

Analyzing Multiple Network Centralities with ViNCent

Björn Zimmer, Ilir Jusufi, and Andreas Kerren

Linnaeus University, School of Computer Science, Physics and Mathematics (DFM), ISOVIS Group,
Vejdes Plats 7, 351 95 Växjö, Sweden

Abstract

The analysis of multivariate networks is an important task in various application domains, such as social network analysis or biochemistry. In this paper, we address the interactive visual analysis of the results of centrality computations in context of networks. An important analytical aspect is to examine nodes according to specific centrality values and to compare them. We present a tool that combines exploratory data visualization with automatic analysis techniques, such as computing a variety of centrality values for network nodes as well as hierarchical clustering or node reordering based on centrality values. Automatic and interactive approaches are seamlessly integrated in one single tool which provides insight into the importance of an individual node or groups of nodes and allows quantifying the network structure.

Categories and Subject Descriptors (according to ACM CCS): E.1 [Data Structures]: —Graphs and networks H.5.2 [Information Interfaces and Presentation]: User Interfaces—Graphical user interfaces (GUI) I.3.6 [Computer Graphics]: Methodology and Techniques—Interaction techniques

1. Introduction

Network centralities are used to discover the relative importance of nodes within a complex network topology. Especially the comparison of several different centrality values in larger networks is an open problem. Identifying communities and central actors in social networks or the calculation of the importance of specific nodes in biochemical networks are some examples where multiple network centralities are used. Typical tasks during network centrality analysis include: finding nodes with high/low centrality values or finding nodes with high values in several centralities across a large number of nodes. Especially the latter task raises challenging analytical problems.

In this paper, we propose analytical extensions for our tool ViNCent [KKZ12] that offers solutions to interactively visualize networks together with their centralities. Our previous work addressed mainly visualization problems, such as the visual representation of the network itself together with the computed centralities, the interactive filtering of (non-)interesting nodes by using histogram brushing, or the minimization of clutter by the implementation of several edge bundling methods that work without a node hierarchy. Here, we discuss a set of extensions which support the analysis process. These are *node reordering* in the circular network drawing for selected centrality measurements, automatic *cluster analyses* over a set of centralities to group

similar nodes together, adding a node-link view to provide the analyst with traditional topological network information, and a complete reimplementing of our tool using *OpenGL* to speed up the analysis process for larger networks.

Background This paragraph provides a brief overview of network centralities in order to facilitate the understanding of the rest of the paper. Because of the page limit, we abandon the inclusion of standard graph definitions and centrality formula. Instead, we refer the reader to various works, such as [GPQX07, DBETT99] for graph definitions or [JKS06, DHK*06, JKL*05] for centrality measurements.

A network centrality C is a function that assigns a value $C(u)$ to a node $u \in V$ of a given graph $G = (V, E)$. In order to compare network centralities according to their importance, u is more important than v if $C(u) > C(v)$, with $u, v \in V$. By using network centralities, analysts are able to better understand the structure of networks and to identify central actors. Typical examples are *degree*, *eccentricity* or *random walk betweenness* [New10]. Whereas the *degree* of a node is a very simple centrality that orders the nodes according to their degree values, *eccentricity* is calculated by using the distance (based on shortest paths) between nodes of an undirected, connected graph. More central nodes have therefore a higher value. Our tool uses the CentiBiN plugin [JKS06] that computes up to 17 different network centralities. Note that not every centrality measure can be applied



Figure 1: Overview of ViNCent. The center shows the radial centrality view of the Les Misérables network. Alternatively, another view with a node-link drawing can be displayed. The right side displays the corresponding histograms of the network centralities as well as detailed values of the network centralities for the currently hovered node “Thenardier”. The left panel allows changing the render settings and displays an overview of the respective node-link layout of the network. A node group has been manually created in the node-link view and is shown as a light-blue stripe along the outer circle in the centrality view as well as in the overview (bottom left) by using a background region of the same color.

to every graph. The problem of choosing the right centralities differs from network to network. For the computation of suitable centralities, data about the functional properties of networks is often missing. This data would allow to choose the “right” centrality measures. Therefore, this analysis is usually done by visual comparison of centrality values on the networks [DHK*06].

