

node *centrality*

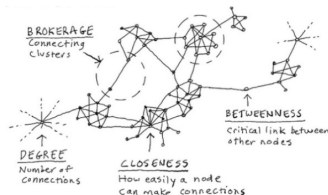
introduction to *network analysis in Python* (*NetPy*)

Lovro Šubelj
University of Ljubljana
10th Dec 2019

centrality *measures*

which *nodes* are most *important*?

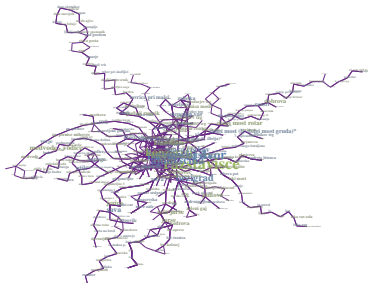
- *node centrality measures* for (*un*)*directed* networks
 - *clustering coefficients* [WS98, SV05, dNMB05]
 - *distance-based* centrality [Fre77, FBW91, New05]
 - *spectral analysis* centrality [Kat53, Bon87, BP98]



- *link analysis algorithms* for *directed* networks

networkology *LPP*

- partial *LPP public bus transport network**
- $n = 416$ bus stops with $\langle k \rangle = 5.62$ connections
- *giant component* 95.4% nodes (6 components)
- “*small-world*” with $\langle C \rangle = 0.09$ and $\langle d \rangle = 14.26$
- “*scale-free*” with $\gamma = 2.62$ for cutoff $k_{min} = 5$



* reduced to largest connected component

centrality *clustering*

important *nodes* are *strongly embedded*

- for *undirected* G *clustering coefficient* C [WS98] of i is
 - t_i is number of *linked neighbors* or *triangles* of i

$$C_i = \frac{2t_i}{k_i(k_i-1)} \quad C_i = 0 \text{ for } k_i \leq 1$$

- μ -*corrected clustering coefficient* C^μ [dNMB05] of i is
 - μ is *maximum* number of *triangles* over *links*

$$C_i^\mu = \frac{2t_i}{k_i\mu} \quad C_i^\mu = 0 \text{ for } k_i = 0$$

centrality *closeness*

important *nodes* are *close to other* nodes

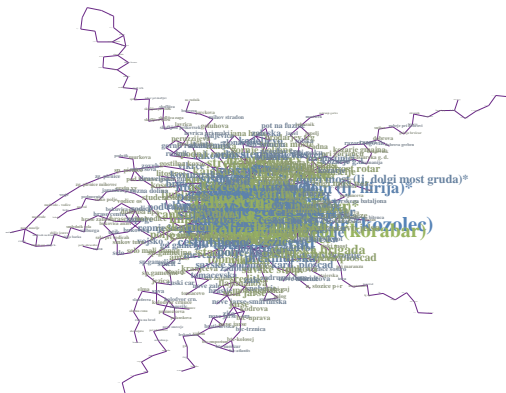
- for (*un*)*directed* G *closeness centrality* ℓ^{-1} [New10] of i is
 - d_{ij} is (*un*)*directed distance* between i and j
 - $d_{ij} = \infty$ for nodes in *different components*

$$\ell_i^{-1} = \frac{1}{n-1} \sum_{j \neq i} \frac{1}{d_{ij}}$$

- ℓ^{-1} spans *small range* in *small-world* networks

networkology *closeness*

- *closeness centrality* ℓ^{-1} in partial LPP network[§]
- *highest* $\ell_i^{-1} = 0.208$ node is *Gosposvetska* with $k_i = 14$



