Edge Centrality via the Holevo Quantity

Joshua Lockhart[1], Giorgia Minello[2], Luca Rossi[3($\boxtimes$)], Simone Severini[1], and Andrea Torsello[2]

[1] Department of Computer Science, University College London, London, UK
[2] DAIS, Università Ca' Foscari Venezia, Venice, Italy
[3] School of Engineering and Applied Science, Aston University, Birmingham, UK
l.rossi@aston.ac.uk

Abstract. In the study of complex networks, vertex centrality measures are used to identify the most important vertices within a graph. A related problem is that of measuring the centrality of an edge. In this paper, we propose a novel edge centrality index rooted in quantum information. More specifically, we measure the importance of an edge in terms of the contribution that it gives to the Von Neumann entropy of the graph. We show that this can be computed in terms of the Holevo quantity, a well known quantum information theoretical measure. While computing the Von Neumann entropy and hence the Holevo quantity requires computing the spectrum of the graph Laplacian, we show how to obtain a simplified measure through a quadratic approximation of the Shannon entropy. This in turns shows that the proposed centrality measure is strongly correlated with the negative degree centrality on the line graph. We evaluate our centrality measure through an extensive set of experiments on real-world as well as synthetic networks, and we compare it against commonly used alternative measures.

Keywords: Edge centrality · Complex networks · Holevo quantity · Quantum information

1 Introduction

The study of complex networks has recently attracted increasing interest in the scientific community, as it allows to model and understand a large number of real-world systems [4]. This is particularly relevant given the growing amount of available data describing the interactions and dynamics of real-world systems. Typical examples of complex networks include metabolic networks [8], protein interactions [7], brain networks [17] and scientific collaboration networks [11].

One of the key problems in network science is that of identifying the most relevant nodes in a network. This importance measure is usually called the *centrality* of a vertex [9]. A number of centrality indices have been introduced in the literature [2,4–6,10,14], each of them capturing different but equally significant aspects of vertex importance. Commonly encountered examples are the degree, closeness and betweenness centrality [5,6,10]. A closely related problem

is that of measuring the centrality of an edge [3,9]. Most edge centrality indices are developed as a variant of vertex centrality ones. A common way to define an edge centrality index is to apply the corresponding vertex centrality to the line graph of the network being studied. Recall that, given a graph $G = (V, E)$, the line graph $\mathcal{L}(G) = (V', E')$ is a dual representation of G where each node $uv \in V'$ corresponds to an edge $(u, v) \in E$, and there exists and edge between two nodes of $\mathcal{L}(G)$ if and only if the corresponding edges of G share a vertex. By measuring the vertex centrality on $\mathcal{L}(G)$, one can map it back to the edges of G to obtain a measure of edge centrality. However, as observed by Koschützki et al. [9], this approach does not yield the same result as the direct definition of the edge centrality on G. Moreover, the size of the line graph is quadratic in the size of the original graph, thus making it hard to scale to large networks when the chosen centrality measure is computationally demanding.

In this paper, we introduce a novel edge centrality measure rooted in quantum information theory. More specifically, we propose to measure the importance of an edge in terms of its contribution to the Von Neumann entropy of the network [13]. This can be measured in terms of the Holevo quantity, a well known quantum information theoretical measure that has recently been applied to the analysis of graph structure [15,16]. We also show how to approximate this quantity in the case of large networks, where computing the exact value of the Von Neumann entropy is not feasible. This in turns highlights a strong connection between the Holevo edge centrality and the negative degree centrality on the line graph. Finally, we perform a series of experiments to evaluate the proposed edge centrality measure on real-world as well as synthetic graphs, and we compare it against a number of widely used alternative measures.

The remainder of the paper is organised as follows: Sect. 2 reviews the necessary quantum information theoretical background and Sect. 3 introduces the proposed edge centrality measure. The experimental evaluation is presented in Sect. 4 and Sect. 5 concludes the paper.

