Deconstructing Data Science

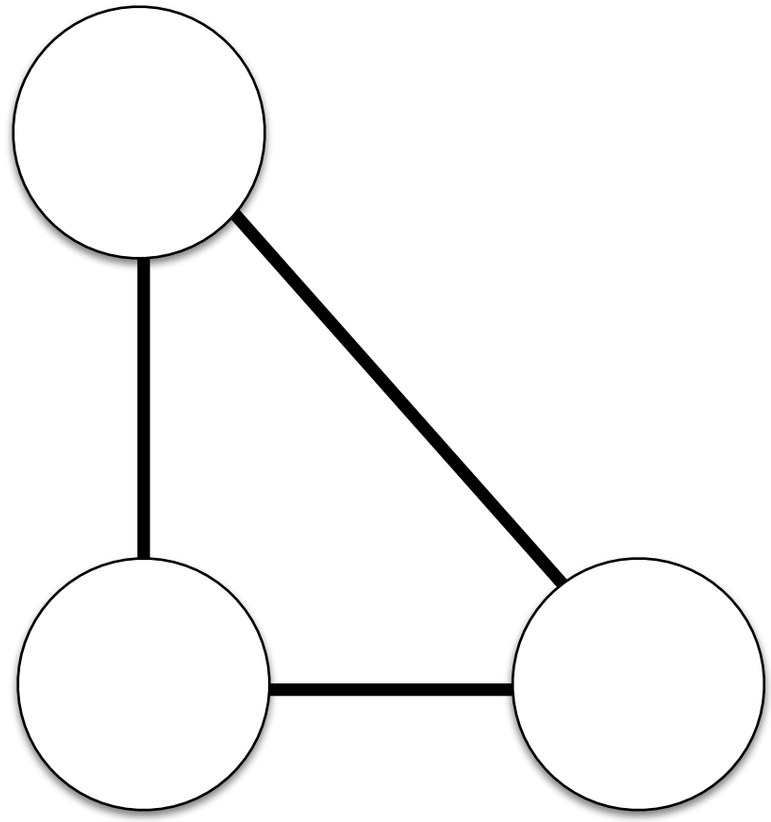
David Bamman, UC Berkeley

Info 290
Lecture 21: Information Diffusion

Apr 11, 2016

Networks

- Nodes
- Edges
- Information



Networks

Network	Nodes	Edges	Information
Social	People		
Internet	Servers		
Citation network	Articles		
Web	Web pages		

Information cascades

- Information effects (herding behavior)
- Direct-benefit effects
- Epidemics

Herding behavior

- Lines outside restaurants/clubs
- Crowd of people looking up (Milgram et al. 1969)

- Inference that observed choices are more powerful than own private information

Urn Game

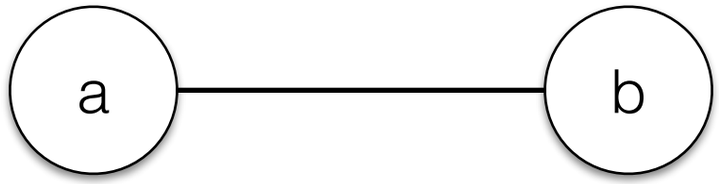
- private information vs. public action

Direct benefit effects

- Direct payoffs for making the same decisions others make

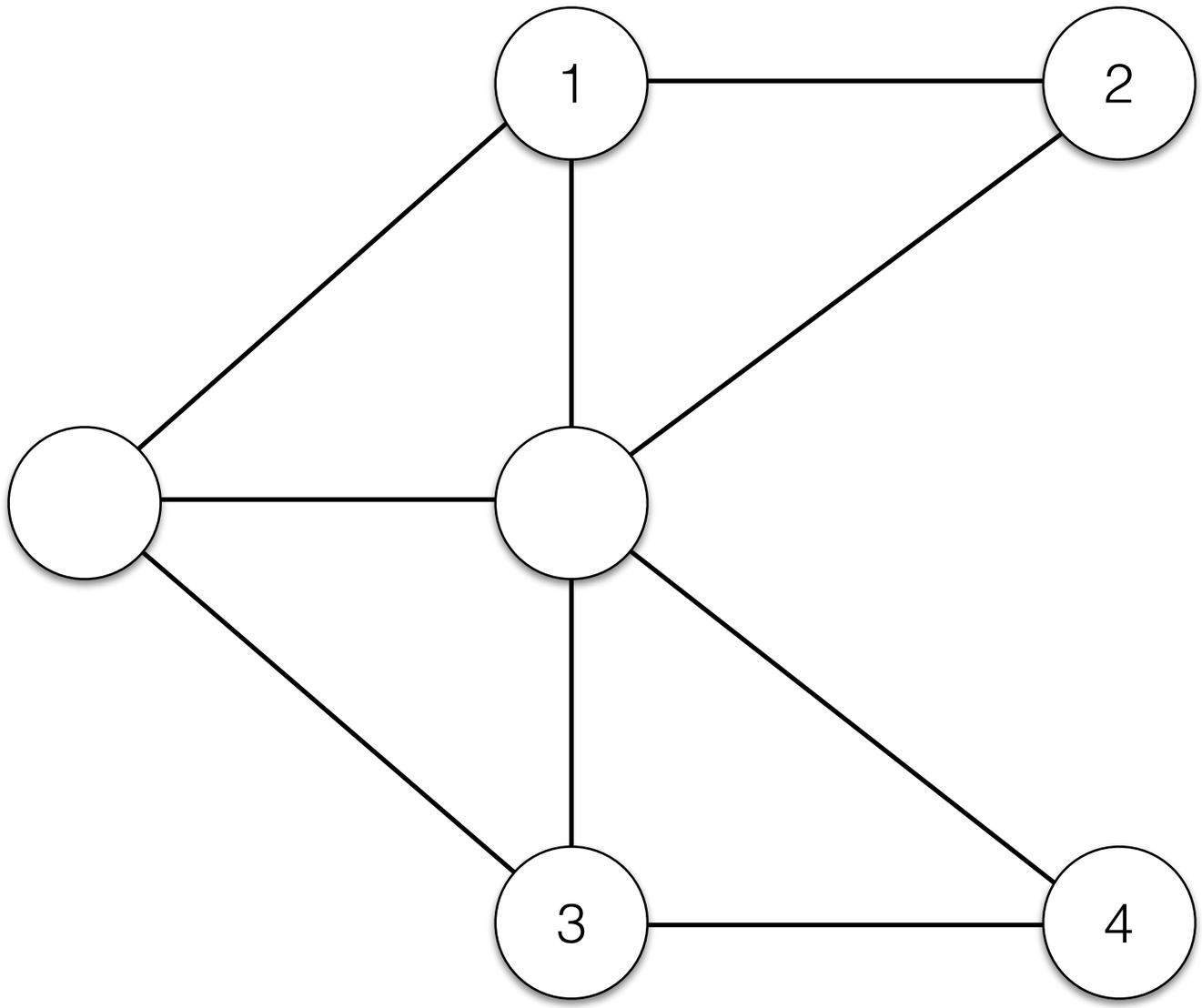
- Social networking sites
- Cell phone providers
- Mac/PC

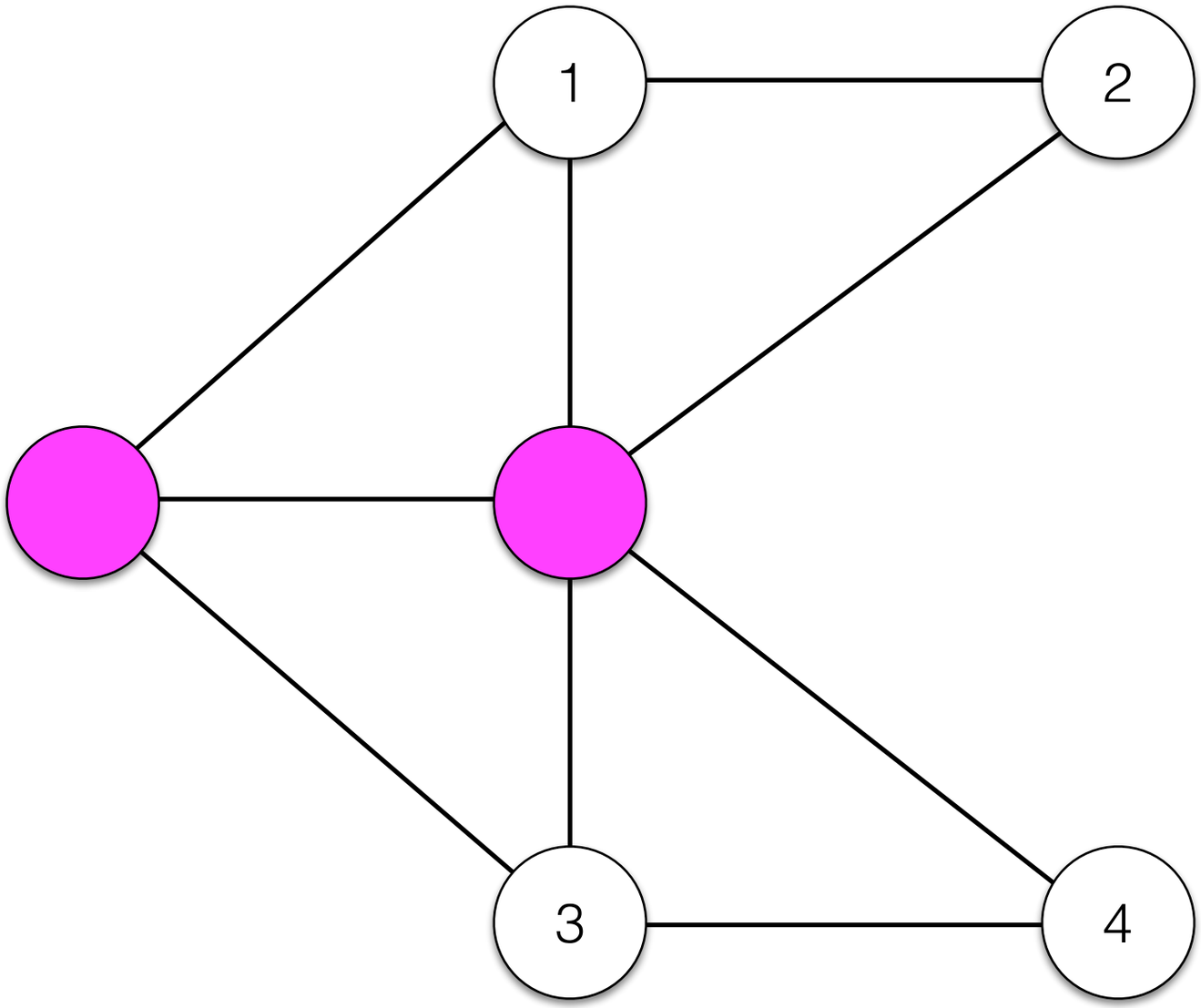
Direct benefit effects



- a and b adopt A, they get a payout of x
- a and b adopt B, they get a payout of y
- otherwise they get a payout of 0

