

SCIENTIFIC REPORTS



OPEN

Eigencentrality based on dissimilarity measures reveals central nodes in complex networks

A. J. Alvarez-Socorro^{1,2}, G. C. Herrera-Almarza^{1,2} & L. A. González-Díaz²

Received: 17 July 2015

Accepted: 21 October 2015

Published: 25 November 2015

One of the most important problems in complex network's theory is the location of the entities that are essential or have a main role within the network. For this purpose, the use of dissimilarity measures (specific to theory of classification and data mining) to enrich the centrality measures in complex networks is proposed. The centrality method used is the eigencentrality which is based on the heuristic that the centrality of a node depends on how central are the nodes in the immediate neighbourhood (like rich get richer phenomenon). This can be described by an eigenvalues problem, however the information of the neighbourhood and the connections between neighbours is not taken in account, neglecting their relevance when is one evaluates the centrality/importance/influence of a node. The contribution calculated by the dissimilarity measure is parameter independent, making the proposed method is also parameter independent. Finally, we perform a comparative study of our method versus other methods reported in the literature, obtaining more accurate and less expensive computational results in most cases.

A large number of systems from nature and others man-made can be described in terms of networks, composed by entities (or nodes) that interact through connections (or links). The topological and statistical properties of the nodes in a network (at microscopic level) tend to be highly heterogeneous, as can be seen by studying their degree distribution, clustering distribution and their degree-degree correlations¹⁻³. On the other hand, there are also heterogeneities at mesoscopic and macroscopic levels, e.g., not all the networks have the same hierarchical structure, community structure or topology⁴⁻⁸. These heterogeneities (at microscopic, mesoscopic and macroscopic levels) have repercussions in the importance of nodes and links in the network. For example, is well known that, in highly modular networks, nodes and links that connect modules or communities, i.e., nodes with neighboring nodes in different communities are more relevant (in terms of global communications) than nodes with neighborhoods fully included in the same community, and links with both extreme nodes in the same module are less relevant than links with extreme nodes in different communities, a fact that has been widely used precisely in the detection of community structures⁹. Thus, there exist nodes that are more important as a result of their position relative to other nodes of the network, giving us relevant information about the properties of the networks. This kind of nodes and links that have a special role in a network are called central with respect to a given role. Thus, one of the ways to address the problem of centrality define first (at least heuristically) the context in which we are talking about "centrality" and then build measures to quantify the definition of centrality used, as in the case of betweenness and closeness¹⁰. However, although there is no consensus on the concept of centrality (because as we mentioned it depends on the system under study and the context or heuristics behind the "centrality" to be measured), we can propose a definition of centrality that involves all existing definitions.

¹Departamento de Física, FCFM, Universidad de Chile, Santiago, Chile. ²Laboratorio de Dinámica No-Lineal y Sistemas Complejos, Centro de Física, Instituto Venezolano de Investigaciones Científicas, Caracas 1020-A, Venezuela. Correspondence and requests for materials should be addressed to A.J.A.-S. (email: alvarez.ciencia@gmail.com)

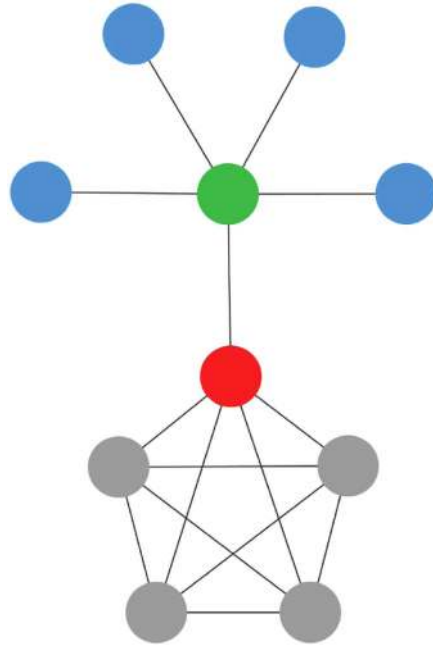


Figure 1. In the illustrated network, green and red node are dissimilar because they do not share neighbors between them. The red node reaches the blue nodes only through the green node and therefore its contribution to the centrality of red node is greater.

Definition. Let $\mathcal{G} = (\mathcal{N}, \mathcal{E})$ be a network and let $\mu : \mathcal{N} \rightarrow \mathbb{R}^+$ be a measure quantifying a desired property. We say that a node $i \in \mathcal{N}$ has a μ -centrality k if $\mu(i) = k$. So that we can talk about closeness - centrality or betweenness - centrality. The reader can find the formula for betweenness and closeness measures in the materials and methods section, equations (11) and (12).