[§] reduced to simple undirected graph

centrality *betweenness*

important *nodes* are *bridges for other* nodes

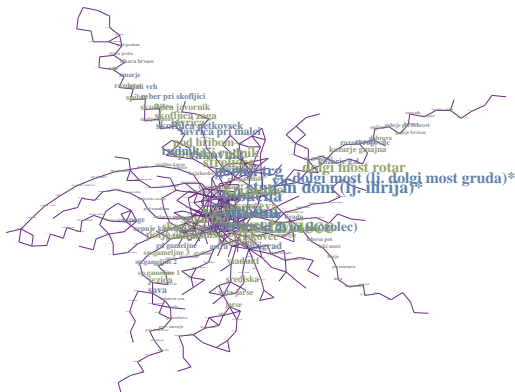
- for (un)directed G *betweenness centrality* σ [Fre77] of i is
 - g_{st} is number of *shortest paths between* s and t
 - g_{st}^i is number of *such shortest paths through* i

$$\sigma_i = \frac{1}{n^2} \sum_{st} \frac{g_{st}^i}{g_{st}}$$

- σ considers *only shortest paths* [FBW91, New05]

networkology *betweenness*

- *betweenness centrality* σ in partial LPP network ¶
- *highest* $\sigma_i = 0.235$ node is *Razstavišče* with $k_i = 11$



¶ reduced to simple undirected graph

centrality *degrees*

important *nodes* are *linked by many* nodes

— for *undirected* G *degree centrality* d of i is

$$d_i = \frac{1}{n-1} \sum_{j \neq i} A_{ij} = \frac{k_i}{n-1}$$

— in *directed* G *in-degree centrality* d^{in} of i is

$$d_i^{in} = \frac{1}{n-1} \sum_{j \neq i} A_{ji} = \frac{k_i^{in}}{n-1}$$

— in *directed* G *out-degree centrality* d^{out} of i is

$$d_i^{out} = \frac{1}{n-1} \sum_{j \neq i} A_{ij} = \frac{k_i^{out}}{n-1}$$

centrality *eigenvector*

important *nodes* are *linked by important* nodes

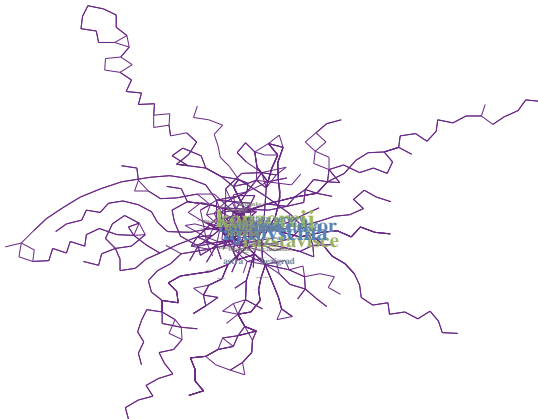
- for (*un*)*directed* G *eigenvector centrality* e [Bon87] of i is
 - e is *leading eigenvector* v_1 of A with *eigenvalue* λ_1^{-1}

$$e_i = \lambda_1^{-1} \sum_j A_{ij} e_j$$

- in *directed* G $e = 0$ for $k^{in} = 0$ *nodes etc.*

networkology *eigenvector*

- *eigenvector centrality* e in partial LPP network
- *highest* $e_i = 0.082$ node is *Konzorcij* with $k_i = 30$



centrality *Katz*

nodes get small amount of *importance* for free

- for (un)directed G *Katz centrality* z [Kat53] of i is
 - α and β are *appropriate positive constants*

$$z_i = \alpha \sum_j A_{ij} z_j + \beta$$

- for *convenience* $\beta = 1$ whereas $\alpha < \lambda_1^{-1}$
 - λ_1 is *leading eigenvalue* of A for *eigenvector* v_1

centrality *PageRank*

nodes distribute equal amount of *importance*

- for (*un*)*directed* *G* *PageRank centrality* p [BP98] of i is
 - α and β are *appropriate positive constants*

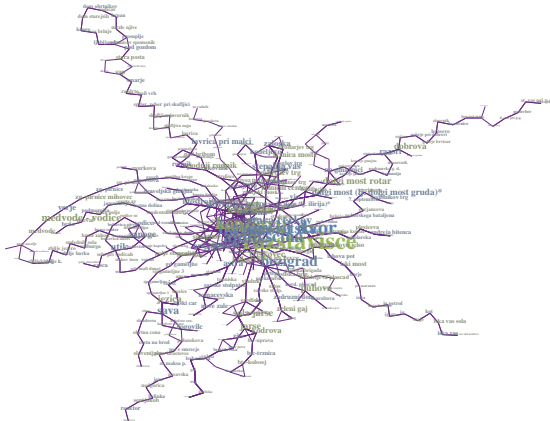
$$p_i = \alpha \sum_j A_{ij} \frac{p_j}{k_j} + \beta$$