Easley and Kleinberg 2010

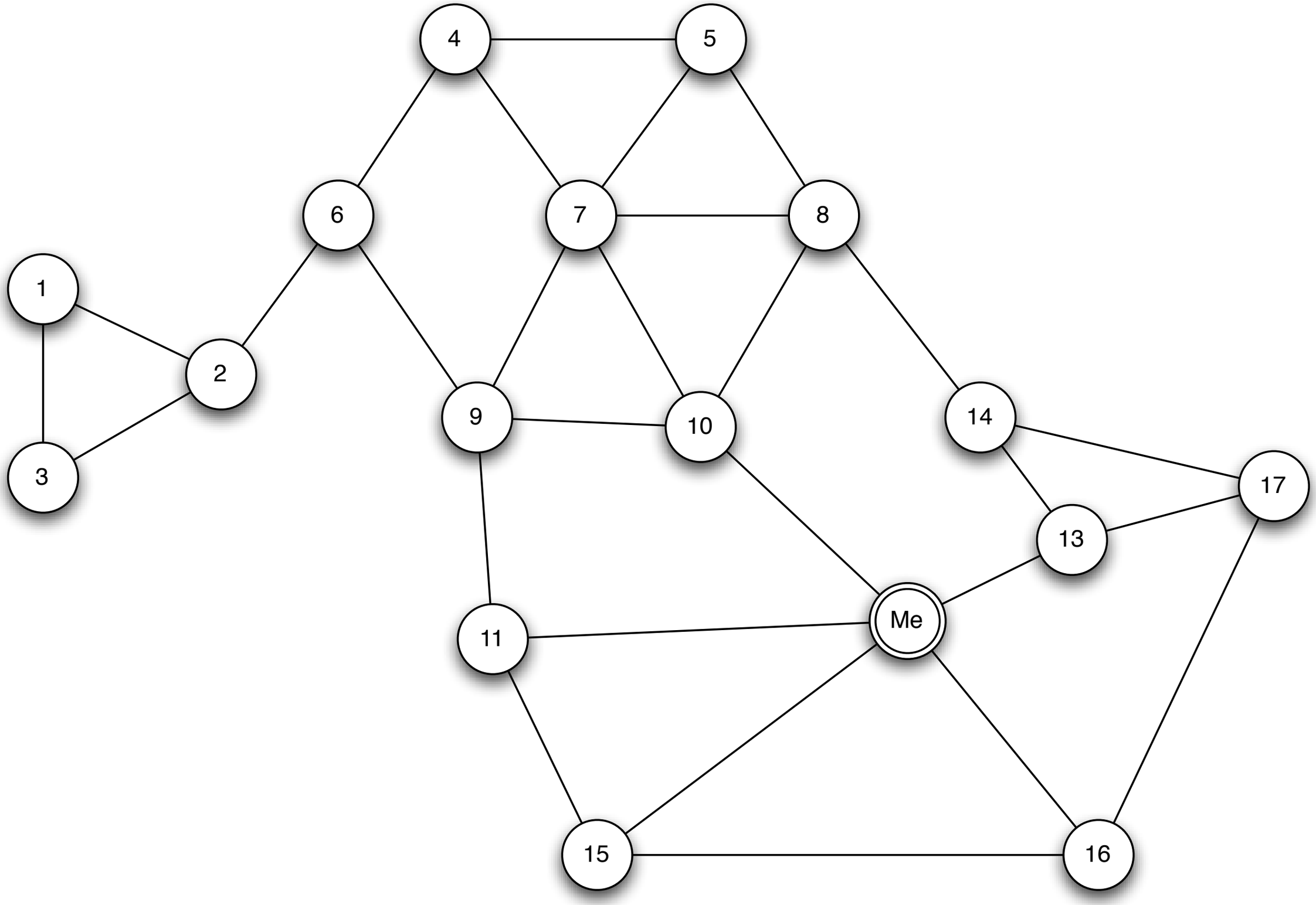

1
2
3
4

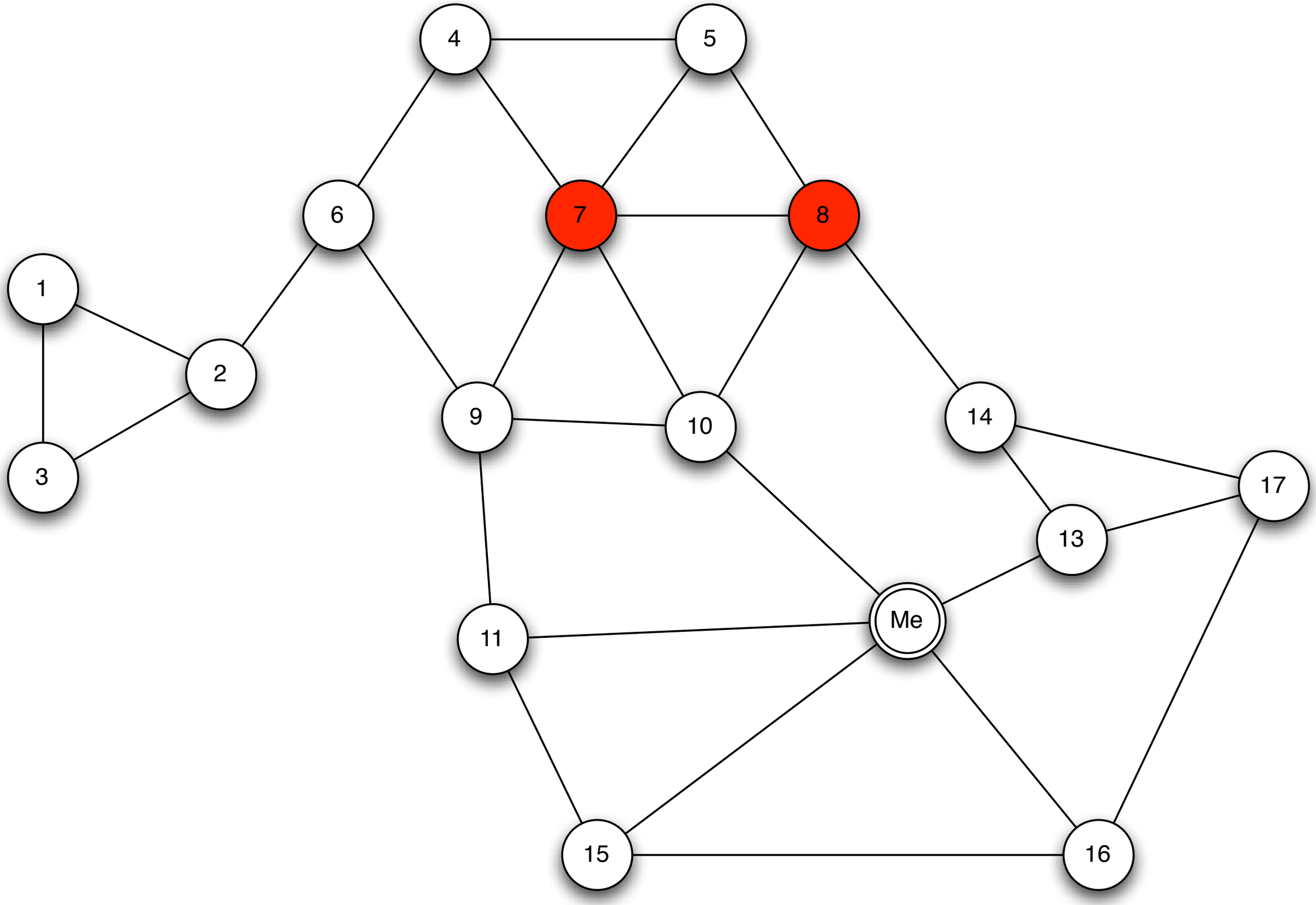


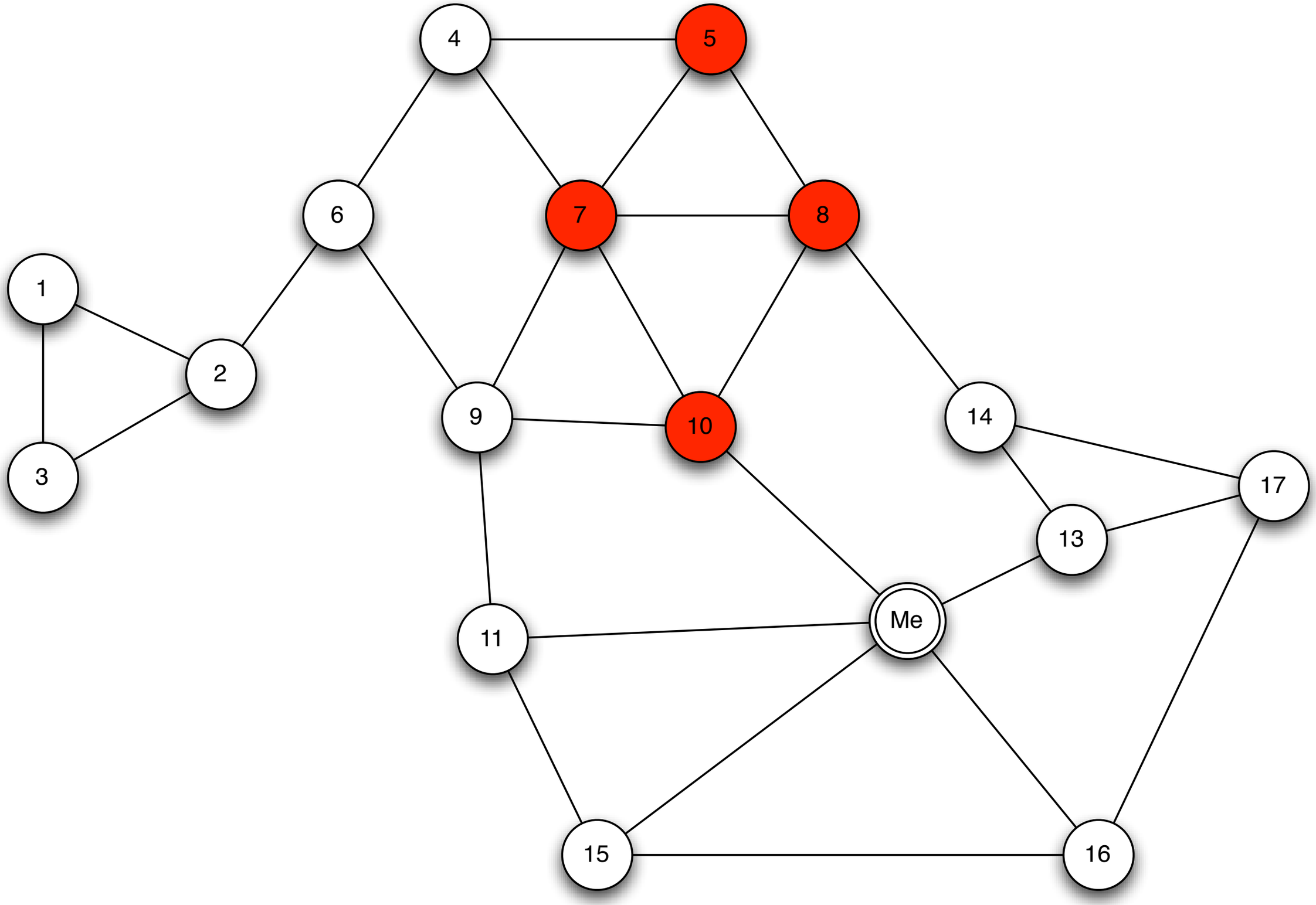
