

node *centrality*

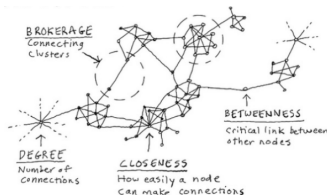
introduction to *network analysis in Python* (*NetPy*)

Lovro Šubelj  
University of Ljubljana  
19th Sep 2019

# centrality *measures*

which *nodes* are most *important*?

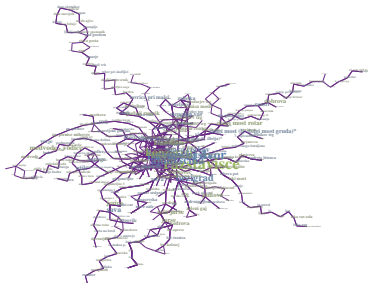
- *node centrality measures* for (*un*)*directed* networks
  - *clustering coefficients* [WS98, SV05, dNMB05]
  - *distance-based* centrality [Fre77, FBW91, New05]
  - *spectral analysis* centrality [Kat53, Bon87, BP98]



- *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  $i$

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

- $\mu$ -*corrected clustering coefficient*  $C^\mu$  [dNMB05] of  $i$  is
  - $\mu$  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

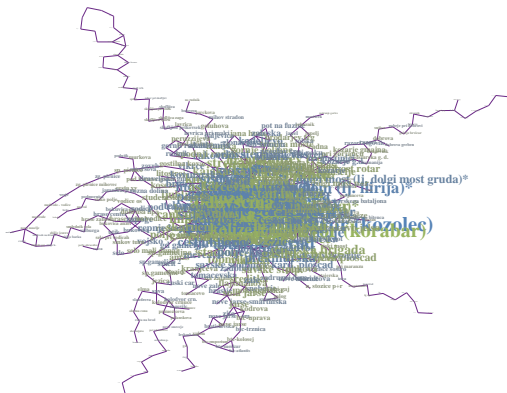
- for (*un*)*directed*  $G$  *closeness centrality*  $\ell^{-1}$  [New10] of  $i$  is
  - $d_{ij}$  is (*un*)*directed distance* between  $i$  and  $j$
  - $d_{ij} = \infty$  for nodes in *different components*

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

- $\ell^{-1}$  spans *small range* in *small-world* networks

# networkology *closeness*

- *closeness centrality*  $\ell^{-1}$  in partial LPP network<sup>§</sup>
- *highest*  $\ell_i^{-1} = 0.208$  node is *Gosposvetska* with  $k_i = 14$



<sup>§</sup> reduced to simple undirected graph



## centrality *betweenness*

important *nodes* are *bridges for other* nodes

- for (un)directed  $G$  *betweenness centrality*  $\sigma$  [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*  $i$

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

- $\sigma$  considers *only shortest paths* [FBW91, New05]



# 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

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

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

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



## centrality *eigenvector*

important *nodes* are *linked by important* nodes

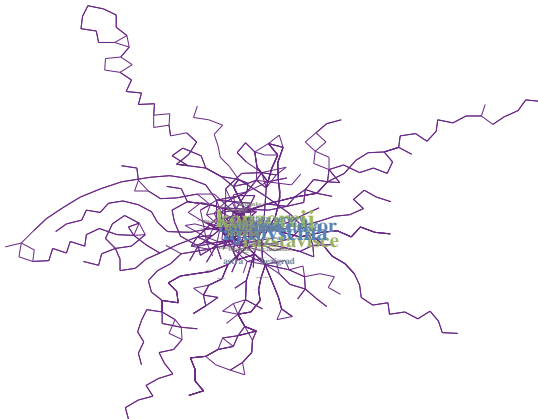
- for (*un*)*directed*  $G$  *eigenvector centrality*  $e$  [Bon87] of  $i$  is
  - $e$  is *leading eigenvector*  $v_1$  of  $A$  with *eigenvalue*  $\lambda_1^{-1}$

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

- 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
  - $\alpha$  and  $\beta$  are *appropriate positive constants*

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

- for *convenience*  $\beta = 1$  whereas  $\alpha < \lambda_1^{-1}$ 
  - $\lambda_1$  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
  - $\alpha$  and  $\beta$  are *appropriate positive constants*

$$p_i = \alpha \sum_j A_{ij} \frac{p_j}{k_j^{\text{out}}} + \beta_j$$

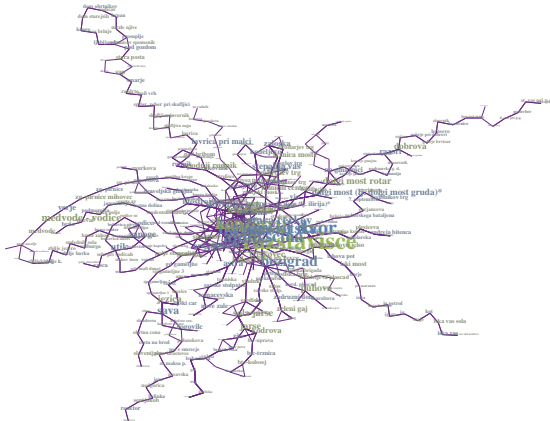
- for *convenience*  $\beta = \frac{1-\alpha}{n}$  whereas  $\alpha = 0.85$

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*?

