality indices for unsinged and undirected networks

In the last part of our analysis, we provide results on how our centrality indices compared to those classic centrality indices developed for unsinged and undirected networks (Bonacich 1987; Stephenson and Zelen 1989; Wasserman and Faust 1994). Degree centrality is the simplest one measuring the number of connection a node has; and eigenvector centrality is a degree-related index that also takes into account the importance of a node's connected neighbours. Betweenness centrality is a shortest path-related measure counting the number of times a node appearing in all shortest paths in a network. Closeness centrality is another shortest path-related measure quantifying the distance between the focal node and all others in the same network (i.e. how close they are). Katz centrality measures the centrality of a node by quantifying the contributions from all other nodes, with more distant nodes being penalized more than less distant ones. All of those indices translate to a notion that important nodes are those occupying important position in a network. Therefore, our intuition suggests that nodes ranking high in those centrality indices should have large total effect on the whole network. The reason is that by being in an important position, a node can affect many others, or effects can spread from it to many others, in a fewer number of steps. Again, we analysed Sampson's monastery social network data and quantified for each monk his degree, betweenness, closeness, eigenvector and Katz centralities (Table 8), and investigated the similarity between those indices and our indices proposed in this paper. Again, like before, we employed Kendall rank correlation coefficient and RBO value to measure the similarity between centrality indices. As expected, total effect correlates with, or produces similar ranking patterns to those classic centrality indices (Table 9). When the direction and the sign of social interactions are taken into account,



(See figure on previous page.)

Fig. 8 Mean total effect (top), mean net effect exerted (middle) and mean net effect received (bottom) plotted against the feeling score of individual students after pooling all results together from our analysis on the Hungarian high school dataset (crosses and bars are means and the 95% confidence intervals respectively)

then such similarity starts to disappear. As suggested by our result, net effect calculated by using our methodology shows weaker correlation or less similarity with those classic centrality indices, and it even shows negative relationship with them (Table 9).

Discussion and conclusion

In this paper we have presented a simple method for quantifying the interaction structure for graphs, signed graphs and signed digraphs. This is achieved by considering how a node can affect another via direct and indirect pathways. Once the interaction structure of the whole network is determined, the nodal centrality can be calculated by considering the total effect a node has on all nodes in the same network. The total effect of a node can be further partitioned into positive and negative components; and subtracting the latter from the former results in a net effect which tells us in what way, whether predominately positive or negative, a node affects the whole network. Our methodology is based on a very different context to those from the literature, and we have demonstrated weak correlations between our indices and those established indices (with the exception of our net effect and the PN index). Therefore our approach here should provide an alternative perspective on nodal centrality.

We note that our methodology of measuring node centrality is in essence a walk-based approach. Two well-known walk-based approaches are Katz centrality (Katz 1953) and exponential centrality (Estrada 2012; Benzi et al. 2013), and their computation involves the multiplication of the adjacency matrix. The number of times the adjacency matrix is being multiplied equals to the length of a walk, and the ij^{th} element of the resulting matrix records the number of walks starting from node i and ending at node j . Our approach also relies on walks between nodes, and this can be represented in the form of matrix multiplication. In fact, our methodology, as presented in great detail in the method section, can be succinctly represented in matrix multiplication form. But what differs here from other walk-based approaches is that we replace adjacency matrix with interaction matrix of step length one, where the ij^{th} element of this matrix is the direct effect of node i on node j , which is determined by the inverse of node j 's in-degree. Ignoring the signs of interaction, centrality values calculated using our approach, or the total effect of individuals nodes, should show some correlation with other walk-based centralities. And this is indeed the case as we have demonstrated using Sampson's monastery social network data, that there is strong correlation between our total effect and Katz centrality. But once the sign of interactions is taken into account to calculate nodal centrality, we then obtain a very different centrality measure (i.e. net effect) to Katz centrality.