a=3

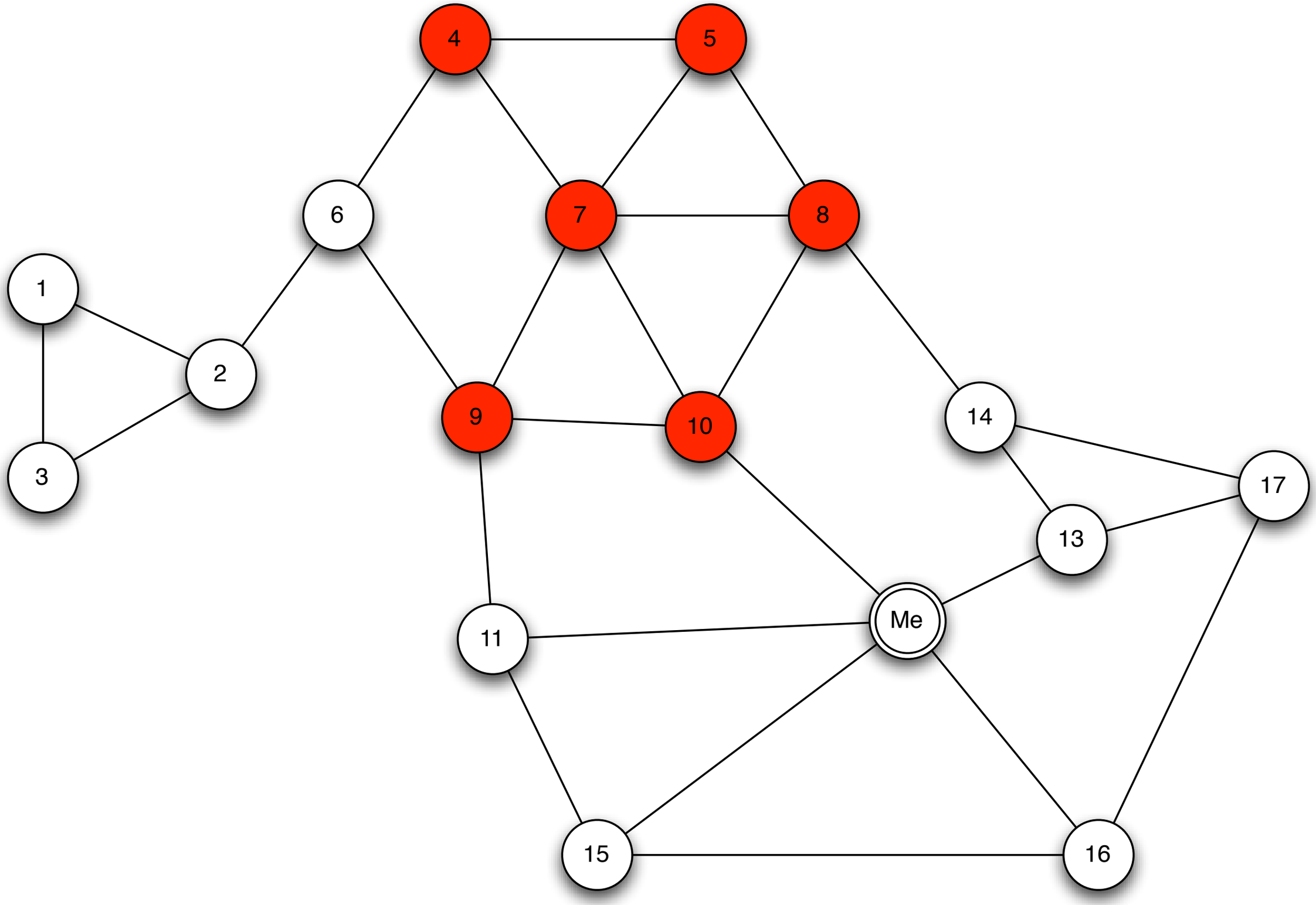
b=2

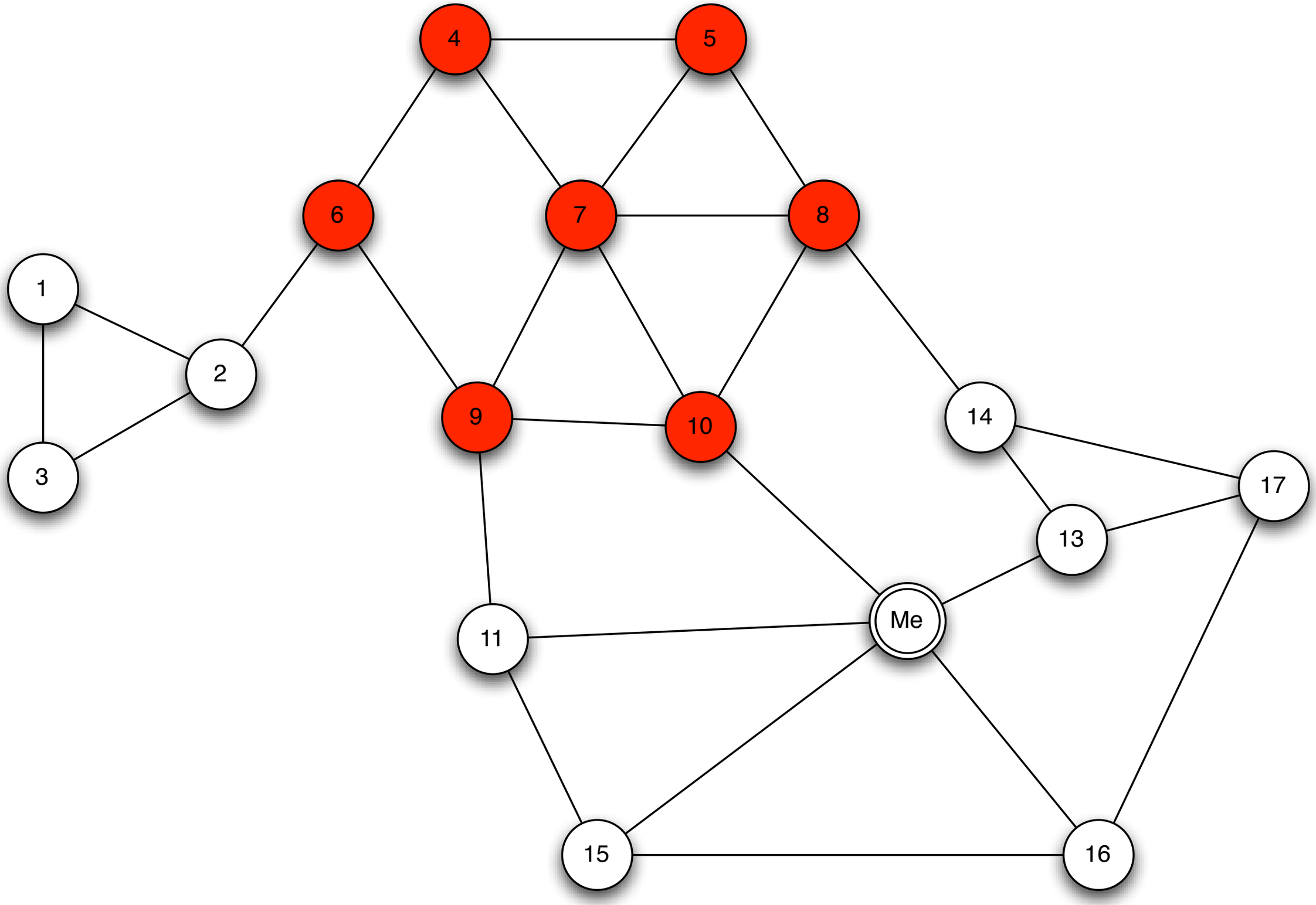
- The topology of the network has consequences for diffusion