Among the applications of measures of centrality in complex networks we have: (i) in social networks, hubs are related to the most influential people on the network¹¹⁻¹³, which is of interest to understand the individual and collective social processes and the information spreading in such networks^{13,14}, (ii) in the protein-protein interaction networks of an organism, central nodes are related to the essential genes, i.e., those genes of an organism that are critical to their survival, which is of broad interest in the research and design of drugs to combat parasitic diseases¹⁵⁻¹⁸, (iii) in air or urban traffic networks, the central nodes are associated with optimal points for the spread of diseases, which are of great interest to effectively prevent and control the spread of diseases, putting at these points, checkpoints and vaccination campaigns¹⁹⁻²¹, among many other applications²²⁻²⁴. The centrality measures are an attempt to locate these nodes, and their goal is to assign a measure (or rank) to each node so that they can be sorted from highest to lowest centrality. Some of the heuristics and statistics used to define centrality are based on¹⁰:

1. How connected a node is,
2. How influential a node is in terms of its neighbourhood,
3. How easily a node can propagate an information,
4. How intermediary is a node as a connector between nodes in the network.

In this paper, we propose a new method for finding the centrality of a node in a given network, based on both the sum of the centralities of the nodes in its neighbourhood and on their dissimilarities. The neighbourhood of the node i will contribute more to its centrality in the measure in which the nodes of the neighbourhood are more dissimilar. In this work, we will consider that two nodes are dissimilar if they do not share neighbours between them (see Fig. 1).

Let R be a network and A its adjacency matrix, i.e., $A_{ij} = 1$ if the link $\{i, j\}$ is in the network and $A_{ij} = 0$ otherwise. This matrix indexes \mathcal{N} .

One method known to find the centrality of the i -th node in a network²⁵ is based on the following heuristic: "The centrality/relevance/influence of a node is proportional to the sum of centralities of its neighbours". Mathematically, this can be written as

$$c_i = \frac{1}{\lambda} \sum_{j=1}^n A_{ij} c_j, \quad i = 1, 2, \dots, n \quad (1)$$

where c_i denotes the centrality of the i -th node and $1/\lambda$ is a constant of proportionality. This leads to the eigenvectors and eigenvalues problem:

$$A\mathbf{c} = \lambda\mathbf{c} \quad (2)$$

Assuming that $\lambda = \lambda_{\max} = \rho(A)$ is the spectral radius of the adjacency matrix of the network, i.e., the largest eigenvalue, then by the Perron-Frobenius theorem²⁶ there exists a unique nonnegative eigenvector \mathbf{c} that satisfies the above equation, obtaining the well-known measure of eigencentality²⁵. Note that to know how central or influential the i -th node is in the network, we only need to know the value of the i -th entry of the vector \mathbf{c} .

The advantages of this approach are: (i) uses local information because the centrality of a node depends explicitly on the centrality of its neighbours, (ii) uses the global information of the network through successive couplings (i.e., the centralities of the nodes in the neighbourhood of a node also depends on their neighbours, and so on), involving all network nodes in the centrality of a given node, (iii) one can analyse large networks quickly, since there are a variety of numerical methods for calculating eigenvalues and eigenvectors fairly efficiently²⁷.

The main disadvantage of this approach lies precisely in its heuristics, since it assumes that all nodes in the neighbourhood of i -th node contribute equally to its centrality, which is in general false (see Fig. 1), leading to a poor ranking, as will be discussed in the results section. Therefore, it is necessary to reformulate this heuristic.

Contribution Centrality

To illustrate the essence of our method, consider the network of Fig. 1. For the red node, the information, relevance or centrality that brings a particular gray node is poor, because the red node can access the rest of the gray nodes directly, without any intermediary (this is because they have almost the same neighbours). However, the green node is essential for red node, because without it, the red node could not access to the blue nodes. One way to quantify this is through a structural dissimilarity measure, e.g. Jaccard dissimilarity²⁸, given by

$$D_{ij} = 1 - \frac{|V^+(i) \cap V^+(j)|}{|V^+(i) \cup V^+(j)|} \quad (3)$$

that allows us to measure the difference between the neighbourhoods of two nodes i, j given. Other measures of structural and dynamical dissimilarities are discussed in the supplementary material. Thus, we can weigh the contribution made by each node j in the neighbourhood of a node i by

$$W_{ij} = A_{ij}D_{ij} \quad (4)$$

This allows us to propose the following heuristic: “the centrality of a node is proportional to the sum of the centralities of the nodes in its neighbourhood, weighed by their contributions.” Mathematically, that is