Given the diversity of centrality measures we have at our disposal nowadays, it is natural to ask which centrality measure is the best one. We believe this depends on the context one is interested in or on the subject matter. For instance, a recently developed and more advanced measure that quantifies centrality in two dimensions has been

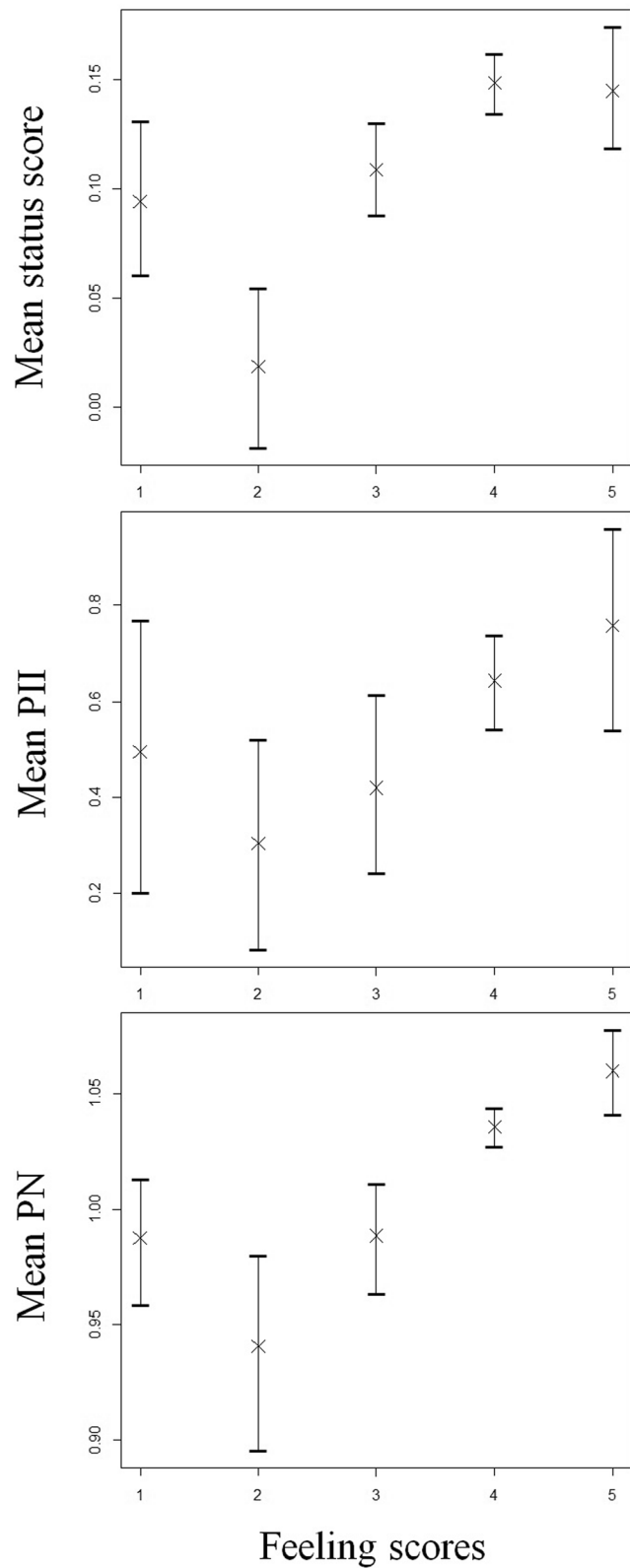
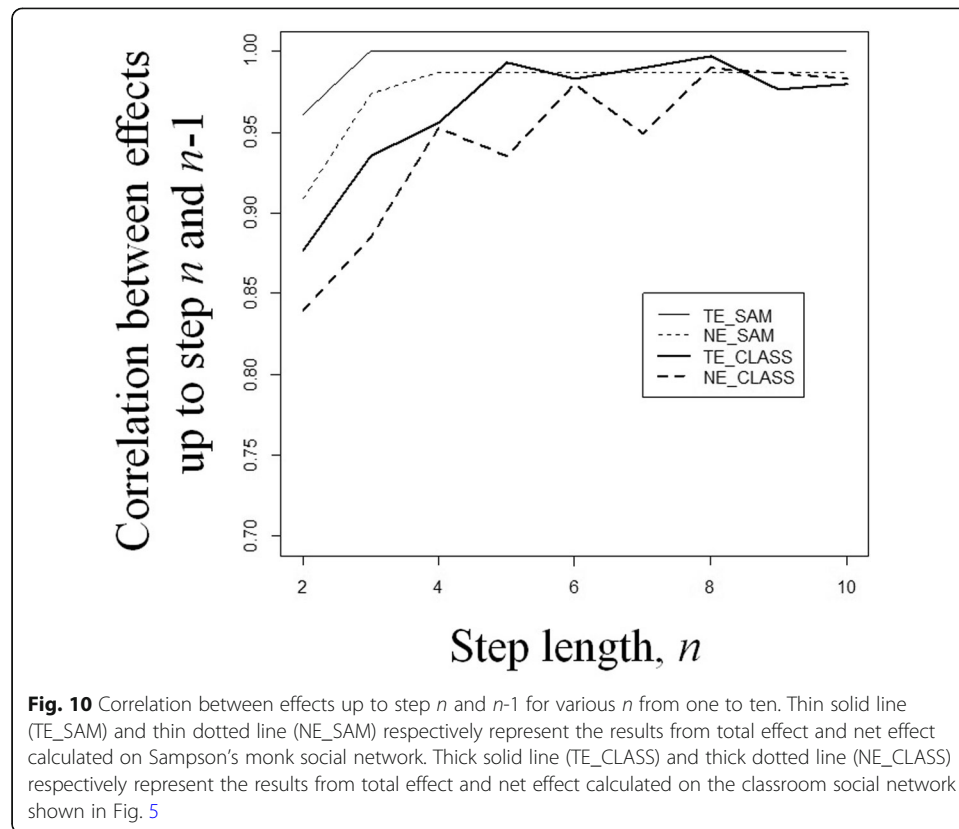