4
5
6
7
8
1
2
9
10
14
3
17
13
Me
11
15
16


4
5
6
7
8
1
2
9
10
14
3
17
13
Me
11
15
16


4
5
6
7
8
1
2
9
10
14
17
3
13
Me
11
15
16

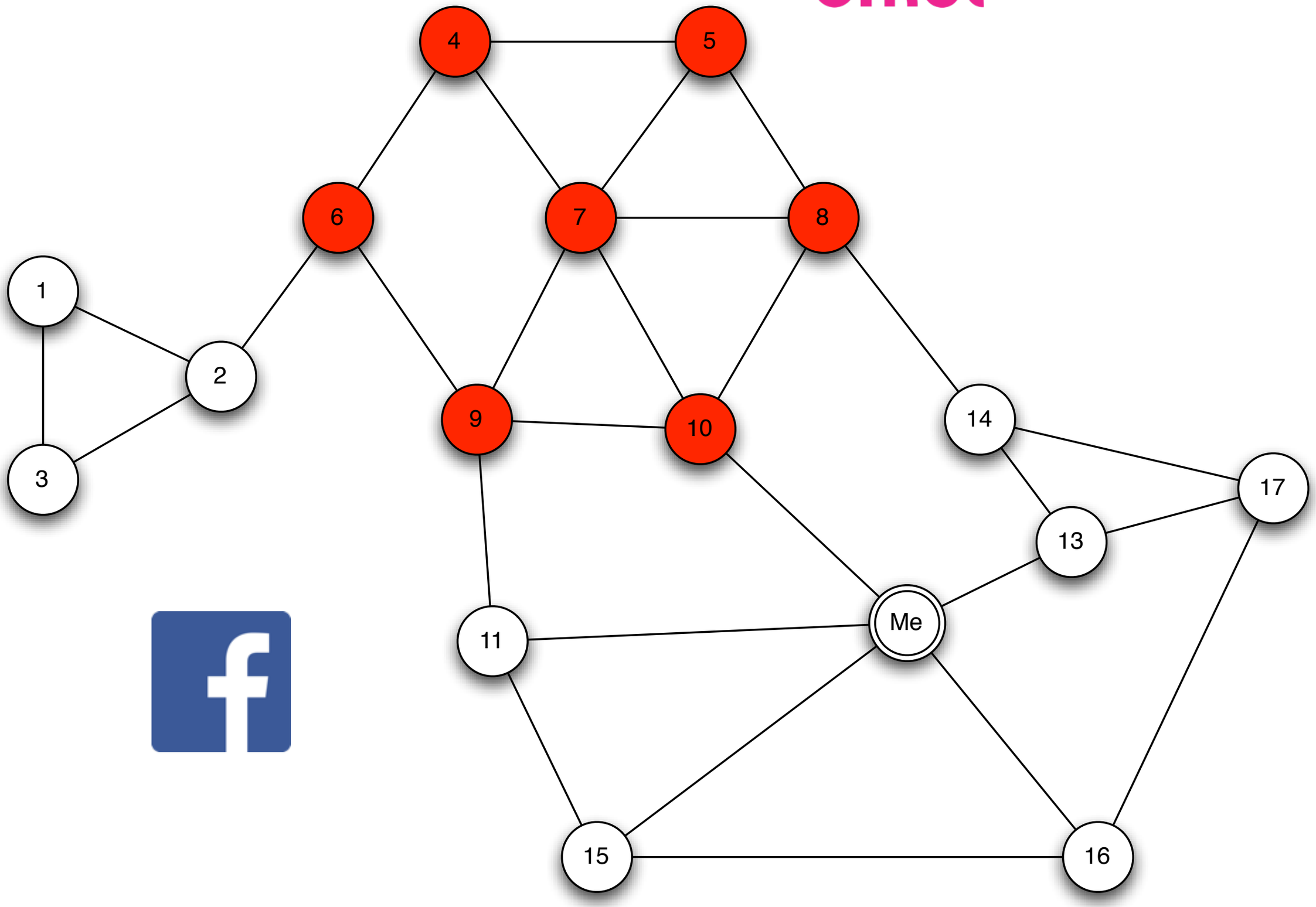

4
5
6
7
8
1
2
9
10
14
17
3
13
Me
11
15
16


4
5
6
7
8
1
2
9
10
14
17
3
13
Me
11
15
16

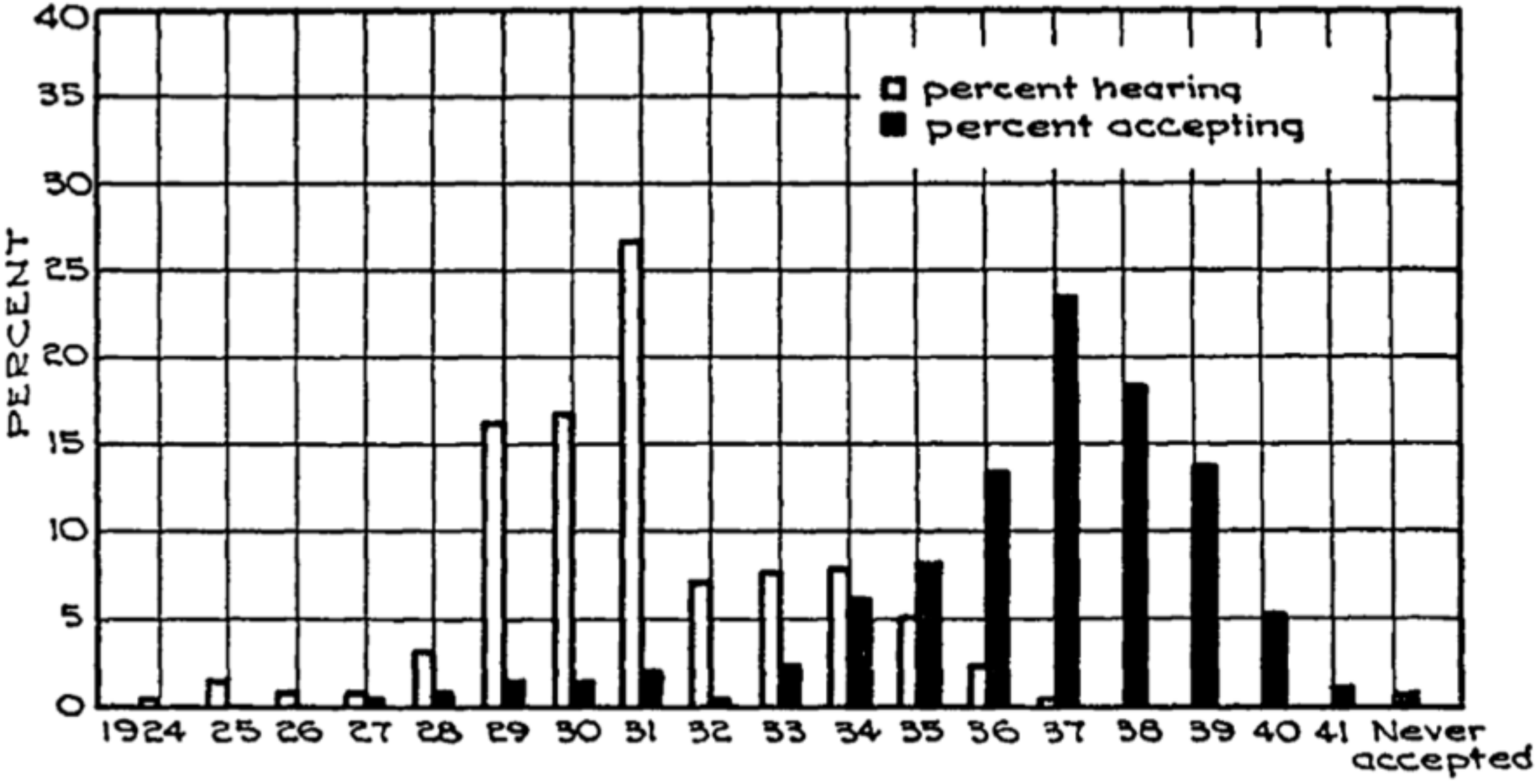
- Tightly connected communities can hinder the spread of innovation

- Viral marketing: how do you choose the nodes where you can maximize adoption in the network?

