




# netrankr: An R package for total, partial, and probabilistic rankings in networks

David Schoch <sup>1</sup>

<sup>1</sup> GESIS - Leibniz Institute for the Social Sciences

DOI: [10.21105/joss.04563](https://doi.org/10.21105/joss.04563)

## Software

- [Review](#) 
- [Repository](#) 
- [Archive](#) 

Editor: [Mehmet Hakan Satman](#) 



## Reviewers:

- [@akbaritabar](#)
- [@corybrunson](#)

Submitted: 04 July 2022

Published: 24 September 2022

## License

Authors of papers retain copyright and release the work under a Creative Commons Attribution 4.0 International License ([CC BY 4.0](#)).

## Summary

One of the key concepts in network science is network centrality. Centrality seeks to provide the answer to the question of who (or what) is important in a network depending on the underlying process forming the network and the empirical phenomenon in question. In a nutshell, an actor in a network is more central if they have better relations, where the definition of better relations depends on the conceptualization of structural importance. Applications of centrality can be found in any field where networks arise. In social networks, we may simply be interested in finding the most popular user. In bioinformatics, centrality is used to detect essential proteins in protein-protein interaction networks ([Jeong et al., 2001](#)). Even in sports, centrality is applied to rank athletes or teams ([Radicchi, 2011](#)). A myriad of indices have been proposed, all with differing interpretations of what constitutes a central position within a network. Although netrankr offers this traditional approach to network centrality, its main focus lies on alternative assessments of centrality based on partial and probabilistic rankings in networks.

## Statement of need

General purpose packages for network analysis such as *igraph* ([Csardi & Nepusz, 2006](#)) and *sna* ([Butts, 2008](#)) implement methods for the most frequently used centrality indices.

However, there also exist dedicated centrality packages, such as *centiserve* ([Jalili et al., 2015](#)), *CINNA* ([Ashtiani et al., 2019](#)), and *influenceR* ([Simon & Aditya, 2015](#)). The biggest in terms of implemented indices is *centiserve* which includes 33 indices. The primary purpose of *CINNA* is to facilitate the choice of indices by visual and statistical tools. *influenceR* is a relatively tiny package which implements only a few specialized measures.

The netrankr package also offers the opportunity to apply more than 30 indices, but the main focus lies on alternative assessment methods. Recent results suggest that a variety of partial rankings exists in networks which are preserved by many index based centrality rankings ([Schoch et al., 2017](#); [Schoch & Brandes, 2015, 2016](#)). These partial rankings can be leveraged to assess centrality on a more general level, without necessarily resorting to indices. Some of these methods and key functionalities of the package are listed below.

- Compute partial rankings based on neighborhood-inclusion and other dominance relations in networks ([Brandes, 2016](#); [Schoch & Brandes, 2015](#))
- Analyze the partial rankings using rank intervals ([Patil & Taillie, 2004](#))
- Calculate rank probabilities to assess the likelihood of certain ranks (e.g. How likely is a node the most central one?) ([Schoch, 2018](#))
- Calculate mutual rank probabilities (e.g. How likely is a node more central than another one?) ([Schoch, 2018](#))

- Compute expected ranks of nodes (e.g. How central do we expect a node to be?)([Schoch, 2018](#))

Note that the most of the tools can also be applied in other empirical settings where partial orders arise.

## Background

In an undirected graph  $G = (V, E)$  with vertex set  $V$  (with cardinality  $n = |V|$ ) and edge set  $E$  (with cardinality  $m = |E|$ ), the neighborhood of a node  $u \in V$  is defined as

$$N(u) = \{w : \{u, w\} \in E\}$$

and its closed neighborhood as  $N[v] = N(v) \cup \{v\}$ . If the neighborhood of the node  $u$  is a subset of the closed neighborhood of the node  $v$ ,  $N(u) \subseteq N[v]$ , then it is called a neighborhood inclusion. This concept defines a binary relation among nodes in a network and consequently  $u$  is dominated by  $v$  if  $N(u) \subseteq N[v]$ . Neighborhood-inclusion thus induces a partial ranking on the vertices of a network.

Schoch & Brandes ([2016](#)) showed that if  $c : V \rightarrow \mathbb{R}$  is a centrality index, then

$$N(u) \subseteq N[v] \implies c(u) \leq c(v)$$

that is, neighborhood inclusion is preserved by indices and the ranking can be viewed as linear extensions of the partial ranking induced by neighborhood-inclusion. Analyzing this partial ranking thus means that all possible centrality rankings can be analyzed at once. More technical details can be found in the dedicated literature ([Brandes, 2020, 2016](#); [Schoch et al., 2017](#); [Schoch & Brandes, 2015](#)).