Fig. 9 Mean status score (top), mean PII centrality (middle) and mean PN centrality (bottom) plotted against the feeling score of individual students after pooling all results together from our analysis on the Hungarian high school dataset (crosses and bars are means and the 95% confidence intervals respectively)



shown to predict better disease control strategy in large social networks with interconnected communities (Cherifi et al. 2019; Ghalmane et al. 2019a, 2019b); while simple indices such as degree centrality can predict essential genes from the protein interaction network of yeast (Jeong et al. 2001). Moreover in biology, it has been shown that host species with high closeness centrality in a food web tend to harbour many parasite species (Chen et al. 2008); whereas closeness centrality performs poorly in predicting the conservation of different enzymes across different bacterial species in a metabolic network (Liu et al. 2007). In a more systematic manner, several studies have investigated the relationship between difference centrality measures, and it has been found that some are more closely related while some are not (Jordán et al. 2006; Oldham et al. 2019), indicating some provide redundant information and some provide complimentary information. And a more recent study has shown that the relationship between different centrality measures varies depending on the topology and the density of networks analysed (Oldham et al. 2019). All in all, given the fact that each centrality index measures a different aspect of network position; we thus suggest that centrality measures should not be used in isolation, and at least several should be tried out, in order to gain a better picture of nodal centrality of the studied network.