orkut



Information vs. adoption



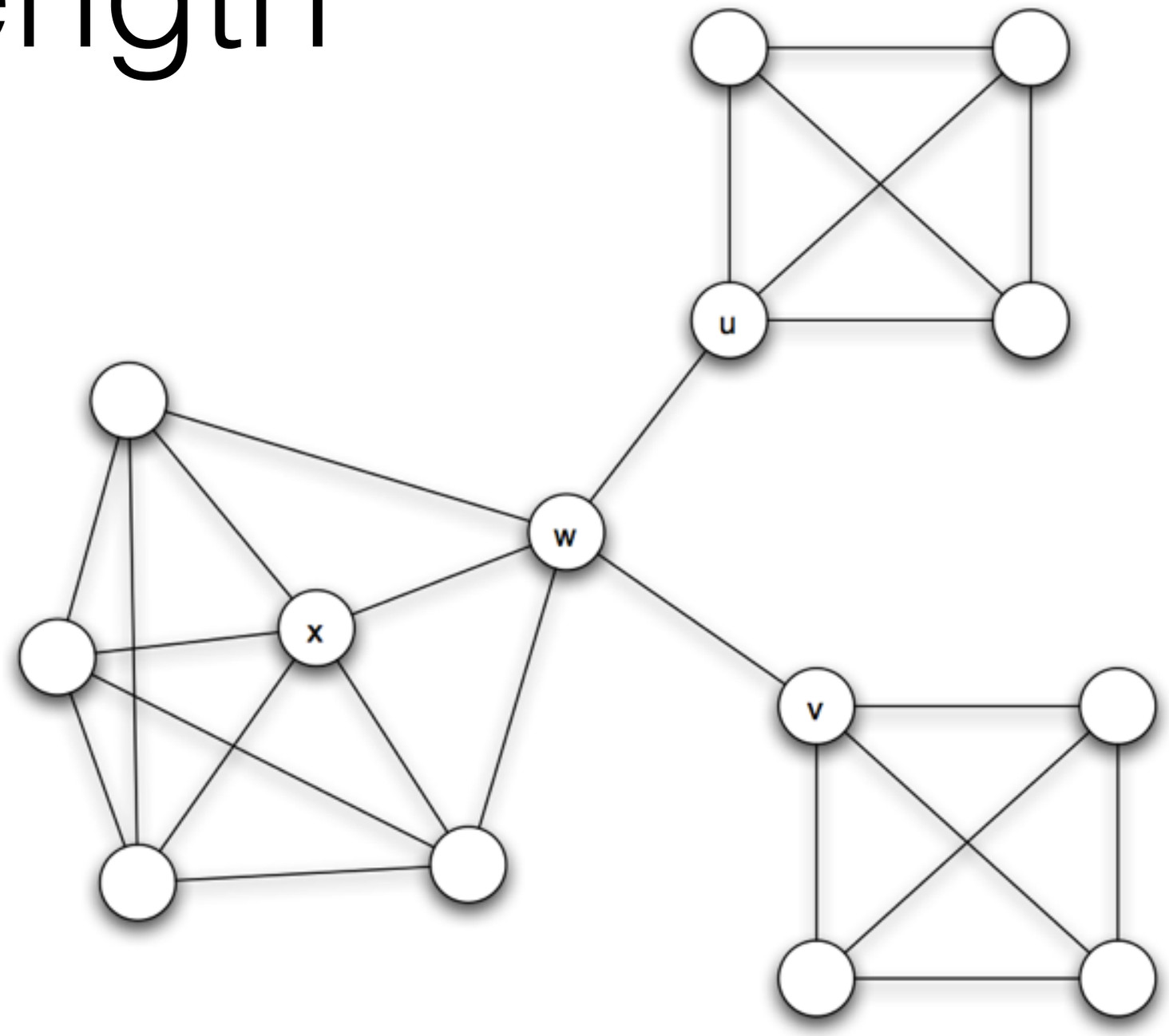
Ryan & Gross (1943), “The Diffusion of Hybrid Seed Corn in Two Iowa Communities,” Rural Sociology

Diffusion of innovations

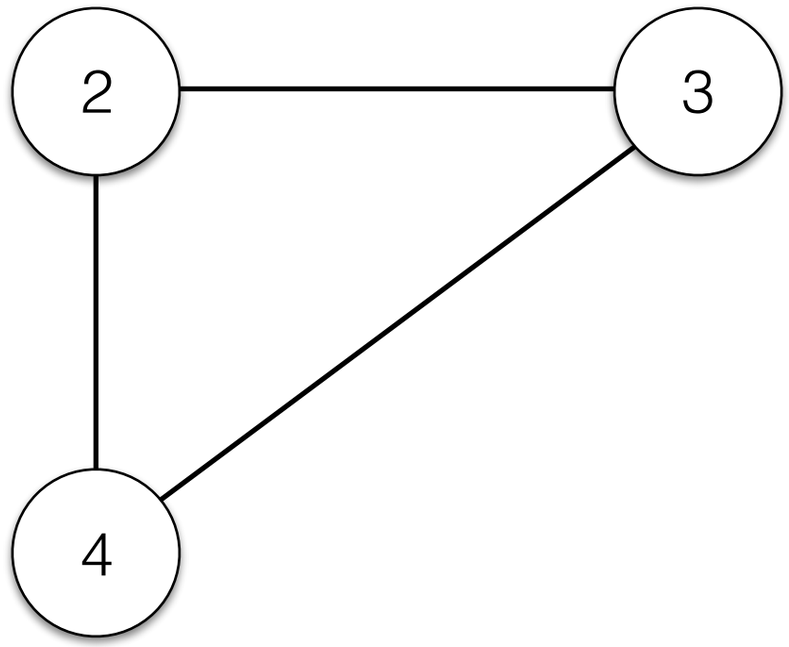
- Spread of a new technology/idea through a social network
- Common principles (Rogers 1995):
  - complexity. How easy can people understand it?
  - observability. How transparent is it when others are using it?
  - trialability. Can it be adopted incrementally?
  - compatibility. How comparable is it with existing practices?

Tie strength

- Hearing about vs. adopting innovation
- Bridges are powerful for conveying awareness, but not uptake



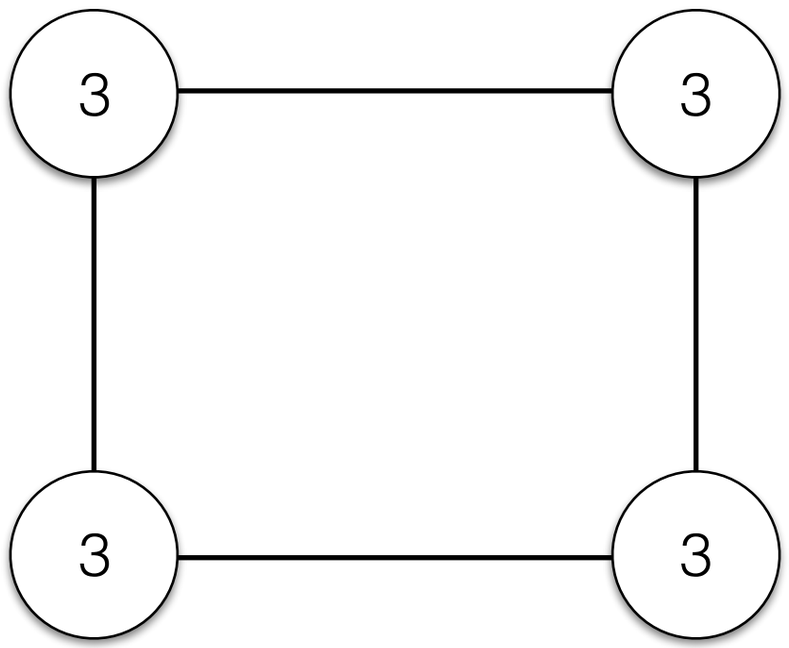
Collective action



“I’ll show up for the protest if at least x other people do too”

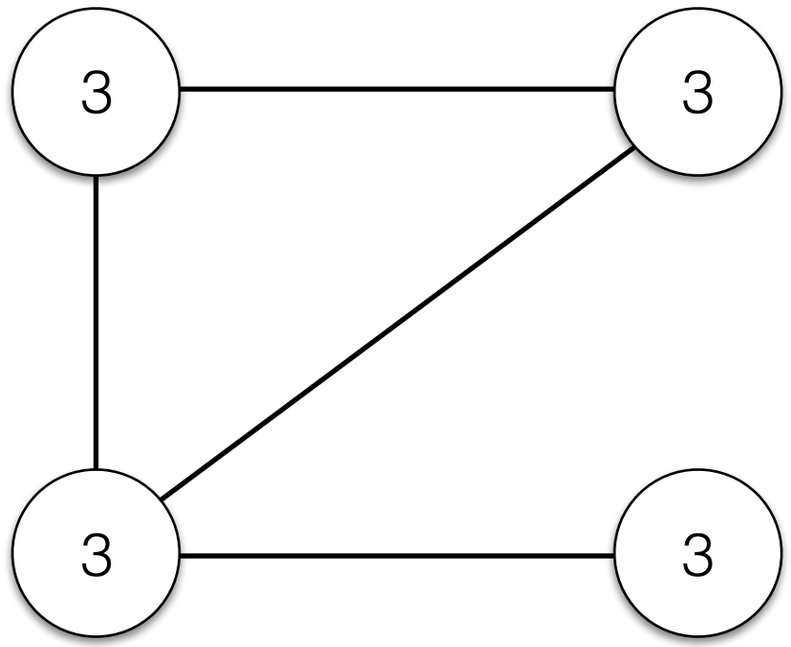
Know the structure of the total social network + the threshold for your friends