2   Quantum Information Theoretical Background

2.1   Quantum States and Von Neumann Entropy

In quantum mechanics, a system can be either in a pure state or a mixed state. Using the Dirac notation, a *pure state* is represented as a column vector $|\psi_i\rangle$. A *mixed state*, on the other hand, is an ensemble of pure quantum states $|\psi_i\rangle$, each with probability p_i. The density operator of such a system is a positive unit-trace matrix defined as

$$\rho = \sum_i p_i |\psi_i\rangle \langle\psi_i|. \tag{1}$$

The *Von Neumann entropy* [12] S of a mixed state is defined in terms of the trace and logarithm of the density operator ρ

$$S(\rho) = -\operatorname{tr}(\rho \ln \rho) = -\sum_i \lambda_i \ln(\lambda_i) \tag{2}$$

where $\lambda_1, \ldots, \lambda_n$ are the eigenvalues of ρ. If $\langle\psi_i|\rho|\psi_i\rangle = 1$, i.e., the quantum system is a pure state $|\psi_i\rangle$ with probability $p_i = 1$, then the Von Neumann entropy $S(\rho) = -\operatorname{tr}(\rho\ln\rho)$ is zero. On other hand, a mixed state always has a non-zero Von Neumann entropy associated with it.

2.2   A Mixed State from the Graph Laplacian

Let $G = (V, E)$ be a simple graph with n vertices and m edges. We assign the vertices of G to the elements of the standard basis of an Hilbert space $\mathcal{H}_G$, $\{|1\rangle, |2\rangle, ..., |n\rangle\}$. Here $|i\rangle$ denotes a column vector where 1 is at the i-th position. The *graph Laplacian* of G is the matrix $L = D - A$, where A is the adjacency matrix of G and D is the diagonal matrix with elements $d(u) = \sum_{v=1}^{n} A(u, v)$. For each edge $e_{i,j}$, we define a pure state

$$|e_{i,j}\rangle := \frac{1}{\sqrt{2}}(|i\rangle - |j\rangle). \tag{3}$$

Then we can define the mixed state $\{\frac{1}{m}, |e_{i,j}\rangle\}$ with density matrix

$$\rho(G) := \frac{1}{m}\sum_{\{i,j\}\in E} |e_{i,j}\rangle\langle e_{i,j}| = \frac{1}{2m}L(G). \tag{4}$$

Let us define the Hilbert spaces $\mathcal{H}_V \cong \mathbb{C}^V$, with orthonormal basis $\mathbf{a}_v$, where $v \in V$, and $\mathcal{H}_E \cong \mathbb{C}^E$, with orthonormal basis $\mathbf{b}_{u,v}$, where $\{u, v\} \in E$. It can be shown that the graph Laplacian corresponds to the partial trace of a rank-1 operator on $\mathcal{H}_V \otimes \mathcal{H}_E$ which is determined by the graph structure [1]. As a consequence, the Von Neumann entropy of $\rho(G)$ can be interpreted as a measure of the amount of entanglement between a system corresponding to the vertices and a system corresponding to the edges of the graph [1].

2.3   Holevo Quantity of a Graph Decomposition

Given a graph G, we can define an ensemble in terms of its subgraphs. Recall that a *decomposition* of a graph G is a set of subgraphs $H_1, H_2, ..., H_k$ that partition the edges of G, i.e., for all i, j, $\bigcup_{i=1}^{k} H_i = G$ and $E(H_i) \cap E(H_j) = \emptyset$, where $E(G)$ denotes the edge set of G. Notice that isolated vertices do not contribute to a decomposition, so each H_i can always be seen a subgraph that contains all the vertices. If we let $\rho(H_1), \rho(H_2), ..., \rho(H_k)$ be the mixed states of the subgraphs, the probability of H_i in the mixture $\rho(G)$ is given by $|E(H_i)|/|E(G)|$. Thus, we can generalise Eq. 4 and write

$$\rho(G) = \sum_{i=1}^{k} \frac{|E(H_i)|}{|E(G)|}\rho(H_i). \tag{5}$$

Consider a graph G and its decomposition $H_1, H_2, ..., H_k$ with corresponding states $\rho(H_1), \rho(H_2), ..., \rho(H_k)$. Let us assign $\rho(H_1), \rho(H_2), ..., \rho(H_k)$ to the elements of an alphabet $\{a_1, a_2, ..., a_k\}$. In quantum information theory, the classical