The next step of research relating to our work here will be on how to correlate results derived from our methodology with sociological theory and phenomenon. A testable hypothesis is that, if a social group is predominately negative in its interaction structure, will this correspond to a particular aspect of group behavior. For instance, balance theory suggests that a balanced triad with the product of signs being positive should promote group stability (Cartwright and Harary 1956; Lerne 2016). We can

Table 8 Different centrality values calculated on the like-dislike relation network from Sampson's Monastery social network dataset by using centrality measures for unsigned and undirected graphs. Deg, Btw, Cls, Eig and Katz are degree, betweenness, closeness, eigenvector and Katz centralities. TE and NE stand for total effect and net effect as derived from our new methodology. Deg, Btw, Cls and Eig were calculated using Netdraw (Borgatti et al. 2002)

Monk	Deg	Btw	Cls	Eig	Katz	TE	NE
Romul	6	1.935	28	0.147	0.066	2.175	0.460
Bonaventure	6	1.885	28	0.151	0.066	2.154	0.937
Ambrose	8	4.493	26	0.199	0.087	2.872	0.796
Berthold	9	3.076	25	0.258	0.099	3.076	0.157
Peter	14	15.235	20	0.348	0.153	5.025	-0.290
Louis	7	2.267	27	0.196	0.077	2.443	0.381
Victor	10	4.127	24	0.281	0.110	3.433	-0.110
Winfried	7	3.350	27	0.172	0.077	2.496	0.897
John	7	3.175	27	0.176	0.077	2.522	-0.134
Gregory	12	7.093	22	0.323	0.132	4.174	0.366
Hugh	7	2.700	27	0.181	0.077	2.478	0.576
Boniface	8	2.600	26	0.220	0.088	2.782	0.727
Mark	9	3.925	25	0.241	0.099	3.149	0.508
Albert	6	0.525	28	0.188	0.067	1.981	0.328
Amand	9	5.735	25	0.227	0.098	3.222	0.156
Basil	12	8.361	22	0.319	0.132	4.196	-0.754
Elias	9	3.269	25	0.256	0.099	3.085	-0.643
Simplicius	8	2.251	26	0.231	0.089	2.736	-0.247

simply adapt this idea and ask whether a social group with a more positive interaction structure tend to have less conflict between its constituent individuals. The only finding so far that relates our result with individuals' attributes is from our analysis on the Hungarian school classroom social networks. In that particular analysis, we have shown that popular individuals and public enemies tend to affect the whole social network more than others do, with the former do so in a positive manner and the latter in a more negative manner. This result seems intuitive and obvious, and what is less

Table 9 Measuring similarity between pairs of centrality indices from Table 8. In each cell, the first value is Kendall rank correlation coefficient and the second value is RBO similarity value. TE and NE stand for total effect and net effect as derived from our new methodology

	Btw	Cls	Eig	Katz	TE	NE
Deg	0.616, 0.626	0.628, 0.786	0.841, 0.928	0.936, 0.917	0.936, 0.960	-0.433, 0.239
Btw		0.820, 0.771	0.490, 0.615	0.556, 0.772	0.791, 0.890	-0.255, 0.334
Cls			0.560, 0.840	0.936, 0.917	0.936, 0.960	-0.433, 0.238
Eig				0.948, 0.999	0.712, 0.875	-0.464, 0.158
Katz					0.765, 0.879	-0.438, 0.158
TE						-0.386, 0.238

obvious is the finding that making friends might have some positive effect on an individual, but making enemies can affect him/her more in a very negative manner. This echoes findings from the literature that negative influence or negative effect tend to have a stronger impact than positive ones (Baumeister et al. 2001; Dong et al. 2011; Cheng et al. 2015). We further found that boys tend to affect the whole social network more than girls do, and this is in line with the general observation that males tend to be more influential than females (Eagly 1983; Carli 2001). Our results show that boys marginally influence the network in a more positive manner than girls do; but the effect received by boys are significantly more positive than that received by girls. Our gender-related findings echo those from previous studies in psychology: mainly, boys tend to report positive influences from peers than girls whereas girls tend to report more negative effects from peers than boys (Fujita et al. 1991; Bagozzi et al. 1999).