Related Work The visualization of network centralities was not much discussed in the visualization literature so far. Typical methods are the use of correlations, scatter plots, and parallel coordinates. These solutions have disadvantages when used with networks, since they do not show where those correlations occur within the network. Dwyer et al. [DHK*06] describe three techniques to visualize network centralities in context of network drawings. These techniques and other related works are described in our previous work [KKZ12].

In the field of information visualization, there is much related work in context of general network analysis, especially in the field of social networks. Some tools include the analysis of centrality measurements, for example, Social-

Action [PS08]. It allows users to rank nodes by centralities and to display them in coordinated views as ordered lists or to compare them by scatterplots. A recent overview of techniques and tools for social network analysis is given by Correa and Ma [CM11].

2. Approach

ViNCent provides multiple, coordinated views on the input data, see Figure 1 for an overview. Concretely, it supports a radial centrality view, a standard node-link view, interactive histograms for filtering as well as a view to display the individual centrality values of hovered nodes. The node-link view uses a standard force-directed layout algorithm to show the general network structure. The centrality view is used to display and analyze all calculated centralities of the input network. Each network node is represented by a small quadrangle that is positioned on a circle. The quadrangle is gray-scaled according to the corresponding node’s degree. Its connections to the other nodes (i.e., the edges) are laid out inside of this circle. The centralities are stacked as additional quadrangles on each node, providing a good comparability of all relative centrality values. In order to distinguish

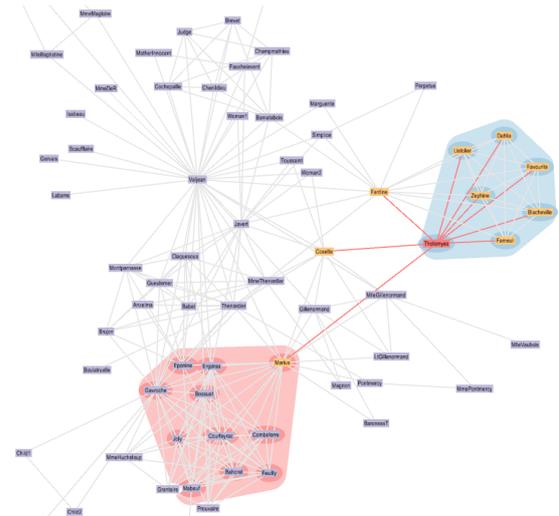
between single centrality values, a specific color schema is employed. A more detailed explanation about the visualization and interaction approaches used in our tool can be found in [KKZ12]. For our sample screenshots, we use the network of characters in the novel *Les Miserables* [Knu93]. It consists of 77 nodes which represent characters and 254 edges which connect characters that appear in the same chapter of the novel.

2.1. Interactive Visual Analysis Methods

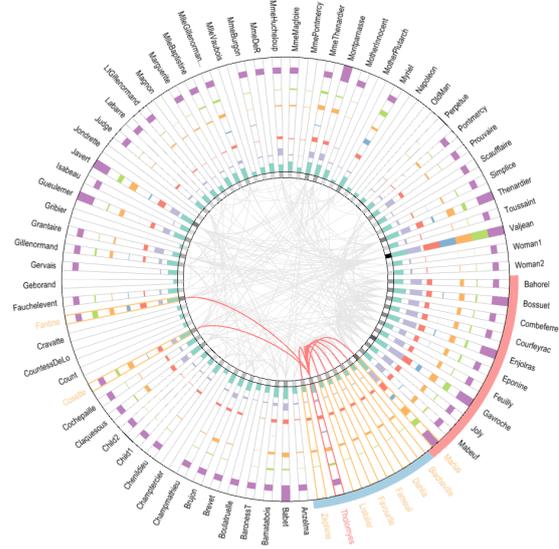
All views of ViNCent are connected to each other via brushing and linking. Hovering an object in the network centrality view, node-link view or histogram bars highlights the respective node or its corresponding data in all other views. Filtering of nodes or centralities is possible by clicking on a histogram bar, by changing the range sliders of the histograms or by clicking on a node in the centrality view. To better facilitate the analysis of centralities in a graph, our reimplementation of ViNCent offers new options to group and to order all or selected nodes in the centrality view based on their centrality values.