- for *convenience* $\beta = \frac{1-\alpha}{n}$ whereas $\alpha = 0.85$

see PageRank algorithm *NetLogo* demo

networkology *PageRank*

- *PageRank centrality* p in partial LPP network
- *highest* $p_i = 0.011$ node is *Razstavišče* with $k_i = 41$



centrality *overview*

which *nodes* are most *important*?

Figure 1: Overview of network centrality measures. The figure displays a grid of 100 centrality measures, categorized by their primary method. The measures are arranged in a grid with rows and columns labeled by their primary method. A legend on the right explains the color coding for the primary methods.

Primary Method	Measure	Secondary Method
Distance-based	1 DC	Distance-based
Distance-based	2 EC	Distance-based
Distance-based	3 BC	Betweenness-based
Distance-based	4 CC	Betweenness-based
Distance-based	5 SC	Subgraph-based
Distance-based	6 CCoef	Clustering Coefficient-based
Distance-based	7 CCoef ⁻¹	Clustering Coefficient-based
Distance-based	8 MNC	Max. Neighbors
Distance-based	9 ECoef	Edge Clustering Coefficient-based
Distance-based	10 PR	PageRank
Distance-based	11 RL	Rang- und Inverse Betweenness
Distance-based	12 IC	Information Centrality
Distance-based	13 IIIA	Subgraph-based
Distance-based	14 IVB	Subgraph-based
Distance-based	15 VB	Subgraph-based
Distance-based	16 VIB	Subgraph-based
Distance-based	17 VIIIB	Subgraph-based
Distance-based	18 VIIIIB	Subgraph-based
Distance-based	19 IB	Subgraph-based
Distance-based	20 IIB	Subgraph-based
Distance-based	21 BN	Betweenness-based
Distance-based	22 RC	Radiality Centrality
Distance-based	23 IG	Integration
Distance-based	24 DCox	Distance-based
Distance-based	25 BCox	Distance-based
Distance-based	26 CCox	Distance-based
Distance-based	27 EOCox	Distance-based
Distance-based	28 KSx	Distance-based
Distance-based	29 PRx	Distance-based
Distance-based	30 IGx	Distance-based
Distance-based	31 RCx	Distance-based
Distance-based	32 DCx	Distance-based
Distance-based	33 BCx	Distance-based
Distance-based	34 Cx	Distance-based
Distance-based	35 LAC	Local Average Clustering Coefficient
Distance-based	36 DMNC	Distance-based
Distance-based	37 SECox	Distance-based
Distance-based	38 LR	LeaderRank
Distance-based	39 RWBC	RandomWalk Betweenness
Distance-based	40 RWCC	RandomWalk Closeness
Distance-based	41 CC _{2,3,4}	Distance-based
Distance-based	42 ECoef _{2,3,4}	Distance-based
Distance-based	43 PRox	Distance-based
Distance-based	44 KSx	Distance-based
Distance-based	45 CCoef	Distance-based
Distance-based	46 RCx	Distance-based
Distance-based	47 IGx	Distance-based
Distance-based	48 DCBC	Distance-based
Distance-based	49 BCCC	Distance-based
Distance-based	50 CKS	Distance-based
Distance-based	51 KSPR	Distance-based
Distance-based	52 DCPR	Distance-based
Distance-based	53 β	Distance-based
Distance-based	54 SC ₂	Distance-based
Distance-based	55 NC	Distance-based
Distance-based	56 EC ₂	Distance-based
Distance-based	57 σ	Distance-based
Distance-based	58 ECC	Distance-based
Distance-based	59 WDC	Distance-based
Distance-based	60 DECC	Distance-based
Distance-based	61 CECC	Distance-based
Distance-based	62 BECC	Distance-based
Distance-based	63 KECC	Distance-based
Distance-based	64 PRECC	Distance-based
Distance-based	65 IGECC	Distance-based
Distance-based	66 DCC	Distance-based
Distance-based	67 BCK	Distance-based
Distance-based	68 CCPR	Distance-based
Distance-based	69 KSG	Distance-based
Distance-based	70 DCCoef	Distance-based
Distance-based	71 SC ₃	Distance-based
Distance-based	72 LI	Distance-based
Distance-based	73 EC ₃	Distance-based
Distance-based	74 BC _{2,3,4}	Distance-based
Distance-based	75 EC ₂	Distance-based
Distance-based	76 SDC	Distance-based
Distance-based	77 DCR	Distance-based
Distance-based	78 CCRC	Distance-based
Distance-based	79 BCRC	Distance-based
Distance-based	80 KSRC	Distance-based
Distance-based	81 PRRC	Distance-based
Distance-based	82 IGRC	Distance-based
Distance-based	83 DCKS	Distance-based
Distance-based	84 BCPR	Distance-based
Distance-based	85 CCIG	Distance-based
Distance-based	86 DCPR	Distance-based
Distance-based	87 BCIG	Distance-based
Distance-based	88 ECRC	Distance-based
Distance-based	89 BCCoef	Distance-based
Distance-based	90 SC ₄	Distance-based
Distance-based	91 EC ₄	Distance-based
Distance-based	92 FC	Distance-based
Distance-based	93 FD	Distance-based
Distance-based	94 US	Distance-based
Distance-based	95 DIS	Distance-based
Distance-based	96 ASS	Distance-based
Distance-based	97 DAM	Distance-based
Distance-based	98 UC	Distance-based
Distance-based	99 EI	Distance-based
Distance-based	100 CM	Distance-based
Distance-based	101 NaC	Distance-based
Distance-based	102 MC	Distance-based
Distance-based	103 HGI	Distance-based
Distance-based	104 HYP	Distance-based
Distance-based	105 HC	Distance-based

