SOFTWARE TOOL ARTICLE

Biological network analysis with CentiScaPe: centralities and experimental dataset integration [version 1; peer review: 2 approved with reservations]

Giovanni Scardoni¹, Gabriele Tosadori¹, Mohammed Faizan², Fausto Spoto³, Franco Fabbri¹, Carlo Laudanna^{1,4}

¹Center for Biomedical Computing, University of Verona, Verona, 37134, Italy

²Birla Institute of Technology & Science, Goa, 403726, India

³Department of Computer Science, University of Verona, Verona, 37134, Italy

⁴Department of Pathology and Diagnostics, University of Verona, Verona, 37134, Italy

V1 First published: 01 Jul 2014, 3:139
<https://doi.org/10.12688/f1000research.4477.1>

Latest published: 07 Jul 2015, 3:139
<https://doi.org/10.12688/f1000research.4477.2>

Abstract

The growing dimension and complexity of available experimental data generating biological networks has increased the need for tools allowing to categorize nodes by their topological relevance in biological networks. Here we present CentiScaPe, a Cytoscape app specifically designed to calculate centrality indexes for the identification of the most important nodes of a network. CentiScaPe is a comprehensive suite of algorithms dedicated to network nodes centrality analysis, computing several centralities for undirected, directed and weighted networks. The results of the topological analysis can be also integrated with data sets from lab experiments, such as expression or phosphorylation levels of the proteins represented in the network, using the graphical features of the tool. This opens a new perspective in the analysis of biological networks, since integration of topological analysis with lab experimental data can increase the predictive power of a bioinformatical analysis.

This article is included in the **Cytoscape** gateway.

Open Peer Review

Approval Status

	1	2
version 2
(revision)
07 Jul 2015
version 1
01 Jul 2014

1. **Ferenc Jordan**, University of Trento, Trento, Italy

2. **Ankush Sharma** , High Performance Computing and Networking Institute, National Research Council, Naples, Italy

Any reports and responses or comments on the article can be found at the end of the article.

Corresponding author: Giovanni Scardoni (giovanni.scardoni@gmail.com)

Competing interests: No competing interests were disclosed.

Grant information: This work was supported by: Italian Association for Cancer Research (AIRC, IG 8690) (C.L.); Fondazione Cariverona; Nanomedicine project University of Verona and Fondazione Cariverona (C.L.). Directed centralities have been developed thanks to the GSoC2013 program.

Copyright: © 2014 Scardoni G *et al.* This is an open access article distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

How to cite this article: Scardoni G, Tosadori G, Faizan M *et al.* **Biological network analysis with CentiScaPe: centralities and experimental dataset integration [version 1; peer review: 2 approved with reservations]** F1000Research 2014, 3:139 <https://doi.org/10.12688/f1000research.4477.1>

First published: 01 Jul 2014, 3:139 <https://doi.org/10.12688/f1000research.4477.1>

Introduction

Biological processes can be displayed as networks where the nodes represent biological entities, and edges represent interactions between these entities. Several kinds of biological networks have been introduced, including metabolic networks, gene networks, signal transduction networks, and protein-protein interaction networks¹. Such networks are a static representation of the dynamics of biological processes, where molecular interactions give rise to cascades of reactions, called pathways, determining the life processes of living organisms. Even if the scientific community is far from being able to simulate the dynamic behavior of such pathways, important information can be extracted from the topological analysis of biological networks^{2,3} since the structure of a network can affect its function⁴. In this context, several global parameters are commonly used to describe the properties of the whole networks, while centralities⁵ are parameters allowing the identification of the most important nodes which could be key regulators in the biological process being studied.

In this paper we present CentiScaPe 2.1, a Cytoscape app^{6,7} for network centralities analysis. While the built in Analyze Network tool is oriented to characterize the global behavior of the network. Provided with several global network statistics, CentiScaPe is designed to identify the most relevant nodes, and it provides a more complete set of centralities. This new version introduces the computation of centralities for directed and weighted networks not available in any other Cytoscape app. As in the previous version, it computes Average Distance, Diameter, Degree, Stress, Betweenness, Radiality, Closeness, Centroid Value, and Eccentricity. Other parameters as Eigenvector, Bridging centrality and Edge Betweenness have been added, making it the most complete app for network centrality analysis.

A web version of CentiScaPe, FastCentiScaPe, is also available (<http://www.cbmc.it/fastcent/>), which performs very fast computations for large networks sending the network to a multiprocessor server. The centrality analysis results are sent to the user by e-mail in xgmml format.

CentiScaPe's main goal is to produce results that should drive further lab experiments, as the high score nodes identified by the computation can be considered as potential targets for drug development and new experiments.