We have also shown that happier individuals tend to exert and receive more positive net effects. We believe this is intuitive and suggests that our methodology is at least sound in relating interaction structure and perceived personal feeling of individuals.

One of the most active fields of research regarding social networks is community detection (Girvan and Newman 2002; Fortunato and Hric 2016; He et al. 2018) where similar actors are being grouped in the same clusters forming various communities. The majority of the methods rely on the pattern of connection in which case actors in the same cluster should have more ties between them than with those from other clusters. Here we conceptually offer a different approach. From our methodology, an interaction matrix can be constructed where the ij^{th} element records the effect of node i on node j . With a user-defined cutoff, node j is node i 's strong interactor if its effect is greater than such a cutoff. From this we then derive a matrix of strong interactors where the ij^{th} element is one if node j is node i 's strong interactor. Note that in this way, two nodes can be strong interactors even if they don't have direct connection between them. Such a matrix can then be subjected to conventional community detection analysis; and nodes that influence each other strongly and positively can be considered to be in the same social group despite there might not be direct linkages between them. However, exploring this possibility is beyond the scope of our paper here and we should defer this till a later work. Another potential use of our methodology is in the role analysis of actors in a social network. Role analysis also has a long tradition in social network analysis and it is often employed to reduce a complex social network to simpler ones (White and Reitz 1983; Borgatti and Everett 1993). Traditionally, actors with the same or similar connection pattern (i.e. structural and regular equivalence) can be grouped in the same role class, this then results in a network of different role classes. Here, we can do similar things by using the interaction matrix derived from our methodology. Again, recorded in the interaction matrix is the effect of one node on another, and the i^{th} row of this matrix can be considered as the interaction profile of node i (i.e. how it interacts with all others, including itself, in the network). Similarity in the interaction profile between node pairs can be quantified using measures often employed in role analysis. This then results in a similarity matrix between nodes that can be used for conventional role analysis. Again, exploring this possibility is beyond the scope of our paper here and we should defer this till a later work.

Finally, recent developments of centrality measurement have emphasized the importance of incorporating the structure of large real-world social networks (Cherifi et al.

2019; Ghalmane et al. 2019a, 2019b). Those networks are characterized by interconnected communities. It has been suggested that centrality in this case should have two components, one local and one global (or at least at some mesoscale), each respective measures how a node influences those in the same community and others in other communities. Although our methodology can investigate effects at the local, meso and global scales by varying the number of steps n , but it does not considered the structure of community-rich social networks in the way that was done in recent literatures. Therefore, a fruitful aspect of future research on centrality analysis will be to extend our methodology to taking into account the community structure of real-world social networks.

Supplementary information

Supplementary information accompanies this paper at <https://doi.org/10.1007/s41109-020-00288-w>.

Additional file 1. The Hungarian school classroom data are given in the form of eight Excel files, each for a particular class. Those eight classes are identified as classes “a”, “b”, “c”, “d”, “e”, “f”, “g” and “h”. For instance, data for class “a” are in the file a.xlsx. Each file has several Excel spreadsheets. A spreadsheet with a label “YYMM” is the social network data collected on month MM in year 20YY, and the data is organized in three columns: the first and the second columns, namely “nominator” and “nominee”, contain the identifiers for the students; while the last column, namely “relation”, contains whether the “nominator” nominates the “nominee” as a friend (coded “1”) or as an enemy (coded “-1”). A spreadsheet with a label “gen” contains the gender information for each student, and it has two columns: the “id” column contains the identifier for a particular student whose gender is recorded in the column “gender” (coded “0” for girls and “1” for boys). A spreadsheet with a label “feel” contains the feeling score for each student when the social network data were collected, and it contains the following columns: the column “id” is the identifies for the students, and the column “feelYYMM” contains the feeling scores for students when the social network data was collected on month MM in year 20YY. The feeling score is a rough measurement of how they feel in the school ranging on an integer scale from very bad (1) to very good (5). Furthermore, a feeling score of 7 means that the corresponding student was not in the class on the day when data were collected; while a score of 9 means that the student was present in the class but didn't give a feeling score. A brief summary of each social network data can be found in the file named “Brief_Summary.xlsx”. It summarises for each social network, the number of students, the number of friend-nominating ties, the number of enemy-nominating ties, the degree (i.e. the sum of the number of friend-nominating ties and the number of enemy-nominating ties), and their means.