Legend:

- Distance-based
- Betweenness-based
- Subgraph-based
- Clustering Coefficient-based
- Edge Clustering Coefficient-based
- Spectral-based
- Miscellaneous

centrality *references*



Phillip Bonacich.

Power and centrality: A family of measures.
American Journal of Sociology, 92(5):1170–1182, 1987.



S. Brin and L. Page.

The anatomy of a large-scale hypertextual Web search engine.
Comput. Networks ISDN, 30(1-7):107–117, 1998.



Wouter de Nooy, Andrej Mrvar, and Vladimir Batagelj.

Exploratory Social Network Analysis with Pajek.
Cambridge University Press, Cambridge, 2005.



David Easley and Jon Kleinberg.

Networks, Crowds, and Markets: Reasoning About a Highly Connected World.
Cambridge University Press, Cambridge, 2010.



Ernesto Estrada and Philip A. Knight.

A First Course in Network Theory.
Oxford University Press, 2015.



Linton C. Freeman, Stephen P. Borgatti, and Douglas R. White.

Centrality in valued graphs: A measure of betweenness based on network flow.
Soc. Networks, 13(2):141–154, 1991.



L. Freeman.

A set of measures of centrality based on betweenness.
Sociometry, 40(1):35–41, 1977.



Leo Katz.

A new status index derived from sociometric analysis.
Psychometrika, 18(1):39–43, 1953.

centrality *references*



M. E. J. Newman.

A measure of betweenness centrality based on random walks.
Soc. Networks, 27(1):39–54, 2005.



Mark E. J. Newman.

Networks: An Introduction.
Oxford University Press, Oxford, 2010.



Sara Nadiv Soffer and Alexei Vázquez.

Network clustering coefficient without degree-correlation biases.
Phys. Rev. E, 71(5):057101, 2005.



D. J. Watts and S. H. Strogatz.

Collective dynamics of 'small-world' networks.
Nature, 393(6684):440–442, 1998.