Collective action



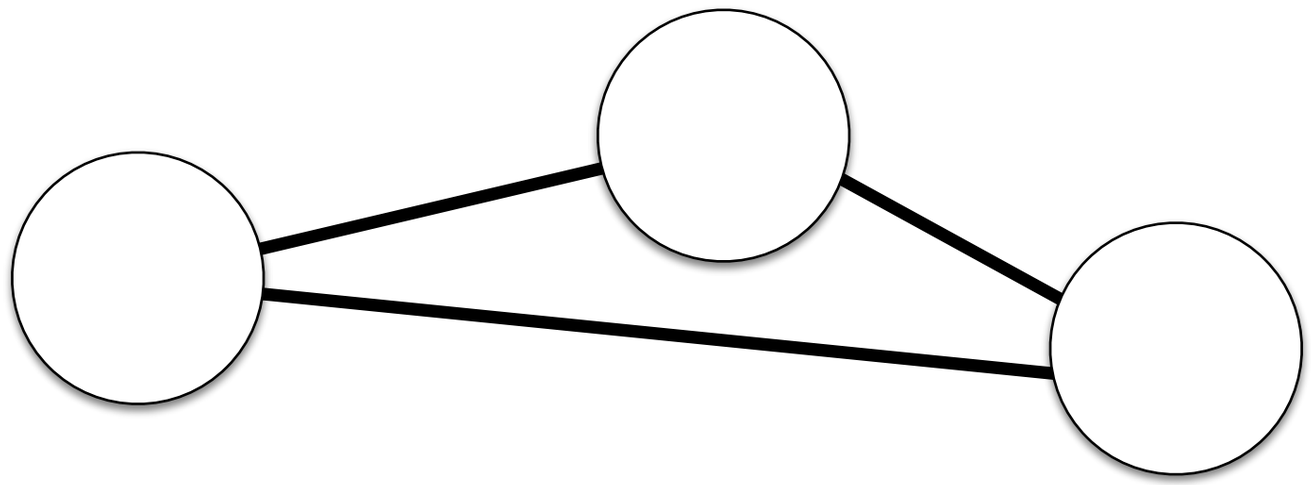
Know the structure of the total social network + the threshold for your friends

Collective action



Know the structure of the total social network + the threshold for your friends

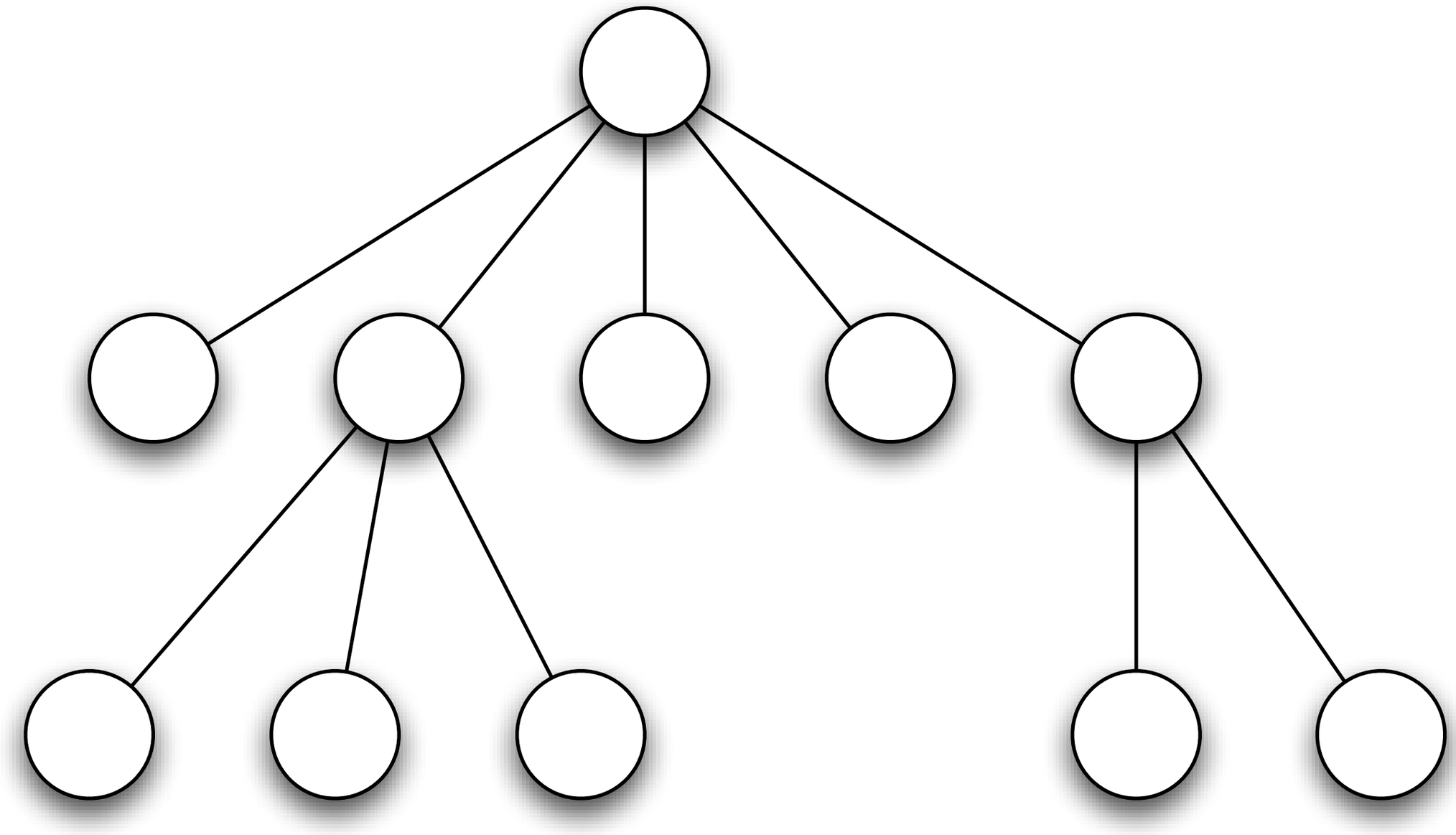
Diffusion as Epidemic



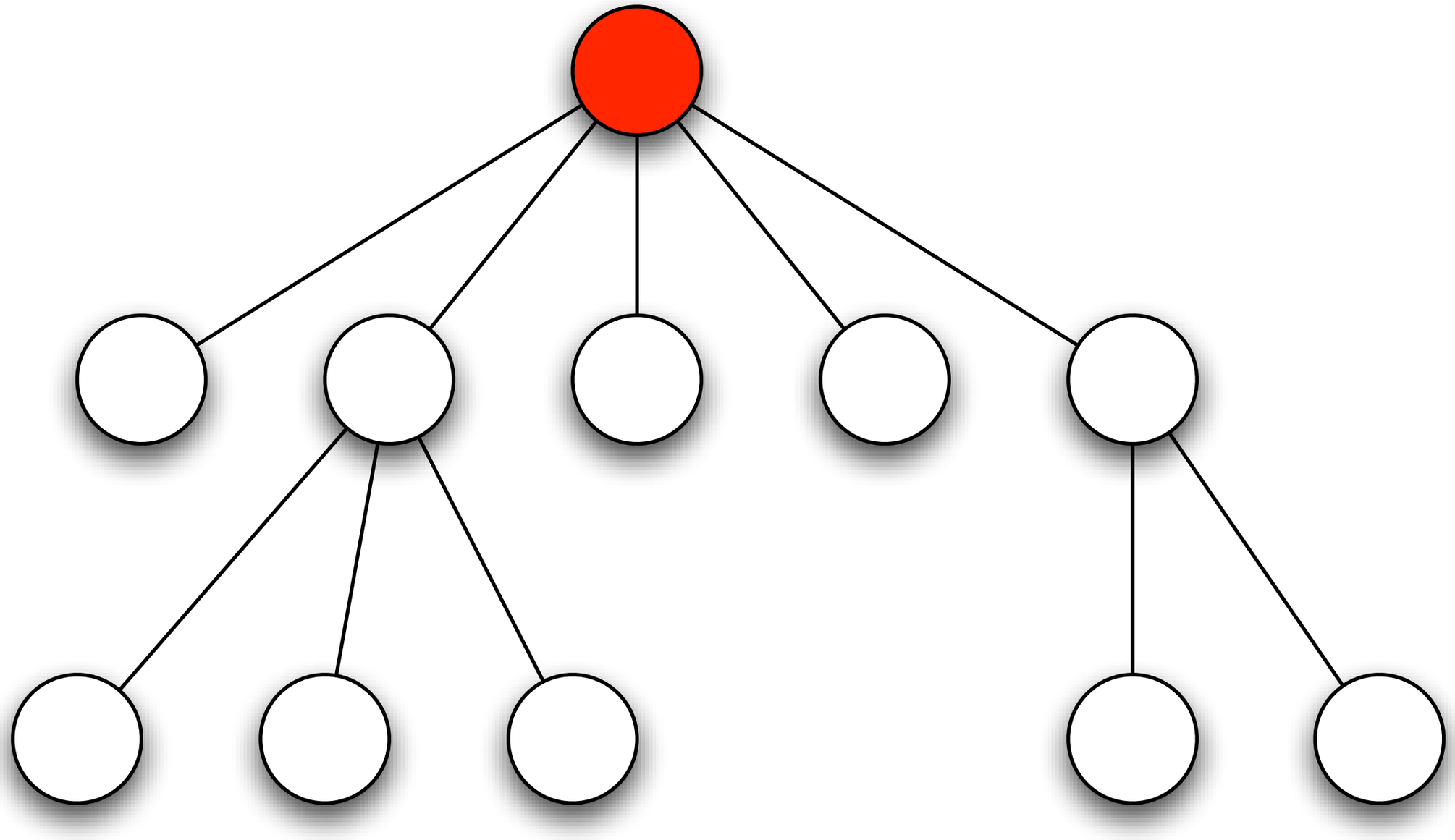
Network	Nodes	Edges	Information
	People		Disease

How does the network change as a function of the disease?