Abbreviations

PII: Political Independence Index; PN: Positive-Negative centrality; TE: Total Effect; NE: Net Effect; ID: Identification; CIs: Confidence intervals; RBO: Rank-biased overlap; Deg: Degree centrality; Btw: Betweenness centrality; CIs: Closeness centrality; Eig: Eigenvector centrality; Katz: Katz centrality; TE_SAM: Total effect for Sampson's dataset; NE_SAM: Net effect for Sampson's dataset; TE_CLASS: Total effect for classroom dataset; NE_CLASS: Net effect for classroom dataset; ASCEND: Analysis, Structure, Characteristics, Experimentation of Network and Databases; ROMUL: Romul; BONAVEN: Bonaventure; AMBROSE: Ambrose; BERTH: Berthold; PETER: Peter; LOUIS: Louis; VICTOR: Victor; WINF: Winfrid; JOHN: John; GREG: Gregory; HUGH: Hugh; BONI: Boniface; MARK: Mark; ALBERT: Albert; AMAND: Amand; BASIL: Basil; ELIAS: Elias; SIMP: Simplicius

Acknowledgements

We are grateful for all Hungarian students who participated in the survey that was supported by Society in Science: the Branco Weiss Fellowship, ETH, Zürich, Switzerland. WCL acknowledges financial support from Thematic Research Program (AS-TP-109-M07) funded by Academia Sinica, Taiwan, project title: Analysis, Structure, Characteristics, Experimentation of Network and Databases (ASCEND).

Authors' contributions

WCL and FJ developed the methodology used in this paper. LCH and CWJL applied the methodology to analyse the data. All authors took part in writing up this paper. The authors read and approved the final manuscript.

Funding

FJ was supported by Society in Science: the Branco Weiss Fellowship, ETH, Zürich, Switzerland. WCL was supported by Thematic Research Program (AS-TP-109-M07) Academia Sinica, Taiwan, project title: Analysis, Structure, Characteristics, Experimentation of Network and Databases (ASCEND).

Availability of data and materials

R codes for our methodology and dataset are available from the corresponding author on request. They are also available for download from GitHub (<https://github.com/WeiChungLiu/SimpleCentralityForSignedDigraph.git>).

Competing interests

The authors declare that they have no competing interests.

Author details

¹Institute of Statistical Science, Academia Sinica, Taipei, Taiwan. ²Department of International Business, Chang Jung Christian University, Tainan, Taiwan. ³Balaton Limnological Institute, MTA Centre for Ecological Research, Klebelsberg Kuno, Tihany 3, 8237, Hungary.