concepts of uncertainty and entropy are extended to deal with quantum states, where uncertainty about the state of a quantum system can be expressed using the density matrix formalism. Assume a source emits letters from the alphabet and that the letter a_i is emitted with probability $p_i = |E(H_i)|/|E(G)|$. An upper bound to the accessible information is given by the *Holevo quantity* of the ensemble $\{p_i, \rho(H_i)\}$:

$$\chi(\{p_i, \rho(H_i)\}) = S\left(\sum_{i=1}^{k} p_i \rho(H_i)\right) - \sum_{i=1}^{k} p_i S(\rho(H_i)) \qquad (6)$$

3 Holevo Edge Centrality

We propose to measure the centrality of an edge as follows. Let $G = (V, E)$ be a graph with $|E| = m$, and let H_e and $H_{\bar{e}}$ denote the subgraphs over edge sets $\{e\}$ and $E \setminus \{e\}$, respectively. Note that $S(\rho(H_e)) = 0$ and

$$\frac{m-1}{m}\rho(H_{\bar{e}}) + \frac{1}{m}\rho(H_e) = \rho(G). \qquad (7)$$

Then the Holevo quantity of the ensemble $\{(m - 1/m, H_{\bar{e}}), (1/m, H_e)\}$ is

$$\chi\left(\left\{\left(\frac{m-1}{m}, H_{\bar{e}}\right), \left(\frac{1}{m}, H_e\right)\right\}\right) = S\left(\rho(G)\right) - \frac{m-1}{m}S\left(\rho(H_{\bar{e}})\right) \qquad (8)$$

Definition 1. *For a graph $G = (V, E)$, the* Holevo edge centrality *of $e \in E$ is*

$$HC(e) = \chi\left(\left\{\left(\frac{m-1}{m}, H_{\bar{e}}\right), \left(\frac{1}{m}, H_e\right)\right\}\right) \qquad (9)$$

When ranking the edges of a graph G, the scaling factor $(m - 1)/m$ is constant for all the edges and thus can be safely ignored. The Holevo edge centrality of an edge e is then a measure of the difference in Von Neumann entropy between the original graph and the graph where e has been removed. In other words, it can be seen as a measure of the contribution of e to the Von Neumann entropy of G. From a physical perspective, this can also be interpreted as the variation of the entanglement between between a system corresponding to the vertices and a system corresponding to the edges of the graph (see the interpretation of the graph Laplacian in Sect. 2).

3.1 Relation with Degree Centrality

In this subsection we investigate the nature of the structural characteristics encapsulated by the Holevo edge centrality. Let $G = (V, E)$ be a graph with n nodes, and let I_n be the identity matrix of size n. We rewrite the Shannon entropy $-\sum_i \lambda_i \ln(\lambda_i)$ using the second order polynomial approximation $k \sum_i \lambda_i(1 - \lambda_i)$, where the value of k depends on the dimension of the simplex. We obtain

$$S(\rho(G)) = -\operatorname{tr}(\rho(G)\ln\rho(G)) \approx \frac{|V|\ln(|V|)}{|V|-1}\operatorname{tr}(\rho(G)(I_n - \rho(G))) \qquad (10)$$

By noting that $\rho(G) = L(G)/(2m)$ and using some simple algebra, we can rewrite Eq. 10 as

$$S(\rho(G)) \approx \frac{|V|\ln(|V|)}{|V|-1}\left(1 - \frac{1}{4m^2}\sum_{v\in V}\left(d^2(v) + d(v)\right)\right) \tag{11}$$

where $d(v)$ denotes the degree of the vertex v. This in turn allows us to approximate Eq. 9 as

$$HC(e) = S(\rho(G)) - S(\rho(H_{\bar{e}})) \approx -\frac{|V|\ln(|V|)}{|V|-1}\frac{d(u) + d(w)}{2m^2} \tag{12}$$

where $e = (u, w)$, we omitted the scaling factor $(m-1)/m$ and we made use of the fact that $1/(4m^2) \approx 1/(4(m-1)^2)$.