Methods and implementation

To calculate all the centralities, the computation of the shortest path between each pair of nodes in the graph is needed. The algorithm for the shortest path is the Dijkstra algorithm⁸, that has been adjusted in order to compute all the shortest paths pairs (needed for Stress and Betweenness computation)⁹ and to use also edge direction and edge weight in the shortest path identification. A good description of this and other centrality algorithms can be found in Koschützki *et al.*⁵. CentiScaPe 2.1 introduces the computation of centralities for directed networks, networks where the edges are considered to have a direction. Consequently, some nodes cannot be reached by others: given two nodes s and t , it is possible that there is no path from s to t , or from t to s or both. Several centrality parameters are based on the computation of the shortest path between each couple of nodes

and are not defined when there are two nodes not connected by a path. In this case CentiScaPe consider the distance from s to t equal to infinity. The centrality definitions have been modified to consider this case (see supplementary materials: CentralitiesTutorial), so the directed centralities can also be used to compute centralities for disconnected networks, i.e. networks where some nodes cannot be reached by others. This gives the use of centrality analysis great flexibility.

In the case of weighted networks the edges are supposed to have a numerical integer attribute depending on experimental data or on some feature of the network. This numerical value is treated as a distance in the computation of the shortest path between two nodes. In an unweighted network the distance between two node connected by an edge is equals to 1, and the distance of two generic nodes is the number of edge of the shortest path connecting them. In a weighted network the distance is the sum of the attribute values of the edges connecting the nodes, so the shortest path is not necessary the one with the lowest number of edges, but the one with the least distance.

CentiScaPe is written in Java as a Cytoscape app, in order to exploit all the excellent features of Cytoscape and to reach the largest number of users. It has a multi-thread core that can exploit multiprocessor architecture. The Java library JFreechart¹⁰ has been used for some of the graphic features.

Results and discussion

The Main use of CentiScaPe is to rank the nodes of a network depending on their topological and experimental relevance. The numerical results are saved as node, edge or network attributes in the Cytoscape attributes browser, depending on the kind of parameters, so all the Cytoscape features for managing attributes are supported. After the computation the centralities are treated as normal Cytoscape attributes. CentiScaPe can be used in undirected networks¹¹, in directed networks and in weighted networks.

Centralities for directed networks (see Supplementary Files: CentralitiesTutorial) are useful in the case of metabolic networks where the direction is from substrates and reactants to the products of the chemical reactions and in signal transduction networks, where the direction depends on the flux of information. Considering direction in the computation of centralities can lead to different and more precise results than the undirected version.

As example, in **Figure 1** the computation of the directed and undirected Stress applied to a network of Oncogenes is shown (see Supplementary Files: Oncogenes.txt and Oncogenes_edge_directions.txt). Results of both the computations are shown. The image, obtained using Cytoscape, represents the different Stress values using the color and size of nodes. The node's size represents the value obtained using the directed Stress, so the bigger the node the higher the value; the color represents the values obtained using the undirected Stress: red is used for the highest values, green for the lowest values. For example a large green node is interesting because it means that a node with a high value of the new algorithm has a low value using CentiScaPe. While analyzing the oncogenes network we saw that the large red node, AKT1, shows how its stress values are high using both

Figure 1. Stress computed on oncogenes. Size represents directed Stress, color undirected Stress (green=low, red=high).

algorithms. However the green medium-sized node, FANCE or RAF1, shows how, using undirected Stress, we obtain a low value for stress, but using the new algorithm we obtain a high stress value. The opposite situation is found in the third highlighted node, the small yellow node, RB1, in the right bottom corner. Here, the value computed with undirected Stress is not very high, similar to the red node, but the value computed with the directed Stress is very low. This can be explained by suggesting that AKT1 is a very important node, that is essential for maintaining connections within this network. For RB1 and CREB1 the situation is not very clear because we have opposite situations. If we use CentiScape, for FANCE, we assume that this node is not essential, but using the directed values it appears to be very central.

Second important features of the new version of CentiScape is the possibility of computing centralities for weighted networks, networks where the edges are provided with an attribute that can be interpreted as a distance between the two connected nodes.