Received: 10 March 2020 Accepted: 17 July 2020

Published online: 05 August 2020

References

- Ahuja MK, Galletta DF, Carley KM (2003) Individual centrality and performance in virtual R&D groups: an empirical study. *Manag Sci* 49(1):21–38
- Albert R, Jeong H, Barabási AL (2000) Error and attack tolerance of complex networks. *Nature* 406(6794):378–382
- Antal T, Krapivsky PL, Redner S (2006) Social balance on networks: the dynamics of friendship and enmity. *Physica D* 224:130–136
- Bagozzi RP, Wong N, Yi Y (1999) The role of culture and gender in the relationship between positive and negative affect. *Cognit Emot* 13(6):641–672
- Baumeister RF, Bratslavsky E, Finkenauer C, Vohs KD (2001) Bad is stronger than good. *Rev Gen Psychol* 5:323–370
- Benzi M, Estrada E, Klymko C (2013) Ranking hubs and authorities using matrix functions. *Linear Algebra Appl* 438(5):2447–2474
- Bonacich P (1987) Power and centrality: a family of measures. *Am J Sociol* 92(5):1170–1182
- Bonacich P, Lloyd P (2004) Calculating status with negative relations. *Soc Networks* 26(4):331–338
- Borgatti SP, Everett MG (1993) Two algorithms for computing regular equivalence. *Soc Networks* 15(4):361–376
- Borgatti SP, Everett MG, Freeman LC (2002) Ucinet for windows: software for social network analysis. Analytic Technologies, Harvard
- Callaway DS, Newman MEJ, Strogatz SH, Watts DJ (2000) Network robustness and fragility: percolation on random graphs. *Phys Rev Lett* 85(25):5468–5471
- Campbell C, Thakar J, Albert R (2011) Network analysis reveals cross-links of the immune pathways activated by bacteria and allergen. *Phys Rev E* 84:031929
- Carli L (2001) Gender and social influence. *Aust J Soc Issues* 57:725–741
- Cartwright D, Harary F (1956) Structural balance: a generalization of Heider's theory. *Psychol Rev* 63(5):277–293
- Chan T-C, Yen T-J, Hu T-H, Fu Y-C, Hwang J-S (2018) Detecting concurrent mood in daily contact networks: an online participatory cohort study with a diary approach. *BMJ Open* 8:e020600
- Chen HW, Liu WC, Davis AJ, Jordan F, Hwang MJ, Shao KT (2008) Network position of hosts in food webs and their parasite diversity. *OIKOS* 117(12):1847–1855
- Cheng S-I, Lin W-h, Phoa FKH, J-s H, W-c L (2015) Analysing the unequal effects of positive and negative information on the behaviour of users of a Taiwanese on-line bulletin board. *PLoS One* 10(9):e0137842
- Cherifi H, Palla G, Szymanski BK, Lu X (2019) On community structure in complex networks: challenges and opportunities. *Appl Network Sci* 4:117
- de Silva S, Ranjeewa ADG, Kryazhimskiy S (2011) The dynamics of social networks among female Asian elephants. *BMC Ecol* 11:17
- Dong G, Zhou H, Zhao X, Lu Q (2011) Early negativity bias occurring prior to experiencing of emotion: an ERP study. *J Psychophysiol* 25:9–17
- Du H, He X, Feldman MW (2016) Structural balance in fully signed networks. *Complexity* 21(S1):497–511
- Dunbar RIM (2018) The anatomy of friendship. *Trends Cogn Sci* 22(1):32–51
- Dunne JA, Williams RJ, Martinez ND (2002) Network structure and biodiversity loss in food webs: robustness increases with connectance. *Ecol Lett* 5(4):558–567
- Eagly AH (1983) Gender and social influence: a social psychological analysis. *Am Psychol* 38(9):971–981
- El-Khatib R, Fogel K, Jandik T (2015) CEO network centrality and merger performance. *J Financ Econ* 116(2):349–382
- Esmailian P, Jalili M (2015) Community detection in signed networks: the role of negative ties in different scales. *Sci Rep* 5:14339
- Estrada E (2007) Food webs robustness to biodiversity loss: the roles of connectance, expansibility and degree distribution. *J Theor Biol* 244(2):296–307
- Estrada E (2012) The structure of complex networks: theory and applications. Oxford University Press, Oxford
- Everett MG, Borgatti SP (2014) Networks containing negative ties. *Soc Networks* 38(1):111–120
- Fortunato S, Hric D (2016) Community detection in networks: a user guide. *Phys Rep* 659:1–44
- Fowler JH, Christakis NA (2008) Dynamic spread of happiness in a large social network: longitudinal analysis over 20 years in the Framingham heart study. *BMJ* 337:a2338
- Fujita F, Diener E, Sandvik E (1991) Gender differences in negative affect and well-being: the case for emotional intensity. *J Pers Soc Psychol* 61(3):427–434
- Ghalmene Z, Cherifi C, Cherifi H, Hassouni ME (2019a) Centrality in complex networks with overlapping community structure. *Sci Rep* 9:10133
- Ghalmene Z, Hassouni ME, Cherifi C, Cherifi H (2019b) Centrality in modular networks. *EJP Data Science* 8:15
- Girvan M, Newman MEJ (2002) Community structure in social and biological networks. *PNAS* 99(12):7821–7826
- He K, Li Y, Soundarajan S, Hopcroft JE (2018) Hidden community detection in social networks. *Inform Sci* 425:92–106
- He X, Du H, Feldman MW, Li G (2019) Information diffusion in signed networks. *PLoS One* 14(10):e0224177
- Hubbell CH (1965) An input–output approach to clique identification. *Sociometry* 28(4):377–399
- Jeong H, Mason SP, Barabási AL, Oltvai ZN (2001) Lethality and centrality in protein networks. *Nature* 411(6833):41–42
- Jordán F (2009) Children in time: community organisation in social and ecological systems. *Curr Sci India* 97(11):1579–1585
- Jordán F, Liu W-C, Davis AJ (2006) Topological keystone species: measures of positional importance in food webs. *Oikos* 112(3):535–546
- Jordán F, Liu W-C, van Veen J (2003) Quantifying the importance of species and their interactions in a host-parasitoid community. *Community Ecol* 4(1):79–88
- Kats L (1953) A new status index derived from sociometric analysis. *Psychometrika* 18(1):39–43