Equation 12 shows that the quadratic approximation of the Holevo centrality is (almost) linearly correlated with the negative edge degree centrality (see Sect. 4). This in turn gives us an important insight into the nature of the Holevo edge centrality. However, the quadratic approximation captures only part of the structural information encapsulated by the exact centrality measure. In particular, Passerini and Severini [13] suggested that those edges that create longer paths, nontrivial symmetries and connected components result in a larger increase of the Von Neumann entropy. Therefore, such edges should have a high centrality value, higher than what the degree information alone would suggest.

Figure 1 shows an example of such a graph, where the central bridge has a high value of the exact Holevo edge centrality, but a relatively low value of the approximated edge centrality. In Fig. 1(b), the blue edges have all the same degree centrality, i.e., they are all adjacent to four other edges. However, from a structural point of view, the removal of the edges connecting the two cliques at the ends of the barbell graph would have a higher impact, as it would disconnect the graph. As shown in Fig. 1(a), the Holevo centrality captures this structural difference, i.e., the weight assigned to the two bridges (blue) is higher than that assigned to the edges in the cliques (red).
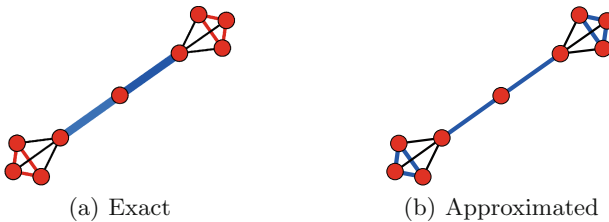

(a) Exact   (b) Approximated

Fig. 1. The Holevo edge centrality and its quadratic approximation on a barbell graph. Here the edge thickness is proportional to the value of the centrality. In (a) the blue edges have a higher centrality than the red edges, but in (b) all these edges (blue) have the same degree centrality. (Color figure online)

4   Experimental Evaluation

In the previous sections we have derived an expression for the Holevo edge centrality, both exact and approximated. Here, we first evaluate this measure on a number of standard networks, and we compare it against other well known edge centralities. We also analyse the behaviour of the proposed centrality measure when graphs endure structural changes.

4.1   Experimental Setup

We perform our experiments on two well known real-world networks, the Florentine families graph and the Karate club network. We then consider the following edge centrality measures:

Degree Centrality: The centrality of an edge e is computed as the degree of the corresponding vertex in the line graph. The idea underpinning the vertex degree centrality is that the importance of a node is proportional to the number of connections it has to other nodes. This is the simplest edge centrality measure, but also the one with the lowest computational complexity.

Betweenness Centrality: The centrality of an edge e is the sum of the fraction of all-pairs shortest paths that pass through e, i.e., $EBC(e) = \sum_{u,v \in V} \frac{\sigma(u,v|e)}{\sigma(u,v)}$ where V is the set of nodes, $\sigma(u,v)$ and $\sigma(u,v|e)$ denote the number of shortest paths between u and v and the number of shortest paths between u and v that pass through e, respectively [3]. An edge with a high betweenness centrality has a large influence on the transfer of information through the network and thus it can be seen as an important bridge-like connector between two parts of a network. Note that the implementation we use does not rely on the line graph, but measure the centrality of an edge directly on the original graph.

Flow Centrality: This centrality measure is also known as random-walk betweenness centrality [10]. While the betweenness centrality measures the importance of an edge e in terms of shortest-paths between pairs of nodes that pass through e, the flow centrality is proportional to the expected number of times a random walk passes through the edge e when going from u to v. Similarly to the betweenness centrality, here we measure the flow centrality directly on the original graph.