1 IA													18 VIIIA																								
1	DC	34											5	9	8																						
	Degree Centrality												SC	Subgraph Centrality	CCoef	Clustering Coefficient	CCoef <sup>-1</sup>	inverse CCoef	MNC	max. neigh. comp.	ECCoef	edge clustering coefficient	PR	PageRank													
2	BC	12	4	13											6	1	9	3	18	2																	
	Betweenness Centrality			Closeness Centrality											SCo	odd Subgraph Centrality	LAC	loc. avg. connectivity	DMNC	dens. max. neigh. comp.	SECcof	sum of ECCoef	LR	LeaderRank													
3	RL	2	12	8											11	1	14	2	15	3	16	3	17	2	18	1											
	Range-Inv. Betweenness Centrality			Information Centrality											CCoef <sup>-1</sup>	SCo	LAC	loc. avg. connectivity	DMNC	dens. max. neigh. comp.	SECcof	sum of ECCoef	LR	LeaderRank													
4	BN	4	29	2											21	1	57	1	90	1	58	1	61	1	62	1	63	1	64	1	65	1	66	1	67	1	
	Betweenness Centrality			Radiality Centrality											IG	Integration	DCoef	BCoef	CCoef	ECCoef	KScof	PRCoef	IGCoef	RCCoef	DCoef	BCoef	SC <sub>1</sub>	odd Subgraph Centrality	LAC	loc. avg. connectivity	DMNC	dens. max. neigh. comp.	SECcof	sum of ECCoef	LR	LeaderRank	
5	RWBC	30	1	93	1	93	1	90	3	94	1	90	1	97	3	96	1	46	1	47	1	48	1	49	1	50	1	51	1	52	1	53	1	54	1		
	Random-Walk Betweenness			RWCC	CC <sub>2,3,4</sub>	ECCoef	CCoef	PRCoef	KScof	CCoef	RCCoef	IGCoef	DCBC	BCCoef	CCS	KSPR	DCPR	$\beta$	Bipartivity	SC <sub>2</sub>	3-localized SC	SC <sub>3</sub>	3-localized SC	LI	Lobby Index	EC <sub>3</sub>	4-localized SC	EC <sub>4</sub>	4-localized EC								
6	$\sigma$	2	56	2	67	71	1	72	1	73	1	74	1	75	1	76	1	77	1	78	1	79	1	80	1	81	1	82	1			83	1	84	1	85	1
	Stress Centrality			ECC	WDC	Weighted Degree	DCEC	CCoef	BCEC	KSECC	PRECC	IGEC	DBCC	BCC	CCS	KSPR	DCPR	$\beta$	Bipartivity	SC <sub>2</sub>	3-localized SC	SC <sub>3</sub>	3-localized SC	LI	Lobby Index	EC <sub>3</sub>	4-localized SC	EC <sub>4</sub>	4-localized EC								
7	BC <sub>2,3,4</sub>	1	93	1	90	1	104	1	105	1	106	1	107	1	108	1	110	1	111	1	112	1	113	1	114	1	115	1	116	1	117	1	118	1			
	2,3,4-localized BC			SDC	Sphere Degree Centrality	DCRC	CCRC	BCRC	KSRC	PRRC	IGRC	DKS	BCPR	CCIG	DCPR	BCIG	ECRC	BCCoef	4-localized SC	SC <sub>4</sub>	4-localized SC	EC <sub>4</sub>	4-localized EC														

Z

mass

C

Name

(Hybrid)

22	2	23	2	24	2	25	1	26	1	27	1	317	1
FC	Functional Centrality	FD	Functional Diversity	US	UniScore	DIS	Pairwise Dis-connectivity	ASS	Asymptotic Mixing	DAM	Damage	UC	United compl. Centrality

28	1	29	1	30	1	31	1	33	1	316	0	318	1
EI	Essentiality Index	CM	Complexity Measure	NoC	Normalized $\alpha$ Centrality	MC	Modular Centrality	HGI	Harary Graph Information	HYP	Hypocritic Index	HC	Homocentric Centrality

Betweenness-based

Distance-based

Linear Combinations

Subgraph-based

Clustering Coefficient-based

Edge Clustering Coefficient-based

Spectral-based

Miscellaneous

@David Schoch (University of Konstanz)

©David Schoch (University of Konstanz)

# 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.