$$c_i = \frac{1}{\lambda} \sum_{j=1}^n W_{ij}c_j, \quad i = 1, 2, \dots, n \quad (5)$$

leading us to the eigenvalues - eigenvectors problem

$$W\mathbf{c} = \lambda\mathbf{c} \quad (6)$$

where $W = A \odot D$, and \odot is the coordinate to coordinate product of matrices A and D . Note that A and D are non-negative matrices, so we can use the Perron-Frobenius theorem to ensure that the above problem has a unique solution for $\lambda = \lambda_{\max}$ with \mathbf{c} non-negative, allowing us to infer the centrality of each node in the network. Thus,

$$c_i = \frac{1}{\lambda_{\max}} \sum_{j=1}^n W_{ij}c_j, \quad i = 1, 2, \dots, n \quad (7)$$

Results

Natural networks that have information about the centrality of the nodes are commonly used as benchmarks²⁹, because, to date, it has not reported a quantitative methodology to study the accuracy of centrality measures used in the detection of the most important nodes, since the concept of importance depends strongly on the system under study and the topological properties of the network considered. However, in order to have information on the runtime of our algorithm, a group of scale-free networks, Barabási - Albert type, with parameter $k=2$, was studied, taking 21 logarithmically spaced values of N , 10 networks and then averaging runtime⁵. Analysed networks from Barabási-Albert model have a number of nodes

from 10^2 to 10^5 nodes. The computation time spent for the analysis of a network of 10^5 nodes proved to be a few minutes, as can be seen in the supplementary material.

Motivated by the fact that centrality measures have been used interchangeably to locate the most important nodes in a network^{12,22,23}, although they measure different properties, we performed a comparative study between our measure and some of the most common centrality measures in different networks.

Florentine Marriages Network. This network was taken from³⁰, it was constructed through data from historical documents on the social relations among renaissance Florentine families. In³¹ is provided evidence that support why the Medici were the most powerful family in the early fifteenth century in Florence.

Applying (7) to this network, we find that the most central node is associated with the Medici family. In Fig. 2, we note that there is a difference between the ranking produced by (7) and the other centrality measures, mainly emphasizing the difference between the results obtained with our method and those obtained with eigencentrality, although both measures have similar heuristics. In Fig. 3, the network is illustrated with the different centrality values produced by (7).

Zachary's Karate Club Network. This network was taken from³² where the nodes are members of a university karate club, and the links represents the presence of ties among the members of the club. For our proposes, we take only the topology, without the weight of the links.

It is known that the most important nodes in this network are 34, 33 and 1, being 34 the president of the students' club, 33 is the vice-president, and 1 the karate sensei, hired by the club. Applying our method to this network, we obtained the same result. In this sense, the contribution centrality provides a more suitable ranking to the importance of the nodes than closeness and betweenness because in this particular network the importance of a node is not directly related with this concepts. Note also that 33 as vice president, inherits the role of 34 in his absence, so his importance in decision-making in the club and his direct connection to 34 makes him the second most important node according to our measure. In Fig. 4, we can see a clear difference between the ranking produced by the contribution centrality and other methods compared here. In Fig. 5, the network is illustrated with the different centrality values produced by (7).

Les Miserables Coappearances Network. Finally, we take the coappearances network of characters in Victor Hugo's novel "Les Miserables" taken from³³. The nodes represent characters and links that connect any pair of characters that appear in the same chapter of the book. We are staying only with the topology, ignoring the number of such coappearances.

In this network, the contribution centrality achieves again, a more suitable ranking than other centrality measures considered. Table 1 shows the first 10 nodes sorted by relevance according to different centrality measures. We found that communicability and eigenvector centrality fail to detect Valjean as the most important node. For the centralities betweenness, closeness and degree, Valjean is detected as the main node, but they are not able to find the second most important node, which is undoubtedly Javert, co-protagonist and antagonist character Valjean in the novel. The information centrality produced good results in this case, however, has misclassified to Enjolras when placed over Marius and Cosette, which does not happen with the contribution centrality. Regarding the latter, it is also noted that Cosette does not appear in the top 10 major nodes in some centralities, for example, eigenvector, degree, communicability, betweenness. Thus, the contribution centrality provides a much better ranking of the characters in the novel Les Miserables, than other methods, just taking the topological structure of the network, not the frequency of co-occurrences. In Fig. 6, the network is illustrated with the different centrality values produced by (7).