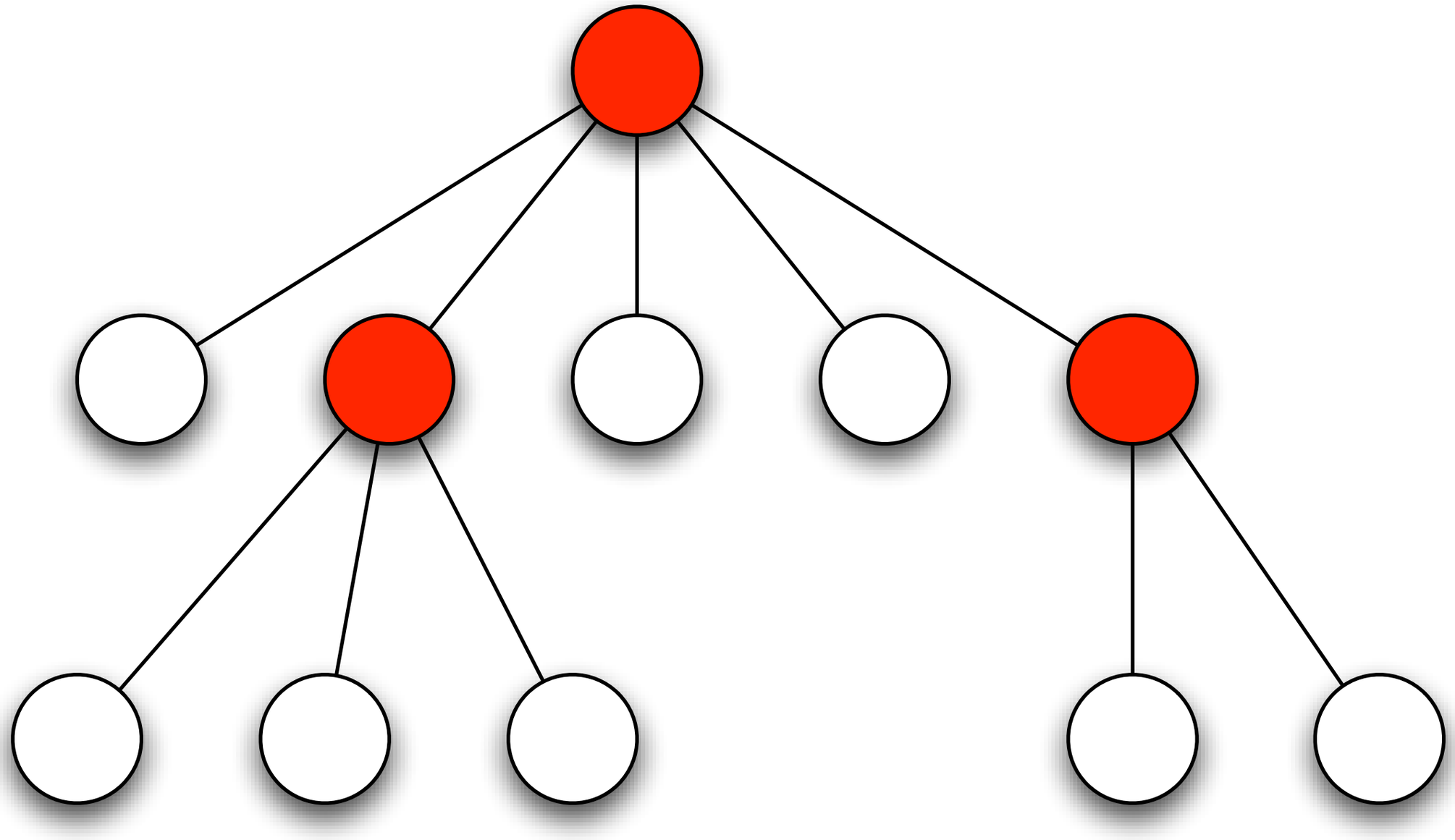
Diffusion as Epidemic



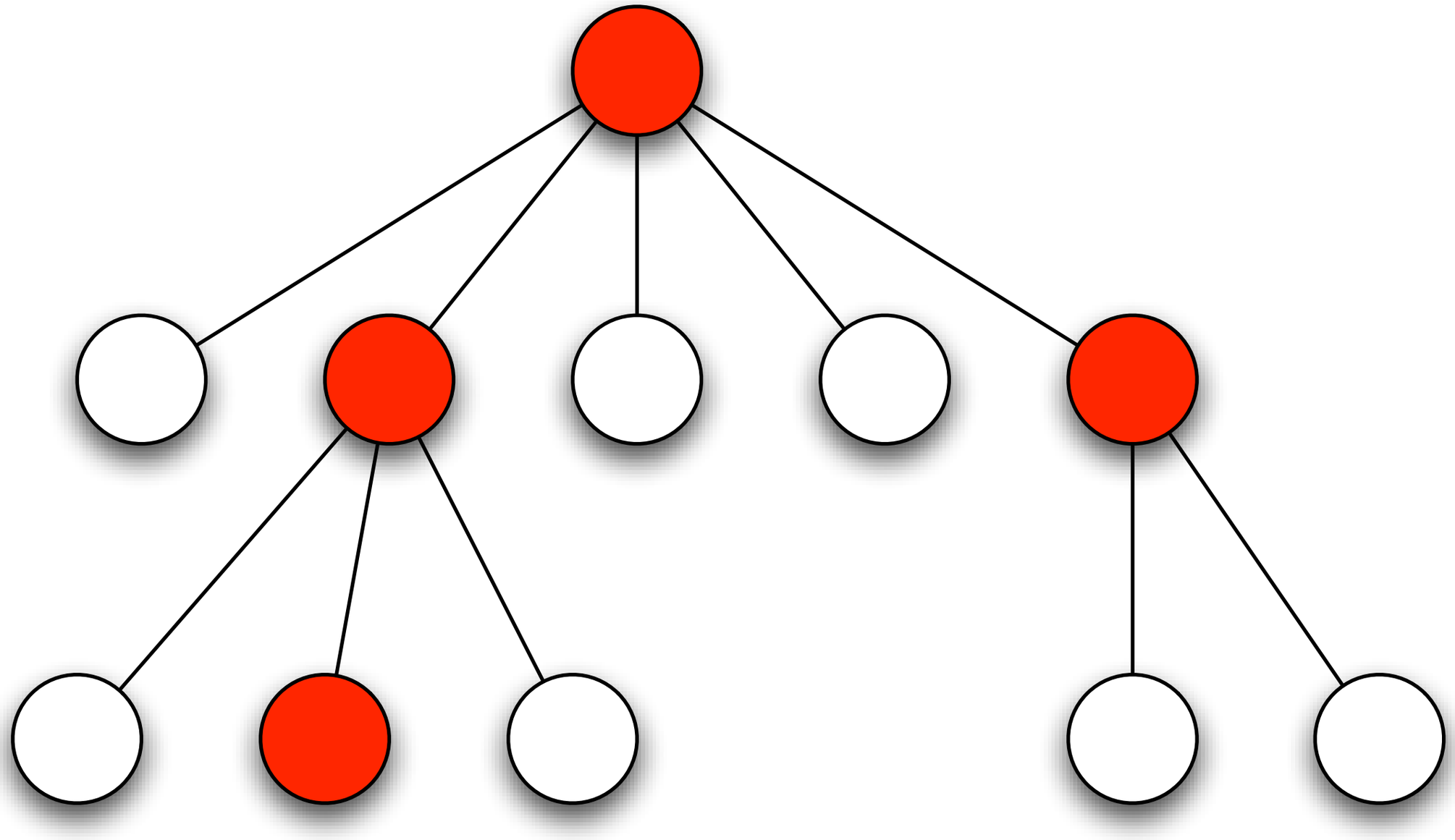
Diffusion as Epidemic



Diffusion as Epidemic



Diffusion as Epidemic



Basic Reproductive Number (R_0)

- Expected number of new infections caused by a randomly selected person in the population

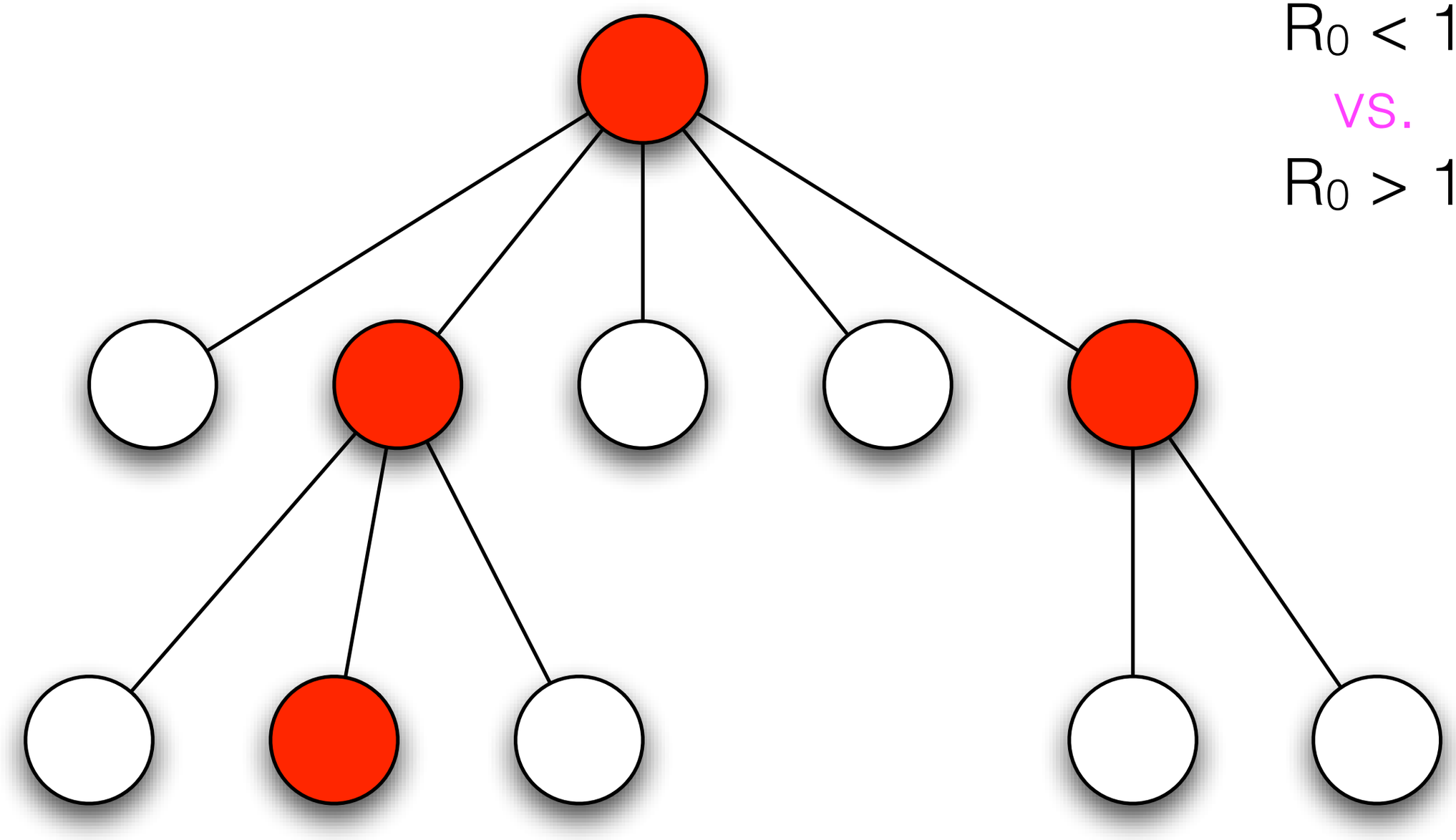
Disease	R_0
1918 Flu	2-3
SARS	2-5
HIV	2-5
Polio	5-7
Smallpox	5-7
Measles	12-18