4.2   Edge Centrality in Real-World Networks

In order to compare the Holevo edge centrality with the measures described in the previous subsection, we compute, for each network, the correlation between the Holevo quantity and the alternative measures. Figure 2 shows the value of these centralities on the Florentine families graph and the Karate club network.
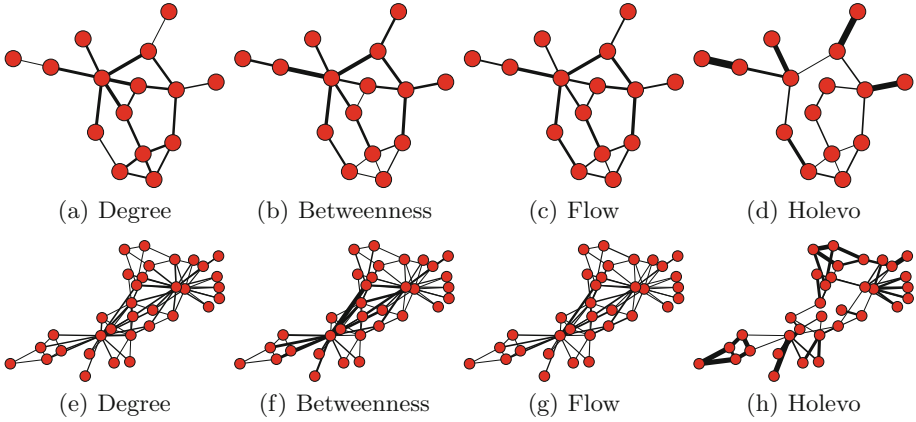
Fig. 2. Edge centralities on the Florentine families network (a–d) and the Karate club network (e–h). A thicker edge indicates a higher value of the centrality.
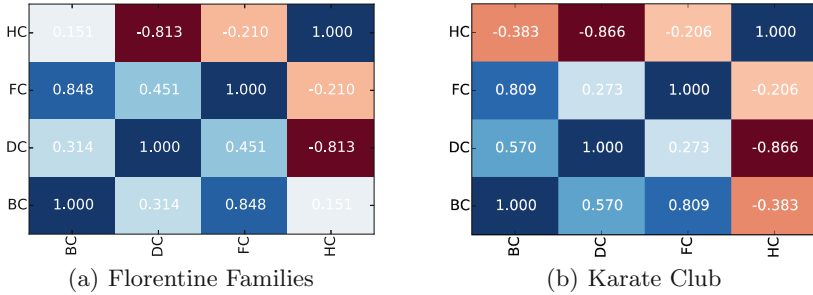

Fig. 3. Correlation matrices for the centrality measure on the Florentine family network and the karate club network. DC, BC, FC, and HC denote the degree, betweenness, flow and Holevo centralities, respectively.

In these plots, the thickness of an edge is proportional to the magnitude of the centrality index. Figure 3, on the other hand, shows the correlation matrix between the different centralities. Here DC, BC, FC and HC denote the degree, betweenness, flow and Holevo centrality, respectively.

The Holevo centrality is always strongly negatively correlated with the degree centrality. This is in accordance with the properties discussed in Sect. 3. However, there are some significant differences. In general, the Holevo centrality is higher on edges that connect low degree nodes. In this sense, it can be seen as a measure of *peripherality*, rather than centrality. However, when two edges have the same degree centrality, edges that would disconnect the network or break structural symmetries are assigned a higher weight, as Fig. 1 shows. Similarly, in Fig. 4(a) the three edges highlighted in blue have the same degree centrality, but the same edges in Fig. 4(b) have different Holevo centralities. In fact, the removal of the red edge does not result in significant structural changes, while the removal of one of the blue edges increases the length of the tail.
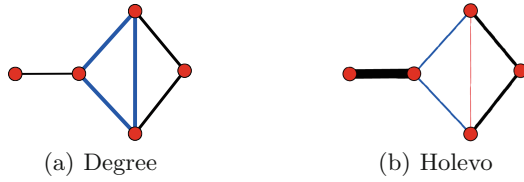
(a) Degree   (b) Holevo

Fig. 4. Toy example showing the difference in the structural information captured by the degree and Holevo centralities. (Color figure online)

4.3   Robustness Analysis

We then investigate the behaviour of the Holevo edge centrality when the graph undergoes structural perturbations. To this end, given an initial graph, we gradually add or delete edges according to an increasing probability p. Figure 5 shows an instance of the noise addition process. Starting from a randomly generated graph, we compute the Holevo edge centrality for all its edges. Then, we perturb the graph structure with a given probability p and again we recompute the Holevo edge centrality for all the graph edges. We compute the correlation between the Holevo centrality of the edges of the original graph and its noisy counterpart. More specifically, we measure the correlation between the centralities of the edges that belong to the intersection of their edge sets. In other words, we analyse how the centrality changes during the perturbation process, with respect to the starting state.