Discussion

In this paper, a general methodology that uses structural dissimilarity measures to enrich centrality methods in complex networks is proposed, illustrating this through eigencentrality method and Jaccard dissimilarity. The combination exploits local and global information network, contributing more to the centrality the most dissimilar neighbours of given node. Note the difference between the eigencentrality method and our method (see Table 1), indicating the important effect of the inclusion of dissimilarity measures as weight parameters when considering the contribution of the neighbors of a given node to its centrality. The eigencentrality method could not detect Valjean as the most important character in the novel Les Miserables, which our method achieved satisfactorily. Overall our measure of centrality behaved differently than the rest of the measures of centrality studied here, which can be clearly seen in Figs 2, 4 and 7. Although compared centrality measures measure different properties, all have been used for detection of the key nodes in the network under consideration^{12,22,23}. In this sense, we compare the results obtained with different centrality measures in each network. Note that our measure tends to adapt well to local and global topological properties of networks studied in this work.

Our method can be extended directly to weighted and directed networks, considering, e.g, the Tanimoto coefficient³⁴, since the only restriction imposed by our methodology is that the W matrix is non-negative to ensure existence and uniqueness of the non-negative eigenvector of centrality. The

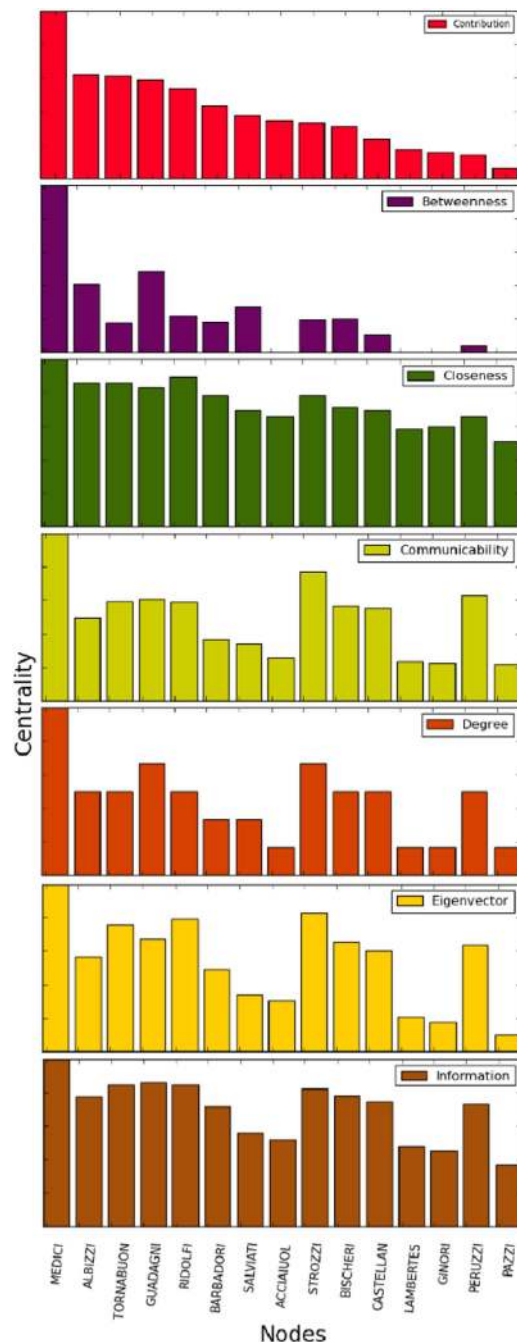


Figure 2. Distributions corresponding to the diverse measures of centrality of the Florentine marriages network.

method is quite simple and powerful. Analysis of networks of over a million nodes was achieved quickly, as can be seen in the supplementary material. Note that the coupling of the matrices A and D for the construction of W ends in some cases in a less dense matrix than A , since if two nodes have exactly the same neighbourhood, the contribution that one provides the other will be zero, eliminating the associated entry in the W matrix, accelerating the calculation of eigenvectors and reducing memory cost. The contribution centrality method can be applied to the development of recommender systems and search engines because metadata nodes (e.g. information from a web page, keywords, number of clicks, etc.) can be introduced into the dissimilarity measure, to obtain a ranking of the nodes in terms of its relevance, similar to Google PageRank³⁵. Note that the strategy of introducing metadata nodes within dissimilarity measure is very general and does not depend on the type of system under study. Finally, the contribution centrality method was applied to social, biological, transport and artificial networks, obtaining in all cases excellent results, as shown in the supplementary material.

