

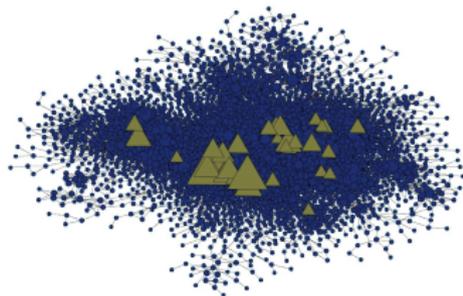
applications *scientometrics*

introduction to *network analysis* (*ina*)

Lovro Šubelj  
University of Ljubljana  
spring 2025/26

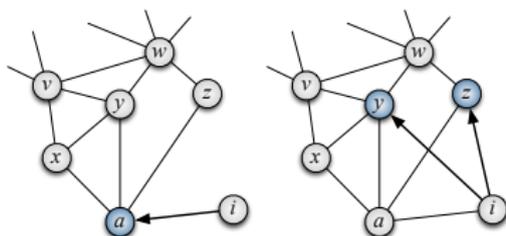
# models *overview*

- *random graph models* for reasoning about *network structure*
  - node mixing [New02], scale-free [BA99] and small-world [WS98] networks
- *generative models* for reasoning also about *network evolution*
  - node aging [HS05], densification laws and shrinking diameters [LKF07]
- next *realistic generative models* of *paper citation networks*



# models *forest fire*

- $G(n, p)$  *forest fire* model [LKF07]
- each new *node*  $i$  *forms links* as follows
  1.  $i$  selects *random ambassador*  $a$  and *links to*  $a$
  2.  $i$  selects  $x_p$  *neighbors*  $a_1, \dots, a_{x_p}$  and *links to*  $a_i$
  3.  $a_1, \dots, a_{x_p}$  are taken as *new ambassadors* of  $i$

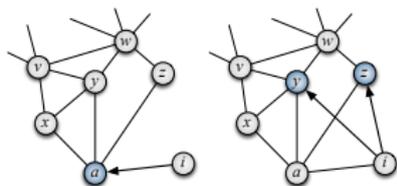


---

\*  $p$  is *burning probability* &  $x_p$  from *geometric distribution*  $G(\frac{p}{1-p})$

## models *dynamics*

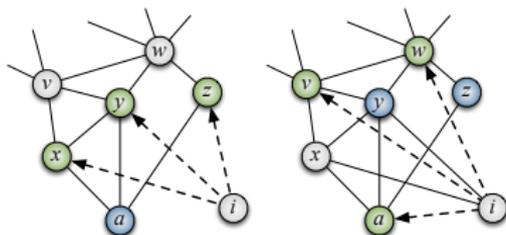
- $G(n, p)$  *author dynamics* model [LKF07]
- each new *paper*  $i$  *forms citations* as follows
  1.  $i$  selects *random read paper*  $a$  and *cites*  $a$
  2.  $i$  selects  $x_p$  *references*  $a_1, \dots, a_{x_p}$  and *cites*  $a_i$
  3.  $a_1, \dots, a_{x_p}$  are taken as *new reading* for  $i$



- then *authors read all cited papers* (vice-versa)
- but only  $\approx 20\%$  of *cited papers are read* [SR03]

## models *citation*

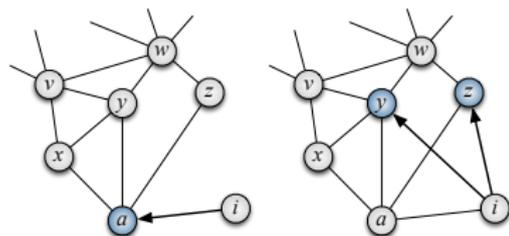
- $G(n, p, q)$  *citation* model [ŠB13]
- each new *node i forms links* as follows
  1.  $i$  selects *random ambassador a* as before
  2.  $i$  selects  $x_p$  *neighbors*  $a_1, \dots, a_{x_p}$  as before
  3.  $i$  selects  $x_q$  *neighbors*  $l_1, \dots, l_{x_q}$  and *links to*  $l_j$
  4.  $a_1, \dots, a_{x_p}$  are taken as *new ambassadors* of  $i$