Since we are interested in the variation of the Holevo centrality as the graph structure changes, we use three different random graph models to generate the initial graph: (1) the Erdös-Rényi model, (2) the Watts-Strogatz model and
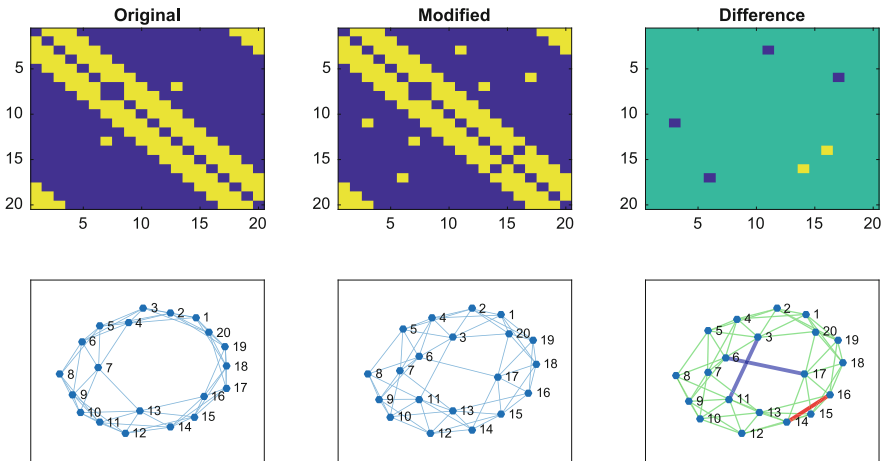

Fig. 5. Perturbation process: on the left, adjacency matrix and plot of the starting graph; in the middle, the edited graph; on the right, the differences between initial and modified graph are highlighted. (Color figure online)
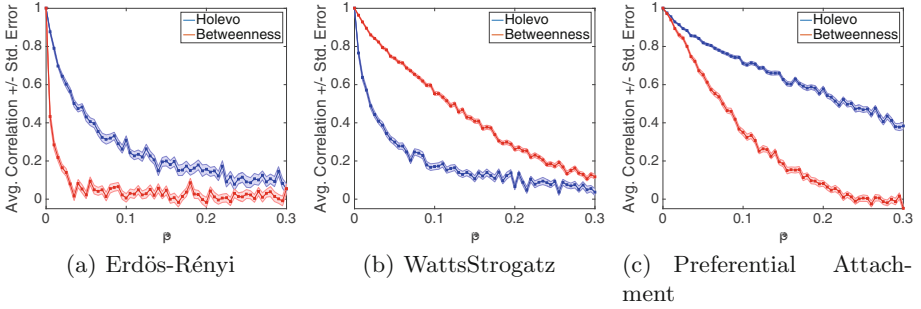
Fig. 6. Average correlation between the centrality of the edges of the original graph and those of increasingly noisy version of it. The different columns refer to different starting graphs: (a) Erdös-Rényi, (b) WattsStrogatz and (c) Preferential Attachment.

(3) the Preferential Attachment model. For each model, we generate a starting graph with the same number of nodes n and we create 100 noisy instances as p varies from 0.01 to 0.3. We perform the same experiment for both the Holevo centrality and the betweenness centrality.

Figure 6 shows the average correlation as we perturb the graph structure, for both the Holevo and betweenness centrality. As expected, in both cases the correlation decreases as the similarity between the original graph and the edited one decreases. However, while the correlation for centrality measures decreases rapidly in the case of ErdösRényi graphs, on scale-free graphs our centrality measure decreases linearly with the value of p, while the betweenness centrality drops significantly more quickly. On the other hand, we observe the opposite behaviour on small-world graphs. This can be explained by noting that in small-world graphs there exist multiple alternative paths between every pair of nodes, and thus the betweenness centrality is less affected by structural modifications. On the other hand, in scale-free graphs most shortest-paths pass through a hub, and thus adding a random edge can create shortcuts that greatly affect the value of the betweenness centrality. The Holevo centrality, however, assigns large weights to long tails and leaves, which are less affected by the structural noise.