## Example usage

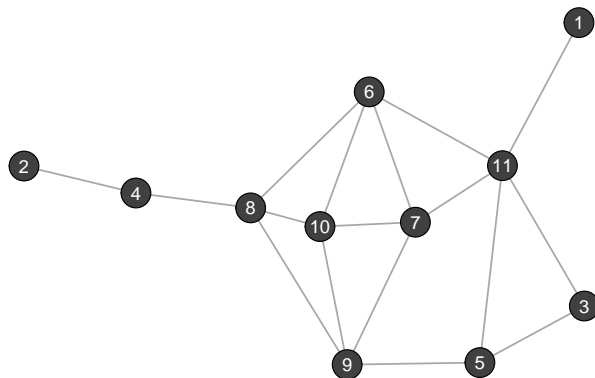
This example briefly explains some of the functionality of the package and the difference to an index driven approach. For more detailed applications see the package vignettes.

We work with a small graph included in the package which was specifically crafted to highlight extreme differences in centrality rankings (see [Figure 1](#)).

```
library(igraph)
library(netrankr)

data("dbces11")
g <- dbces11

# Code to reproduce Figure 1
library(ggraph)
ggraph(g, "stress") +
  geom_edge_link0(edge_color = "grey66") +
  geom_node_point(shape = 21, fill = "grey25", size = 8) +
  geom_node_text(label = 1:11, color = "white")
```



**Figure 1:** Toy graph included in netrankr used to illustrate the functionality of the package.

Say we are interested in the most central node of the graph and simply compute some standard centrality scores with the *igraph* package.

```
cent_scores <- data.frame(
  degree = degree(g),
  betweenness = round(betweenness(g), 4),
  closeness = round(closeness(g), 4),
  eigenvector = round(eigen centrality(g)$vector, 4),
  subgraph = round(subgraph centrality(g), 4))

# What are the most central nodes for each index?
apply(cent_scores, 2, which.max)
#>      degree betweenness  closeness eigenvector  subgraph
#>         11           8         6           7         10
```

Each index assigns the highest value to a different vertex and it is not clear which one is the correct choice.

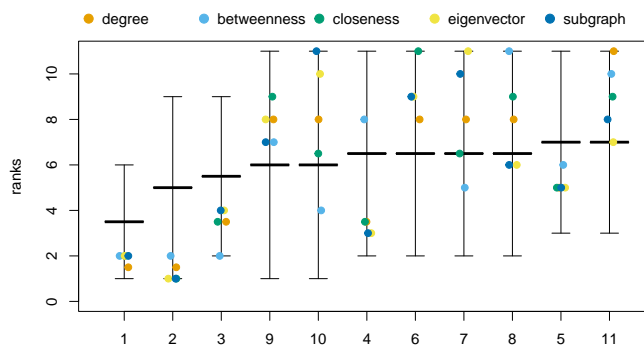
A more general assessment using *netrankr* starts by calculating the neighborhood inclusion preorder.

```
P <- neighborhood_inclusion(g)
P
#>    1 2 3 4 5 6 7 8 9 10 11
#> 1  0 0 1 0 1 1 1 0 0 0 1
#> 2  0 0 0 1 0 0 0 0 1 0 0
#> 3  0 0 0 0 1 0 0 0 0 0 1
#> 4  0 0 0 0 0 0 0 0 0 0 0
#> 5  0 0 0 0 0 0 0 0 0 0 0
#> 6  0 0 0 0 0 0 0 0 0 0 0
#> 7  0 0 0 0 0 0 0 0 0 0 0
#> 8  0 0 0 0 0 0 0 0 0 0 0
#> 9  0 0 0 0 0 0 0 0 0 0 0
#> 10 0 0 0 0 0 0 0 0 0 0 0
#> 11 0 0 0 0 0 0 0 0 0 0 0
```

$P[u, v] = 1$  if  $N(u) \subseteq N[v]$ . Hence,  $u$  is always ranked below  $v$ . If  $P[u, v] = 0$ , then there may be indices that rank  $u$  above  $v$  and vice versa.