In the network depicted in Figure 2 we have a distance (dist) attribute for each edge. We have $\text{dist}(A,B)=2$, $\text{dist}(B,C)=3$ and $\text{dist}(A,C)=7$. Since A and C are connected by a single edge, in an unweighted computation, the distance from A to C is equal to 1. But if the attributes of the edges are considered as distances, the shortest path between A and C is the one passing through B ($=2+3=5$) since it is shorter than the one connecting A directly to C ($=7$). The computation of weighted shortest paths will result in completely different

Figure 2. A weighted network. If the weight of the edge is not considered the shortest path from A to C is the directed edge, and the distance is 1. If the weights are considered the shortest path from A to C is the one passing through B, and the distance is equal to 5.

values compared to a situation where the weight is not considered. The user should take care that the weight is used in cases where close nodes are more important than distant nodes. So, depending on the meaning of the attributes, one can use the value or its reciprocal. For example if the attribute represents the speed of a reaction instead of a distance, the reciprocal should be used.

An example where weighted networks centrality analysis is used, is provided by Currie *et al.*¹² where an euclidean distance is given to each edge depending on the difference between the phosphorylation level of the proteins connected by that edge.

All the graphical features of the previous version of CentiScaPe, such as the plot by node, the plot by centrality and the boolean-based result panel have been maintained in the new version. A complete guide can be found in Scardoni *et al.*¹¹ or in the CentiScaPe userguide available from the website.

Conclusions

CentiScaPe 2.1 has been enriched with new centrality parameters such as Eigenvector, Bridging centrality and Edge Betweenness centrality, and with the possibility to analyze directed and weighted

networks. It allows integrating centrality-based network analysis with experimental data. The results of the computation can be used and exported as Cytoscape attributes, allowing the user to exploit all the other features of Cytoscape and its apps. Compared to the built-in Analyze Network tool of Cytoscape, CentiScaPe is an excellent integrative tool allowing the identification of potential target nodes from both the topological and the experimental point of view, and can be considered as an essential instrument for the characterizations of nodes in order to drive further experiments.

Software availability

Software available from the Cytoscape App Store: <http://apps.cytoscape.org/apps/centiscape>

Latest source code: <https://bitbucket.org/giovanniscardoni/centiscapepublic/>

Source code as at the time of publication: <https://bitbucket.org/F1000Research/centiscapepublic-archive>

Archived source code as at the time of publication: <http://dx.doi.org/10.5281/zenodo.10652>¹³

License: Lesser GNU Public License 3.0: <https://www.gnu.org/licenses/lgpl.html>

Author contributions

GS is the main designer and developer of CentiScaPe. CL contributed to the design of CentiScaPe and performed the experiments. GT contributed to the definition of the directed centralities and performed the examples of usage in the Supplementary files. MF developed the directed centralities. FS contributed to the development of the last version of CentiScaPe. FF is the main developer of the web version of CentiScaPe.

Competing interests

No competing interests were disclosed.

Grant information

This work was supported by: Italian Association for Cancer Research (AIRC, IG 8690) (C.L.); Fondazione Cariverona; Nanomedicine project University of Verona and Fondazione Cariverona (C.L.). Directed centralities have been developed thanks to the GSOC2013 program.

Supplementary files

Supplementary files available from: <https://f1000researchdata.s3.amazonaws.com/supplementary/4477/dd4541d4-bdfe-461a-be92-f2e2cf9f2d57.zip>

References

1. Guido Caldarelli. **Scale-Free Networks: Complex Webs in Nature and Technology (Oxford Finance)**. Oxford University Press, USA. 2007.
[Reference Source](#)
2. Jeong H, Tombor B, Albert R, *et al.*: **The large-scale organization of metabolic networks**. *Nature*. 2000; **407**(6804): 651–654.
[PubMed Abstract](#) | [Publisher Full Text](#)
3. Barabási AL, Oltvai ZN: **Network biology: understanding the cell's functional organization**. *Nat Rev Genet*. 2004; **5**(2): 101–113.
[PubMed Abstract](#) | [Publisher Full Text](#)
4. Strogatz SH: **Exploring complex networks**. *Nature*. 2001; **410**(6825): 268–276.
[PubMed Abstract](#) | [Publisher Full Text](#)
5. Dirk K, Lehmann KA, Peeters L, *et al.*: **Centrality indices**. In Ulrik Brandes and Thomas Erlebach, editors, *Network Analysis: Methodological Foundations*, Springer. 2005; 16–61.
[Publisher Full Text](#)
6. Cline MS, Smoot M, Cerami E, *et al.*: **Integration of biological networks and gene expression data using Cytoscape**. *Nat Protoc*. 2007; **2**(10): 2366–2382.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
7. Saito R, Smoot ME, Ono K, *et al.*: **A travel guide to Cytoscape plugins**. *Nat Methods*. 2012; **9**(11): 1069–76.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
8. Dijkstra EW: **A note on two problems in connexion with graphs**. *Numerische Mathematik*. 1959; **1**(1): 269–271.
[Publisher Full Text](#)
9. Scardoni G, Laudanna C: **Centralities based analysis of complex networks**. In Yagang Zhang, editor, *New Frontiers in Graph Theory*. InTech. 2012.
[Publisher Full Text](#)
10. Gilbert D: **JFreeChart**.
[Reference Source](#)
11. Scardoni G, Petterlini M, Laudanna C: **Analyzing biological network parameters with CentiScaPe**. *Bioinformatics*. 2009; **25**(21): 2857–2859.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
12. Currie HN, Vrana JA, Han AA, *et al.*: **An approach to investigate intracellular protein network responses**. *Chem Res Toxicol*. 2014; **27**(1): 17–26.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
13. Scardoni G, Tosadori G, Faizan M, *et al.*: **F1000Research-centiscapepublic-archive**. *ZENODO*. 2014.
[Data Source](#)