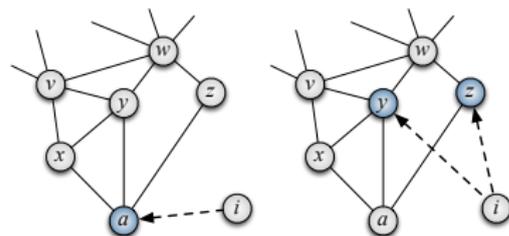
---

<sup>†</sup>  $q$  is *linking probability* &  $x_q$  from *geometric distribution*  $G(\frac{q}{1-q})$

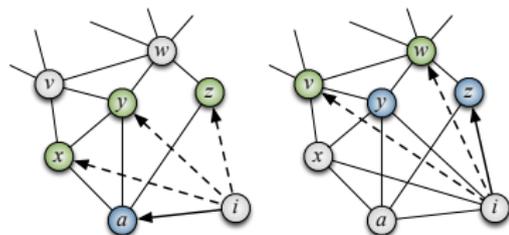
# models *undirected*



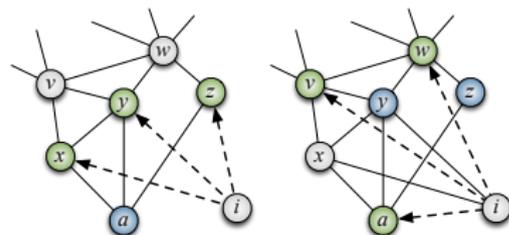
*forest fire* model [LKF07]  $A = L$



*butterfly* model [MAF08]  $A \supseteq L$



*copying* model [KR05]  $A \subseteq L$

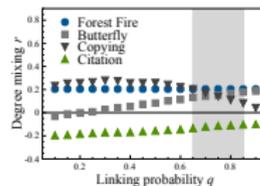
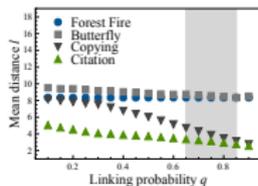
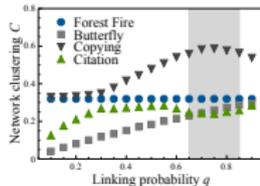
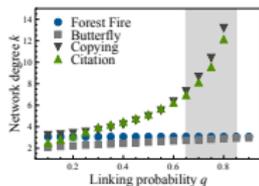
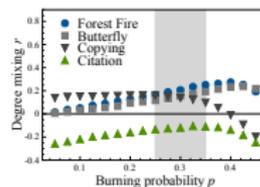
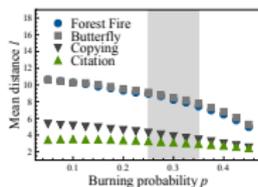
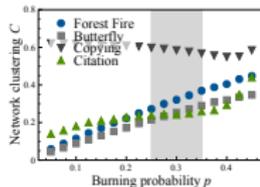
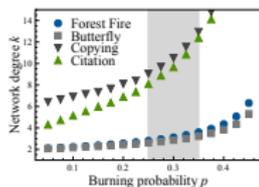


*citation* model [ŠB13]  $A$  &  $L$  arbitrary

‡  $A$  is set of *ambassadors* &  $L$  is set of *linked nodes*

# models *comparison*

- *all models* generate *modular scale-free/small-world* networks
- *only citation* model generates *degree disassortative* networks



§ shaded regions show likely parameter values [LJT<sup>+</sup>11]

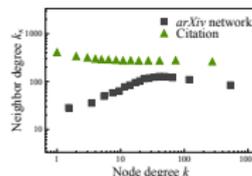
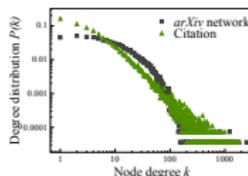
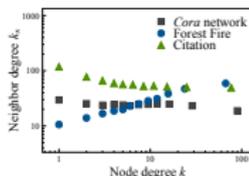
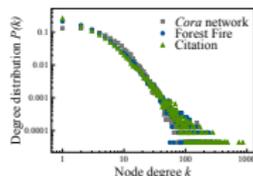
# models *examples*

— *modeling* of *Cora citation* network [MNRS00]