Node Grouping If the initial layout of a graph displayed in the node-link view reveals a couple of nodes that look interesting, the user can manually select these nodes and group them together. All grouped nodes will be drawn side by side in the centrality view, increasing the comparability of all centralities in a group. The group markers are drawn as small colored stripes along the outer circle in the centrality view. Figure 2 shows an example with two manually selected groups. The user can also automatically group all nodes by using a k -Means clustering algorithm [Mir05]. For this, he/she chooses which centralities should be used for the clustering algorithm and how many clusters (k) are desired. Automatic clustering is especially useful while analyzing large networks, where it is difficult to see central nodes by just looking at the node-link representation. Figure 3 shows the centrality view after clustering the *Les Miserables* graph.

Node Ordering By clicking on a rectangle of a specific centrality value in the centrality view, all nodes are arranged by their corresponding values in descending order. This makes it easier to compare centrality values within the same magnitude. Figure 3 shows a small use case based on our sample data set. In Figure 3(a), the user has clicked on the *centroid* centrality for the node representation of the character “Marius”. All nodes are now aligned in descending order of their respective *centroid* centrality values, whereas the ordering is done for each group individually. In this case, “Marius” has the highest *centroid* centrality and is therefore drawn at the first position of his group. This rule applies for all following groups on the circle, providing the possibility to quickly identify the highest/lowest centrality values for each group individually. Figure 3(b) shows the same data set with a node ordering based on the *eigenvector* centrality.



(a) Node-link view with two groups manually created by the user.

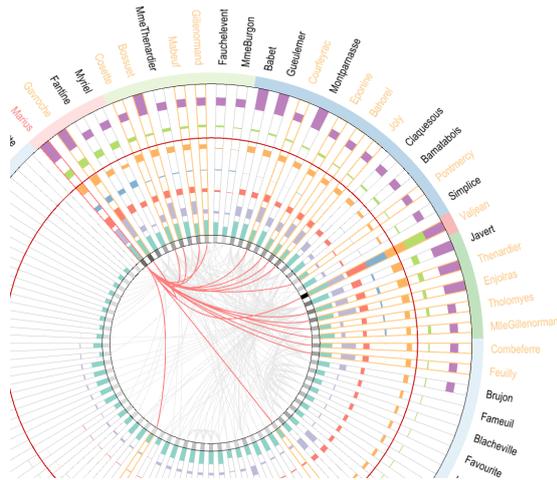


(b) Centrality view showing the selected groups. The group markers are drawn as colored stripes along the outer circle.

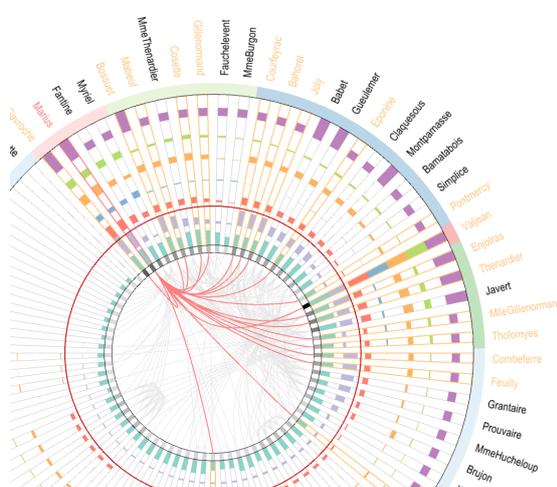
Figure 2: Two manually grouped nodes displayed in the node-link (a) and centrality view (b). Color brushing is used to identify the groups in each view. The name “Tholomyes” is hovered and simultaneously highlighted in red color within both views. All direct neighbors of the corresponding node, i.e., in graph-theoretic distance of one, are highlighted in orange color together with the connecting edges.

3. Conclusions and Future Work

The additional node-link view enables the user to better perceive the network’s topology in order to interactively pre-select interesting nodes. It also supports node grouping to analyze the nodes in the centrality view afterwards. With the help of our new node ordering functionality, nodes with



(a) Centrality view ordered by *centroid* centrality (orange).



(b) Centrality view ordered by *eigenvector* centrality (lavender).