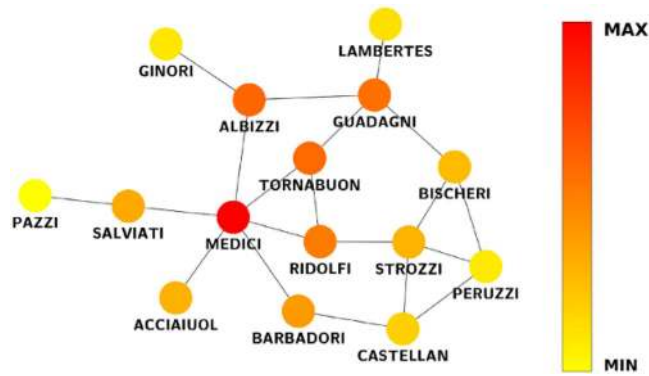


Figure 3. Florentine marriages network.

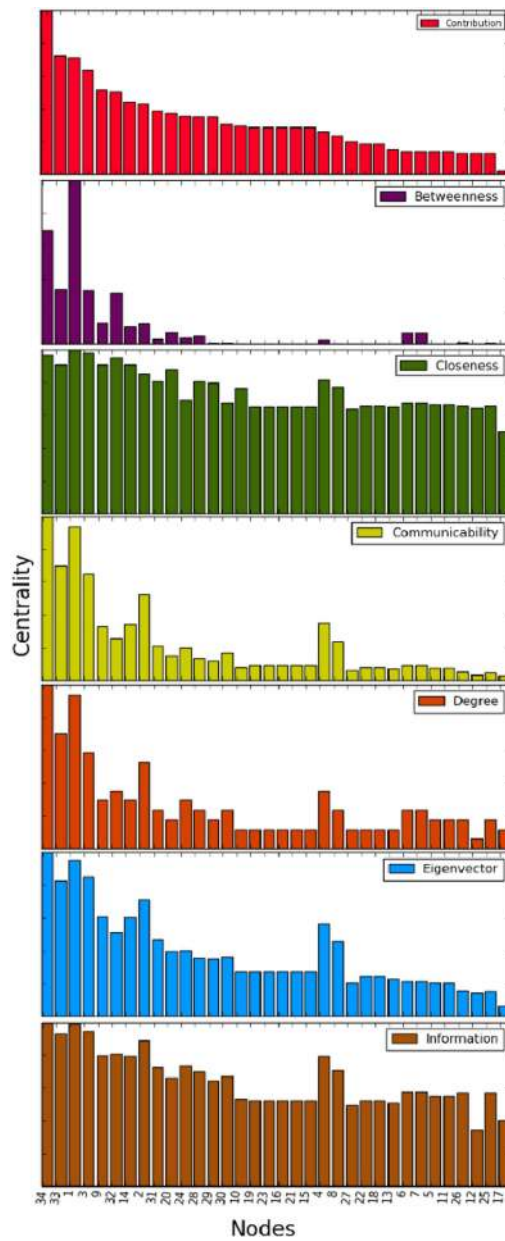


Figure 4. Distributions corresponding to the diverse centrality measures of the network of Zachary's karate club.

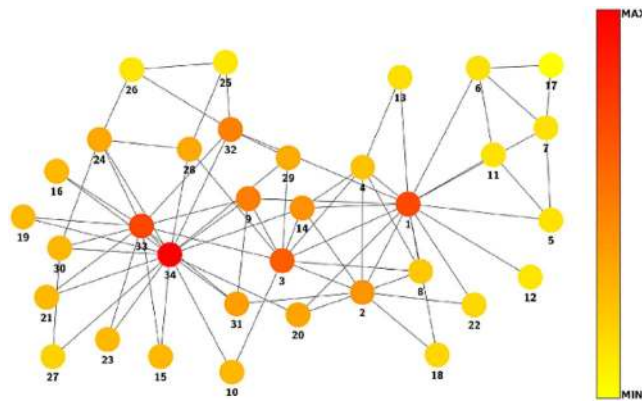


Figure 5. Zachary's karate club network.

Index	Contribution	Betweenness	Closeness	Communicability	Degree	Eienvector	Information
1	Valjean	Valjean	Valjean	Gavroche	Valjean	Gavroche	Valjean
2	Javert	Myriel	Marius	Valjean	Gavroche	Valjean	Javert
3	Gavroche	Gavroche	Thenardier	Enjolras	Marius	Enjolras	Marius
4	Thenardier	Marius	Javert	Marius	Javert	Marius	Gavroche
5	Marius	Fantine	Gavroche	Bossuet	Thenardier	Bossuet	Thenardier
6	Cosette	Thenardier	Enjolras	Courfeyrac	Fantine	Courfeyrac	Enjolras
7	Fantine	Javert	Cosette	Bahorel	Enjolras	Bahorel	Cosette
8	MmeThenardier	MlleGillenormand	Bossuet	Joly	Courfeyrac	Joly	MmeThenardier
9	Enjolras	Enjolras	Gueulemer	Combeferre	Bossuet	Feuilly	Bossuet
10	Claquesous	Tholomyes	Babet	Feuilly	Joly	Combeferre	Fantine

Table 1. First 10 nodes sorted by relevance according to different measures of centrality in the network of co-occurrences of characters in the novel by Victor Hugo's Les Miserables.