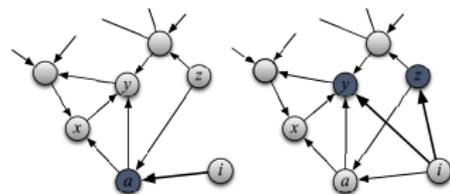
model	nodes $n$	degree $\langle k \rangle$	mixing $r$
<i>forest fire</i>	23,166	7.669	0.211
<i>citation</i>	23,166	7.760	-0.047
	23,166	7.697	-0.055

— *modeling* of *arXiv citation* network [LKF05]

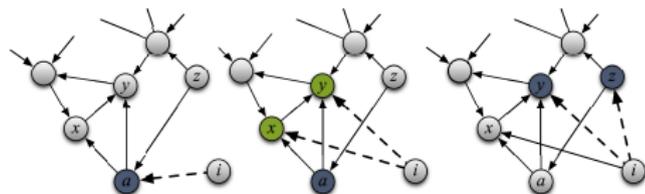
model	nodes $n$	degree $\langle k \rangle$	mixing $r$
<i>citation</i>	27,400	25.598	-0.068
	27,400	25.695	-0.030



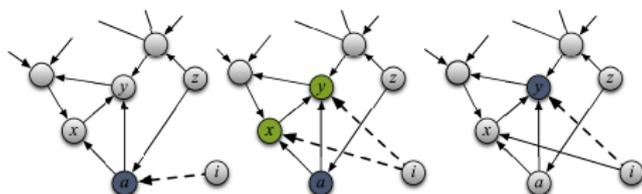
# models *directed*



*forest fire* model [LKF07]



*citation* model [ŠŽB14]



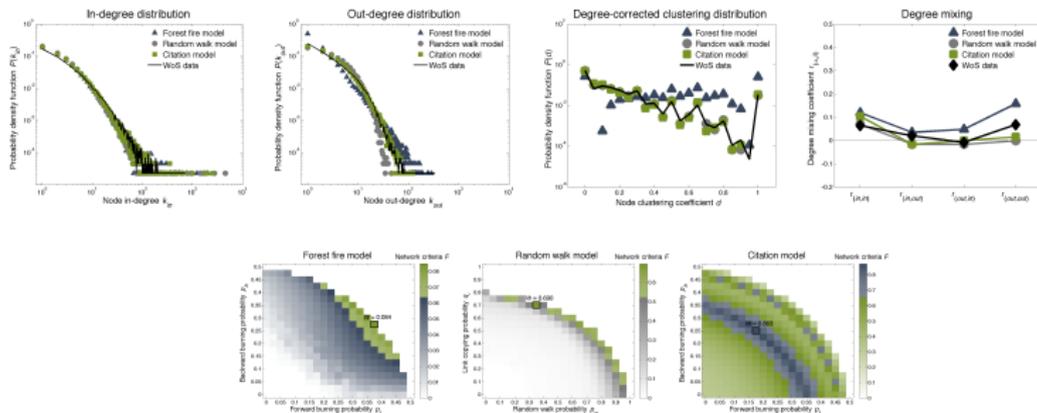
*random walk* model [Váz03]

---

$p_{fwd}/p_{bck}$  are *burning probabilities* &  $q_{amb}/q_{lnk}$  are *linking probabilities*

— modeling of *WoS citation* network [ŠŽB14]

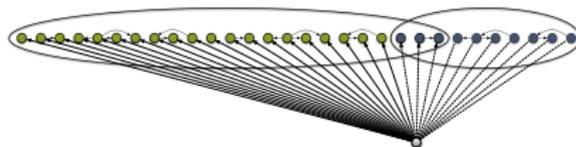
model	degree $\langle k \rangle$	power-law $\gamma$		mixing $r$		clustering $C$	
<i>forest fire</i>	7.88	2.51	2.45	0.13	0.17	0.40	0.52
<i>random walk</i>	7.91	2.61	4.68	0.07	0.00	0.11	0.13
<i>citation</i>	7.97	2.58	3.31	0.12	0.01	0.11	0.13
	8.05	2.47	3.35	0.06	0.07	0.11	0.14