We can examine the minimal and maximal possible rank of each node for rankings that extend the preorder to a total order using rank intervals. The bigger the intervals are, the more freedom exists for indices to rank nodes differently (see Figure 2).

```
plot(rank_intervals(P), cent_scores = cent_scores, ties.method = "average")
```



**Figure 2:** Rank intervals for the toy graph showing the range of potential centrality ranks including the rankings of five indices.

Note that the ranks of nodes are not uniformly distributed in the intervals. The exact probabilities, can be obtained with `exact_rank_prob()`.

```
res <- exact_rank_prob(P)
```

`res$rank.prob` contains the probabilities for each node to occupy a certain rank. For instance, the probability for each node to be the most central one is as follows:

```
round(res$rank.prob[,11], 2)
#>  1  2  3  4  5  6  7  8  9 10 11
#> 0.00 0.00 0.00 0.14 0.16 0.11 0.11 0.14 0.09 0.09 0.16
```

The entry `res$relative.rank[u, v]` indicates how likely it is that `v` is more central than `u`.

```
# How likely is it, that 6 is more central than 3?
round(res$relative.rank[3, 6], 2)
#> [1] 0.75
```

`res$expected.ranks` contains the expected centrality ranks for all nodes.

```
round(res$expected.rank, 2)
#>  1  2  3  4  5  6  7  8  9 10 11
#> 1.71 3.00 4.29 7.50 8.14 6.86 6.86 7.50 6.00 6.00 8.14
```

The higher the value, the more central a node is expected to be.

## References

- Ashtiani, M., Mirzaie, M., & Jafari, M. (2019). CINNA: An R/CRAN package to decipher central informative nodes in network analysis. *Bioinformatics*, 35(8), 1436–1437. <https://doi.org/10.1093/bioinformatics/bty819>
- Brandes, U. (2020). Central Positions in Social Networks. In H. Fernau (Ed.), *Computer Science – Theory and Applications* (pp. 30–45). Springer International Publishing.

- Brandes, U. (2016). Network positions. *Methodological Innovations*, 9, 2059799116630650. <https://doi.org/10.1177/2059799116630650>
- Butts, C. T. (2008). Social Network Analysis with SNA. *Journal of Statistical Software*, 24(6). <https://doi.org/10.18637/jss.v024.i06>
- Csardi, G., & Nepusz, T. (2006). The igraph software package for complex network research. *InterJournal, Complex Systems*, 1695(5), 1–9.
- Jalili, M., Salehzadeh-Yazdi, A., Asgari, Y., Arab, S. S., Yaghmaie, M., Ghavamzadeh, A., & Alimoghaddam, K. (2015). CentiServer: A Comprehensive Resource, Web-Based Application and R Package for Centrality Analysis. *PLOS ONE*, 10(11), e0143111. <https://doi.org/10.1371/journal.pone.0143111>
- Jeong, H., Mason, S. P., Barabási, A.-L., & Oltvai, Z. N. (2001). Lethality and centrality in protein networks. *Nature*, 411(6833), 41–42. <https://doi.org/10.1038/35075138>
- Patil, G., & Taillie, C. (2004). Multiple indicators, partially ordered sets, and linear extensions: Multi-criterion ranking and prioritization. *Environmental and Ecological Statistics*, 11(2), 199–228. <https://doi.org/10.1023/b:eest.0000027209.93218.d9>
- Radicchi, F. (2011). Who Is the Best Player Ever? A Complex Network Analysis of the History of Professional Tennis. *PLOS ONE*, 6(2), e17249. <https://doi.org/10.1371/journal.pone.0017249>
- Schoch, D. (2018). Centrality without indices: Partial rankings and rank probabilities in networks. *Social Networks*, 54, 50–60. <https://doi.org/10.1016/j.socnet.2017.12.003>
- Schoch, D., & Brandes, U. (2015). Stars, neighborhood inclusion, and network centrality. *SIAM Workshop on Network Science*.
- Schoch, D., & Brandes, U. (2016). Re-conceptualizing centrality in social networks. *European Journal of Applied Mathematics*, 27(6), 971–985. <https://doi.org/10.1017/S0956792516000401>
- Schoch, D., Valente, T. W., & Brandes, U. (2017). Correlations among centrality indices and a class of uniquely ranked graphs. *Social Networks*, 50(Supplement C), 46–54. <https://doi.org/10.1016/j.socnet.2017.03.010>
- Simon, J., & Aditya, K. (2015). *influenceR: Software tools to quantify structural importance of nodes in a network* [Manual].