Open Peer Review

Current Peer Review Status: ? ?

Version 1

Reviewer Report 11 July 2014

<https://doi.org/10.5256/f1000research.4788.r5341>

© 2014 Sharma A. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Ankush Sharma

Laboratory for Genomics, Transcriptomics and Proteomics, High Performance Computing and Networking Institute, National Research Council, Naples, Italy

In this research article, the authors describe the CentiScaPe app for Cytoscape (and its web version FastCentiScaPe) for computing the several centrality parameters in networks, analyzing the elemental level importance of each node in the network, based on network topology. The addition of centralities parameters, especially Bridging Centrality and Eigenvector centrality, appears to be useful to improve metrics, to infer the informational flow across overlapping modules, and as a measure of the influence of a node in a network. The article is well written with an appropriate title and an abstract which provides sufficient details. The description of the purpose, implementation and usage of CentiScaPe app are informative and detailed for the audience.

Minor comments:

1. The example of calculating stress in the network of oncogenes in the manuscript should comply with experimental functional relevance, e.g. a decrease in bridging centrality is observed for key intermodular nodes upon heat shock response or oxidative stress, partially disintegrating overlapping modules into local communities which alters the centrality statistics of the nodes ^{1,2}. It would be good to see such implementation in the example showing applicability of CentiScaPe.
2. I suggest the authors have a small example of the analysis comparing connected and disconnected networks, to support computation of centralities for networks with a disconnected component.
3. It would be better to implement a user-friendly exploration of results and user registration feature instead of email authorization, in order to avoid delays in using the web version FastCentiScaPe.

4. Correction in the name of built-in tool i.e. NetworkAnalyzer , instead of "Analyze Network" tool.
5. Usage of appropriate words such as "more comprehensive" in place of "more complete" in introduction section.

References

1. Lehtinen S, Marsellach FX, Codlin S, Schmidt A, et al.: Stress induces remodelling of yeast interaction and co-expression networks. *Mol Biosyst.* 2013; **9** (7): 1697-1707 [PubMed Abstract](#) | [Publisher Full Text](#) | [Reference Source](#)
2. Mihalik A, Csermely P: Heat shock partially dissociates the overlapping modules of the yeast protein-protein interaction network: a systems level model of adaptation. *PLoS Comput Biol.* 2011; **7** (10). [PubMed Abstract](#) | [Free Full Text](#) | [Publisher Full Text](#) | [Reference Source](#)

Competing Interests: No competing interests were disclosed.

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Reviewer Report 07 July 2014

<https://doi.org/10.5256/f1000research.4788.r5358>

© 2014 Jordan F. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Ferenc Jordan

Centre for Computational and Systems Biology, University of Trento, Trento, Italy

Analysing several centrality measures in parallel is important and there is a need for this, however, it would be even nicer to compare them (see

<http://www.sciencedirect.com/science/article/pii/S0304380007001184>) and discuss better their biological relevance (i.e. what is Eigenvector centrality good for?). Central nodes can be of key regulator function but this is not a result, this is still a hypothesis. It should be mentioned briefly that either simulations or lab experiments should reinforce these findings. For infinite distances, using the reciprocal distance matrix has already been suggested as an elegant solution (see <http://link.springer.com/article/10.1007/BF01164642#page-1>). It should be clarified a bit more that a large weight can be considered as a short or as a long path, depending on its biological

meaning. In Figure 1, the grid layout algorithm could be replaced by some better one, I think. Finally, a quick English check would be welcome.

Competing Interests: No competing interests were disclosed.

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

The benefits of publishing with F1000Research:

- Your article is published within days, with no editorial bias
- You can publish traditional articles, null/negative results, case reports, data notes and more
- The peer review process is transparent and collaborative
- Your article is indexed in PubMed after passing peer review
- Dedicated customer support at every stage

For pre-submission enquiries, contact research@f1000.com

F1000Research