- Lerne J (2016) Structural balance in signed networks: separating the probability to interact from the tendency to fight. *Soc Networks* 45:66–77
- Li Y, Chen W, Wang Y, Zhang ZL (2015) Voter model on signed social networks. *Internet Math* 11(2):93–133
- Liu W-C, Chen H-W, Jordán F, Lin W-H, Liu CW-J (2010) Quantifying the interaction structure and the topological importance of species in food webs: a signed digraph approach. *J Theor Biol* 267(3):355–362
- Liu WC, Lin WH, Davis AJ, Jordan F, Yang HT, Hwang MJ (2007) A network perspective on the topological importance of enzymes and their phylogenetic conservation. *BMC Bioinformatics* 8:121
- Müller CB, Adriaanse ICT, Belshaw R, Godfray HCJ (1999) The structure of an aphid–parasitoid community. *J Anim Ecol* 68(2): 346–370
- Oldham S, Fulcher B, Parkes L, Arnatkeviciute A, Suo C, Fornito A (2019) Consistency and differences between centrality measures across distinct classes of networks. *PLoS One* 14(7):e0220061
- Sampson SP (1968) A novitiate in a period of change: an experimental and case study of social relationships Dissertation, Cornell University
- Smith JM, Halgin DS, Kidwell-Lopez V, Labianca G, Brass DJ, Borgatti SP (2014) Power in politically charged networks. *Soc Networks* 36(1):162–176
- Stephenson K, Zelen M (1989) Rethinking centrality: methods and examples. *Soc Networks* 11(1):1–37
- Traag VA, Bruggeman J (2009) Community detection in networks with positive and negative links. *Phys Rev E* 80:036115
- Wasserman S, Faust K (1994) *Social network analysis: methods and applications*. Cambridge University Press, Cambridge
- Webber W, Moffat A, Zobel J (2010) A similarity measure for indefinite rankings. *ACM Trans Inf Syst* 28(4):20
- Wey T, Blumstein DT, Shen W, Jordán F (2008) Social network analysis of animal behaviour: a promising tool for the study of sociality. *Anim Behav* 75(2):333–344
- White D, Reitz K (1983) Graph and semigroup homomorphisms. *Soc Networks* 5(2):193–234

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► [springeropen.com](https://www.springeropen.com)