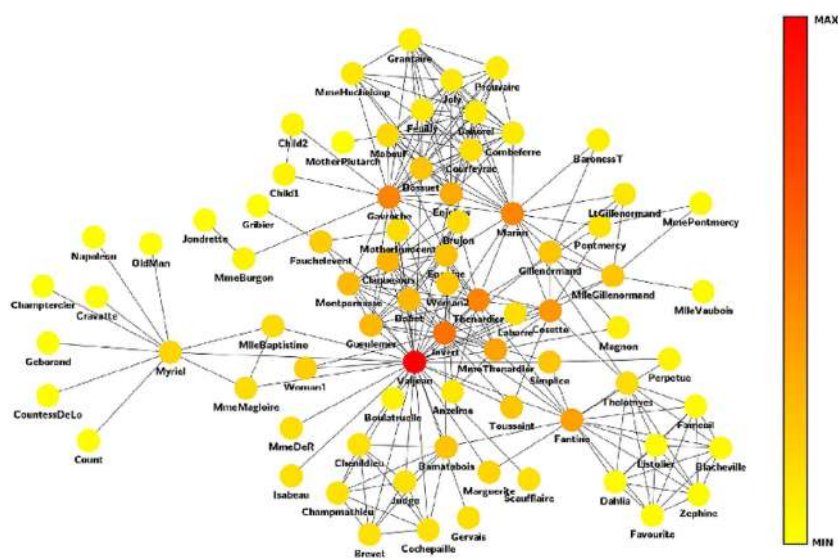


Figure 6. Les Miserables coappearances network.

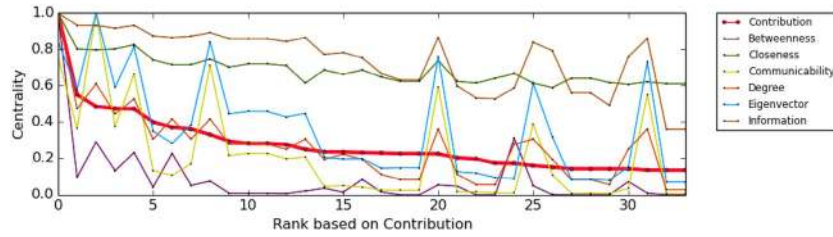


Figure 7. Distributions corresponding to the diverse measures of centrality of the Les Miserables coappearances network.

Materials and Methods

In the following, $\mathcal{G} = (\mathcal{N}, \mathcal{E})$ will denote a unweighted, undirected and connected network, with $|\mathcal{N}| = n$ nodes and $|\mathcal{E}| = m$ links and A will denote the adjacency matrix of the network \mathcal{G} .

Degree Centrality. The degree centrality³⁶ is the most simple centrality measure of all. Its heuristic is based on the idea that the most connected node is the most central. Hence, we can consider the centrality of a node i as proportional to the degree (or number of connections) of that node, i.e.,

$$C_D(i) = \frac{k_i}{n - 1} \tag{8}$$

In terms of the adjacency matrix, we can write the above equation as

$$C_D(i) = \frac{\sum_{k=1}^n A_{ik}}{n - 1} \tag{9}$$

Eigenvector Centrality. The heuristic behind the eigencentality measure²⁵ is based on the key idea that the centrality of a node is proportional to the sum of centralities of its neighbourhood and therefore, a node connected to central nodes, will also be central. As explained in the introductory section of this paper, this heuristic leads us to calculate the centrality of a node i by an equation of the form

$$c_i = \frac{1}{\lambda_{\max}} \sum_{j \in N(i)} c_j = \frac{1}{\lambda_{\max}} \sum_{j=1}^n A_{ij} c_j, \quad i = 1, 2, \dots, n, \tag{10}$$

where λ_{\max} is the largest eigenvalue of the adjacency matrix A . Note that the above equation leads directly to the eigenvector problem:

$$A\mathbf{c} = \lambda_{\max} \mathbf{c},$$

where \mathbf{c} exist and it is positive as consequence of the Perron-Frobenius theorem²⁶.