5   Conclusion

In this paper we have introduced a novel edge centrality measure based on the quantum information theoretical concept of Holevo quantity. We measured the importance of an edge in terms of the difference in Von Neumann entropy between the original graph and the graph where that edge has been remove. We showed that by taking a quadratic approximation of the Von Neumann entropy we obtain an approximated value of the Holevo centrality that is proportional to the negative degree centrality. We performed a series of experiments on both real-world and synthetic networks and we compared the proposed centrality measure to widely used alternatives. Future work will investigate higher order

approximations of this centrality measure as well as the possibility of defining network growth models based on the Holevo quantity.

References

1. de Beaudrap, N., Giovannetti, V., Severini, S., Wilson, R.: Interpreting the Von Neumann entropy of graph laplacians, and coentropic graphs. Panor. Math. Pure Appl. **658**, 227 (2016)
2. Bonacich, P.: Power and centrality: a family of measures. Am. J. Sociol. **92**, 1170–1182 (1987)
3. Brandes, U.: On variants of shortest-path betweenness centrality and their generic computation. Soc. Netw. **30**(2), 136–145 (2008)
4. Estrada, E.: The Structure of Complex Networks. Oxford University Press, Oxford (2011)
5. Freeman, L.C.: A set of measures of centrality based on betweenness. Sociometry **40**, 35–41 (1977)
6. Freeman, L.C.: Centrality in social networks conceptual clarification. Soc. Netw. **1**(3), 215–239 (1979)
7. Ito, T., Chiba, T., Ozawa, R., Yoshida, M., Hattori, M., Sakaki, Y.: A comprehensive two-hybrid analysis to explore the yeast protein interactome. Proc. Natl. Acad. Sci. **98**(8), 4569 (2001)
8. Jeong, H., Tombor, B., Albert, R., Oltvai, Z., Barabási, A.: The large-scale organization of metabolic networks. Nature **407**(6804), 651–654 (2000)
9. Koschützki, D., Lehmann, K.A., Peeters, L., Richter, S., Tenfelde-Podehl, D., Zlotowski, O.: Centrality indices. In: Brandes, U., Erlebach, T. (eds.) Network Analysis. LNCS, vol. 3418, pp. 16–61. Springer, Heidelberg (2005). doi:10.1007/978-3-540-31955-9_3
10. Newman, M.E.: A measure of betweenness centrality based on random walks. Soc. Netw. **27**(1), 39–54 (2005)
11. Newman, M.: Scientific collaboration networks. I. Network construction and fundamental results. Phys. Rev. E **64**(1), 016131 (2001)
12. Nielsen, M.A., Chuang, I.L.: Quantum Computation and Quantum Information. Cambridge University Press, Cambridge (2010)
13. Passerini, F., Severini, S.: Quantifying complexity in networks: the Von Neumann entropy. Int. J. Agent Technol. Syst. (IJATS) **1**(4), 58–67 (2009)
14. Rossi, L., Torsello, A., Hancock, E.R.: Node centrality for continuous-time quantum walks. In: Fränti, P., Brown, G., Loog, M., Escolano, F., Pelillo, M. (eds.) S+SSPR 2014. LNCS, vol. 8621, pp. 103–112. Springer, Heidelberg (2014). doi:10.1007/978-3-662-44415-3_11
15. Rossi, L., Torsello, A., Hancock, E.R.: Measuring graph similarity through continuous-time quantum walks and the quantum Jensen-Shannon divergence. Phys. Rev. E **91**(2), 022815 (2015)
16. Rossi, L., Torsello, A., Hancock, E.R., Wilson, R.C.: Characterizing graph symmetries through quantum Jensen-Shannon divergence. Phys. Rev. E **88**(3), 032806 (2013)
17. Sporns, O.: Network analysis, complexity, and brain function. Complexity **8**(1), 56–60 (2002)