Figure 3: Both centrality views show the groups created by a *k*-means clustering algorithm and the node ordering for two different centrality values. Here, nodes with low centrality values are in the largest group on the lower part of the circle, whereas all nodes with higher centralities are grouped into five smaller clusters on the upper right part. Thus, node grouping and ordering provide an easy way to analyze and compare the highest centrality values. While a centrality of a node is hovered (here: “Marius”) a red circle is drawn at the corresponding centrality value to allow a better comparison of this value with those of other nodes.

similar centralities can be arranged side by side to increase the comparability of their specific network centralities. Also, the lack of immediate feedback (less than five frames per second while exploring larger networks) in our previous work [KKZ12] was solved by using OpenGL to render the centrality and node-link visualizations. This increased the performance to 25 frames per second on a 2.2 Ghz CPU using a Radeon HD 6750M graphics card. For scalability

reasons, we plan to embed centrality views into the node-link representation to either manually or automatically group nodes and replace them by our circular centrality view, similar to the NodeTrix approach [HFM07]. Using hierarchical clustering layouts [NL05] in this context should decrease the number of overlapping groups in the node-link view. Additionally, we will evaluate our tool together with domain experts in biological and social networks to clarify arising challenges while analyzing the centralities of large networks.

References

[CM11] CORREA C. D., MA K.-L.: Visualizing social networks. In *Social Network Data Analytics*, Aggarwal C., (Ed.). Springer, 2011, pp. 307–326. 88

[DBETT99] DI BATTISTA G., EADES P., TAMASSIA R., TOLLIS I. G.: *Graph Drawing: Algorithms for the Visualization of Graphs*. Prentice Hall, 1999. 87

[DHK*06] DWYER T., HONG S.-H., KOSCHÜTZKI D., SCHREIBER F., XU K.: Visual analysis of network centralities. In *Proceedings of the 2006 Asia-Pacific Symposium on Information Visualisation (APVis’06)* (Darlinghurst, Australia, 2006), Misue K., Sugiyama K., Tanaka J., (Eds.), Australian Computer Society, ACM International Conference Proceeding Series, vol. 164, pp. 189–198. 87, 88

[GPQX07] GÖRG C., POHL M., QELI E., XU K.: Visual Representations. In *Human-Centered Visualization Environments* (2007), Kerren A., Ebert A., Meyer J., (Eds.), LNCS Tutorial 4417, Springer, pp. 163–230. 87

[HFM07] HENRY N., FEKETE J.-D., MCGUFFIN M. J.: NodeTrix: a hybrid visualization of social networks. *IEEE Transactions on Visualization and Computer Graphics (IEEE Visualization Conference and IEEE Conference on Information Visualization) Proceedings 13* (2007), 1302–1309. 90

[JKL*05] JACOB R., KOSCHÜTZKI D., LEHMANN K. A., PEETERS L., TENFELDE-PODEHL D.: Algorithms for centrality indices. In *Network Analysis*, Brandes U., Erlebach T., (Eds.). Springer, 2005, pp. 62–82. 87

[JKS06] JUNKER B., KOSCHÜTZKI D., SCHREIBER F.: Exploration of biological network centralities with CentiBiN. *BMC Bioinformatics* 7, 1 (2006), 219. 87

[KKZ12] KERREN A., KÖSTINGER H., ZIMMER B.: Vincent - visualisation of network centralities. In *Proceedings of the International Conference on Information Visualization Theory and Applications (IVAPP ’12)* (2012), INSTICC, pp. 703–712. 87, 88, 89, 90

[Knu93] KNUTH D. E.: *The Stanford GraphBase: a platform for combinatorial computing*. ACM, New York, NY, USA, 1993. 89

[Mir05] MIRKIN B.: *Clustering for Data Mining: A Data Recovery Approach*. Chapman & Hall/CRC, Boca Raton, FL, USA, 2005. 89

[New10] NEWMAN M. E. J.: *Networks: An Introduction*. Oxford University Press, 2010. 87

[NL05] NOACK A., LEWERENTZ C.: A space of layout styles for hierarchical graph models of software systems. In *Proceedings of the 2005 ACM symposium on Software visualization* (New York, NY, USA, 2005), SoftVis ’05, ACM, pp. 155–164. 90

[PS08] PERER A., SHNEIDERMAN B.: Systematic yet flexible discovery: guiding domain experts through exploratory data analysis. In *Proceedings of the 13th International Conference on Intelligent User Interfaces* (New York, NY, USA, 2008), IUI ’08, ACM, pp. 109–118. 88