## — evolution/comparison of WoS *citation* network [ŠŽB14]

period	paper citation		paper study		paper discovery by		
	# cites	% copied	# reads	% cited	% papers	% services	% others
1970–1980	2.23	52.1%	3.39	33.5%	41.4%	0.0%	58.5%
1980–1990	2.62	65.1%	2.96	33.0%	48.3%	1.1%	50.6%
1990–2000	3.42	81.6%	2.38	29.0%	40.3%	23.2%	36.5%
2000–2010	5.06	83.6%	2.90	32.2%	40.7%	27.5%	31.7%

field	paper citation		paper study		paper discovery by		
	# cites	% copied	# reads	% cited	% papers	% services	% others
theory	2.93	79.7%	1.47	45.2%	74.7%	0.5%	24.9%
AI	4.52	87.3%	1.47	40.9%	25.8%	47.6%	26.6%
software	2.78	81.5%	1.58	36.4%	68.8%	2.0%	29.2%
cybernetics	2.18	69.6%	1.59	43.2%	24.5%	37.8%	37.6%



$$\| \langle k_{out} \rangle \text{ papers cited} \ \& \ s = \left( 1 - \frac{p_{fwd}}{1 - p_{fwd}} - \frac{p_{bck}}{1 - p_{bck}} \right)^{-1} \text{ papers read} \ \& \ s \frac{q_{lnk}}{1 - q_{lnk}} \text{ citations copied etc.}$$

# models *references*



A.-L. Barabási and R. Albert.  
Emergence of scaling in random networks.  
*Science*, 286(5439):509–512, 1999.



A.-L. Barabási.  
*Network Science*.  
Cambridge University Press, Cambridge, 2016.



Wouter de Nooy, Andrej Mrvar, and Vladimir Batagelj.  
*Exploratory Social Network Analysis with Pajek: Expanded and Revised Second Edition*.  
Cambridge University Press, Cambridge, 2011.



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.



Kamalika Basu Hajra and Parongama Sen.  
Modelling aging characteristics in citation networks.  
*Physica A*, 368(2):575–582, 2005.



P. L. Krapivsky and S. Redner.  
Network growth by copying.  
*Phys. Rev. E*, 71(3):036118, 2005.



Paul J. Laurienti, Karen E. Joyce, Qawi K. Telesford, Jonathan H. Burdette, and Satoru Hayasaka.  
Universal fractal scaling of self-organized networks.  
*Physica A*, 390(20):3608–3613, 2011.

# models *references*



Jure Leskovec, Jon Kleinberg, and C. Faloutsos.

Graphs over time: Densification laws, shrinking diameters and possible explanations.

In *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 177–187, Chicago, IL, USA, 2005.



Jure Leskovec, Jon Kleinberg, and Christos Faloutsos.

Graph evolution: Densification and shrinking diameters.

*ACM Trans. Knowl. Discov. Data*, 1(1):1–41, 2007.



Mary McGlohon, Leman Akoglu, and Christos Faloutsos.

Weighted graphs and disconnected components: Patterns and a generator.

In *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 524–532, New York, NY, USA, 2008.



Andrew Kachites McCallum, Kamal Nigam, Jason Rennie, and Kristie Seymore.

Automating the construction of Internet portals with machine learning.

*Inform. Retrieval*, 3(2):127–163, 2000.



M. E. J. Newman.

Assortative mixing in networks.

*Phys. Rev. Lett.*, 89(20):208701, 2002.



Mark E. J. Newman.

*Networks*.

Oxford University Press, Oxford, 2nd edition, 2018.



Lovro Šubelj and Marko Bajec.

Model of complex networks based on citation dynamics.

In *Proceedings of the WWW Workshop on Large Scale Network Analysis*, pages 527–530, Rio de Janeiro, Brazil, 2013.

# models *references*



M. V. Simkin and V. P. Roychowdhury.

Read before you cite!

*Compl. Syst.*, 14(3):269–274, 2003.



Lovro Šubelj, Slavko Žitnik, and Marko Bajec.

Who reads and who cites? Unveiling author citation dynamics by modeling citation networks.

In *Proceedings of the International Conference on Network Science*, page 1, Berkeley, CA, USA, 2014.



Alexei Vázquez.

Growing network with local rules: Preferential attachment, clustering hierarchy, and degree correlations.

*Phys. Rev. E*, 67(5):056104, 2003.



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

Collective dynamics of 'small-world' networks.

*Nature*, 393(6684):440–442, 1998.