Betweenness Centrality. This is one of the most popular measures of centrality in the literature³⁷ and is based on the idea that nodes for which more information flows, will be higher values of centrality, under the assumption that the information always travels along the shortest paths connecting any two nodes in the network. Mathematically, this is

$$B(i) = \sum_{\substack{j,k \in \mathcal{N} \\ j \neq k}} \frac{\gamma_{jk}(i)}{\gamma_{jk}}, \tag{11}$$

where γ_{jk} is the number of shorter paths ranging from the node j to node k and $\gamma_{jk}(i)$ is the number of shorter paths ranging from the node j to node k and passing through the node i .

Closeness Centrality. The heuristic principle behind this measure is based on the concept of closeness³⁸. A node is central when it can reach any node in the network in few steps, i.e., is closer to all nodes. Thus, the node with higher centrality value will be one with the lowest average length of shortest paths, i.e.,

$$C_c(i) = \frac{n - 1}{\sum_{\substack{j=1 \\ j \neq i}}^n d(i, j)} \tag{12}$$

where $d(i, j)$ is the shortest path distance between the nodes i and j . In the case where the network is unconnected, closeness centrality is calculated separately for each connected component is calculated separately.

Information Centrality. The information centrality³⁹ is based on the study of how information flows between all pairs of nodes in the network. For a given node i , the information centrality is calculated through the harmonic mean of the combined paths information, for all nodes j in the network, meaning by combined path between i and j the set of all paths joining this pair of nodes. Thus, we have mathematically

$$\bar{I}_i = \frac{n}{\sum_{j=1}^n I_{ij}} \quad (13)$$

The information of a path is defined as the inverse of its length. If paths that make up a combined path are independent (i.e not shared links) then the combined path information I_{ij} is given by the sum of the information of the paths that compose it. Otherwise, it is necessary to calculate the matrix $D(i, j)$ containing the number of links that share the paths in combined path and its information is given by $I_{ij} = \sum_{r,s} [D(i, j)^{-1}]_{rs}$.

Communicability Centrality. The communicability centrality^{40,41}, also called subgraph centrality exploit all closed paths of all lengths that start and end at a node i , having greater influence on the centrality of node i paths with shorter length. Mathematically, the communicability of a node $i \in \mathcal{G}$ is calculated by the exponential of the adjacency matrix of the network as:

$$\text{Com}(i) = (e^A)_{ii} \quad (14)$$

References

- Boccaletti, S., Latora, V., Moreno, Y., Chavez, M. & Hwang, D.-U. Complex networks: Structure and dynamics. *Phys. Rep.* **424**, 175–308 (2006).
- Newman, M. E. J. The Structure and function of complex networks. *SIAM Rev.* **45**, 167–256 (2003). Plath, S. (2000). The unabridged journals. Kukil, K. V. (Ed.). New York, NY: Anchor.
- Dorogovtsev, S. N. & Mendes, J. F. F. *Evolution of Networks: From Biological Nets to the Internet and WWW*. (Oxford University Press, Oxford, 2003).
- Watts, D. J. & Strogatz, S. H. Collective dynamics of “small-world” networks. *Nature* **393**, 440–442 (1998).
- Barabási, A.-L. & Albert, R. Emergence of scaling in random networks. *Science* **286**, 509–512 (1999).
- Fortunato, S. Community detection in graphs. *Phys. Rep.* **486**, 75–174 (2010).
- Mones, E., Vicsek, L. & Vicsek, T. Hierarchy measure for complex networks. *PLoS One.* **7**, 33799 (2012).
- Barabási, A.-L. & Ravasz, E. Hierarchical organization in complex networks. *Phys. Rev. E.* **67**, 026112 (2003).
- Newman, M. E. J. & Girvan, M. Finding and evaluating community structure in networks. *Phys. Rev. E.* **69**, 026113 (2004).
- Koschützki, D., Lehmann, K. A. & Peeters, L. In Brandes, U. & Erlebach, T. (Eds.), *Network analysis*, LNCS 3418, 16–61 (Springer, New York, 2005).
- Ibarra, H. & Andrews, S. B. Power, social influence, and sense making: Effects of network centrality and proximity on employee perceptions. *Admin. Sci. Quart.* **38**, 277 (1993).
- Wasserman, S. & Faust, K. Social Network analysis: Methods and applications. *Soc. Net.* **8**, 825 (1994).
- Borgatti, S. P. Centrality and network flow. *Soc. Net.* **27**, 55–71 (2005).
- Shah, D. & Zaman, T. Rumors in a network: Whos the culprit? *IEEE Trans. Inf. Theory* **57**, 5163–5181 (2011).
- Jeong, H., Mason, S. P., Barabási, A. L. & Oltvai, Z. N. Lethality and centrality in protein networks. *Nature* **411**, 41–42 (2001).
- Tew, K. L., Li, X.-L. & Tan, S.-H. Functional centrality: detecting lethality of proteins in protein interaction networks. *Gen. inf. Int. Conf. Gen. Inf.* **19**, 166–177 (2007).
- Rio, G. del, Koschützki, D. & Coello, G. How to identify essential genes from molecular networks? *BMC Systems Biology.* **3**, 102 (2009).
- Raman, K., Damaraju, N. & Joshi, G. K. The organisational structure of protein networks: Revisiting the centrality-lethality hypothesis. *Syst. Synth. Bio.* **8**, 73–81 (2014).
- Pastor-Satorras, R. & Vespignani, A. Epidemic spreading in scale-free networks. *Phys. Rev. Lett.* **86**, 3200 (2001).
- Moreno, Y. & Vázquez, A. Disease spreading in structured scale-free networks. *Eur. Phys. Jour. B.* **31**, 265–271 (2003).
- Colizza, V., Barrat, A., Barthélemy, M. & Vespignani, A. The role of the airline transportation network in the prediction and predictability of global epidemics. *Proc. Natl. Acad. Sci. USA* **103**, 2015–2020 (2006).
- Newman, M. *Networks: An Introduction*. (OUP, Oxford, 2010).
- Estrada, E. *The Structure of Complex Networks: Theory and Applications*. (Oxford University Press, 2011).
- Cohen, R. & Havlin S. *Complex Networks: Structure, Robustness and Function*. (Cambridge University Press, 2010).
- Bonacich, P. Power and centrality: A family of measures. *Am. Jour. Soc.* **92**, 1170 (1987).
- Meyer, C. *Matrix Analysis and Applied Linear Algebra*. (SIAM, Philadelphia, 2000).
- Saad, Y. *Numerical Methods for Large Eigenvalue Problems: Revised Edition*. (SIAM, 2011).
- Levandowsky, M. & Winter, D. Distance between sets, *Nature* **234**, 34–35 (1971).
- Batool, K. & Niazi, M. Towards a Methodology for Validation of Centrality Measures in Complex Networks. *PLoS One.* **9**(4) (2014).
- Padgett, J. F. & Ansell, C. K. Robust action and the rise of the Medici, 1400–1434. *Am. Jour. Soc.* **98**, 1259 (1993).
- Mathew, J. *Social and Economic Networks*. (Princeton University Press, 2010).
- Zachary W. W. An information flow model for conflict and fission in small groups. *Jour. Anth. Res.* **33**, 452–473 (1977).
- Knuth D. E. *The Stanford GraphBase: A Platform for Combinatorial Computing*. (Addison-Wesley, Reading, MA, 1993).
- Lipkus, A. A proof of the triangle inequality for the Tanimoto distance. *Jour. Math. Chem.* **26**, 263–265 (1999).

35. Brin, S. & Page, L. The anatomy of a large-scale hypertextual web search engine. *Computer networks and ISDN systems* **30**, 107–117 (1998).
36. Wasserman, S. & Faust, K. *Social Network Analysis: Methods and Applications*. (Cambridge University Press, 1994).
37. Brandes, U. A faster algorithm for betweenness centrality. *The Jour. Math. Socio.* **25**, 163–177 (2001).
38. Sabidussi, G. The centrality index of a graph. *Psychometrika*. **31**, 4, 581–603 (1996).
39. Stephenson, K. & Zelen, M. Rethinking centrality: Methods and examples. *Social Networks* **11**, 1–37 (1989).
40. Estrada, E. & Rodriguez-Velázquez, J. A. Subgraph centrality and clustering in complex hyper-networks. *Physica A*. **364**, 581–594 (2006).
41. Estrada, E. & Hatano, N. Communicability in complex networks. *Phys. Rev. E*. **77**, 036111 (2008).

Acknowledgements

This work was supported by Venezuelan Institute for Scientific Research (IVIC) under project No. 1089.

Author Contributions

A.J.A.-S. conceived the project, A.J.A.-S. and G.C.H.-A. performed the simulations. G.C.H.-A. prepared all figures and tables. A.J.A.-S. wrote the manuscript, all authors discussed the results and reviewed the manuscript. G.A.G.-D. supervised the work.

Additional Information

Supplementary information accompanies this paper at <http://www.nature.com/srep>

Competing financial interests: The authors declare no competing financial interests.

How to cite this article: Alvarez-Socorro, A. J. *et al.* Eigencentrality based on dissimilarity measures reveals central nodes in complex networks. *Sci. Rep.* **5**, 17095; doi: 10.1038/srep17095 (2015).



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>