Diffusion as Epidemic

$R_0 < 1$

vs.

$R_0 > 1$



Basic Reproductive Number (R_0)

- In tree models, R_0 = p x k

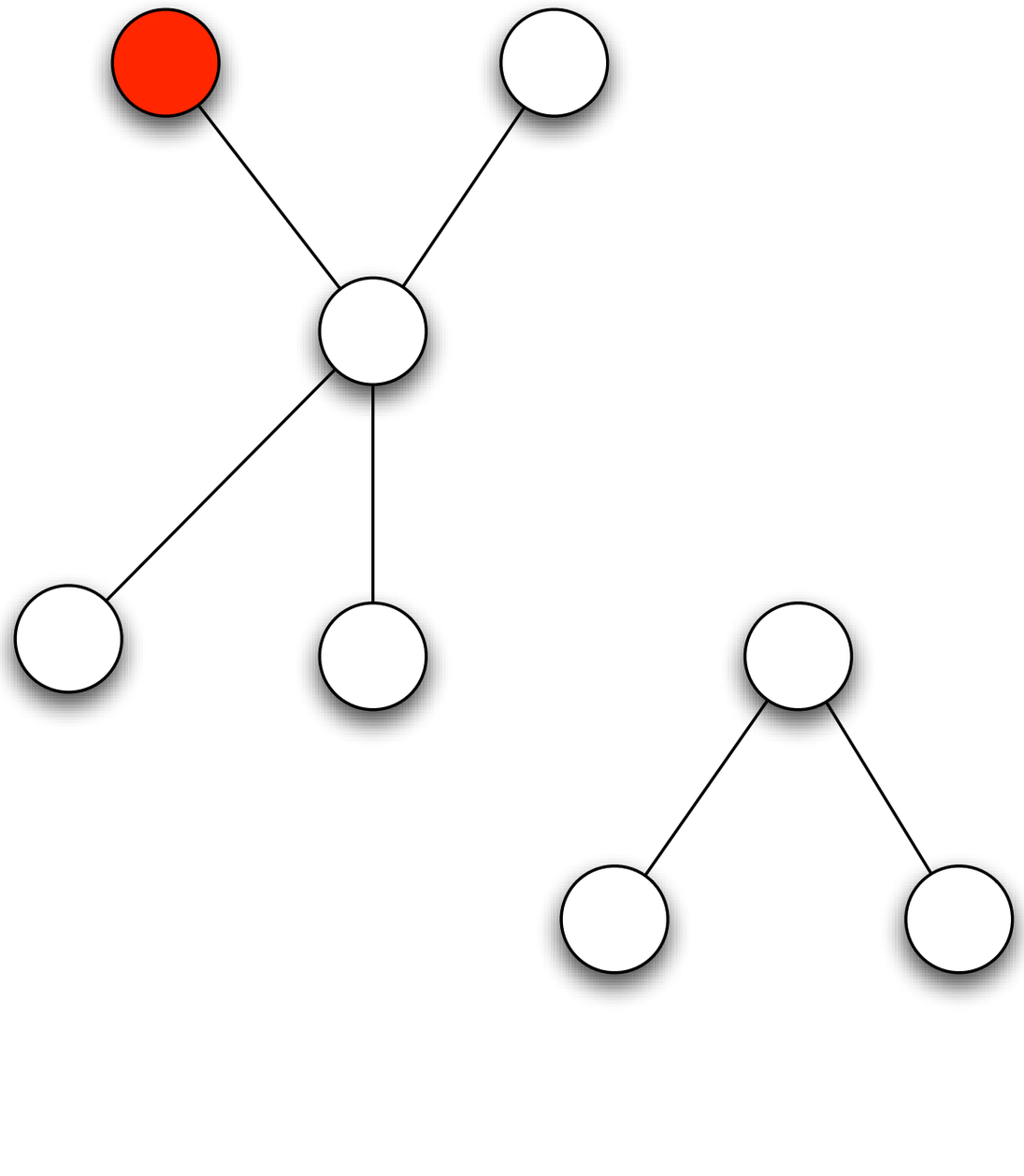
- p = probability of infecting 1 person

- k = number of people in contact with

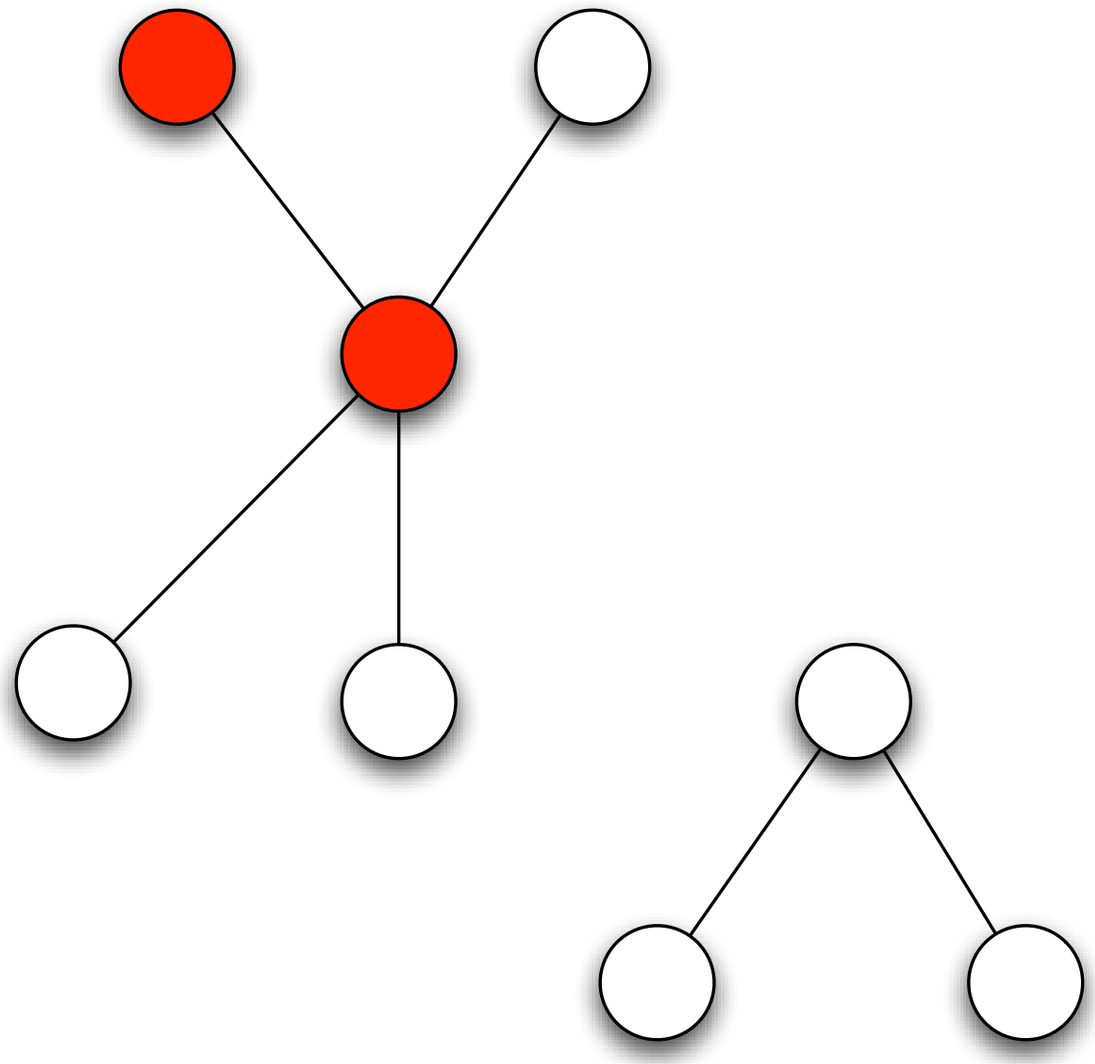
decrease p by preventing spread of disease

decrease k by quarantine

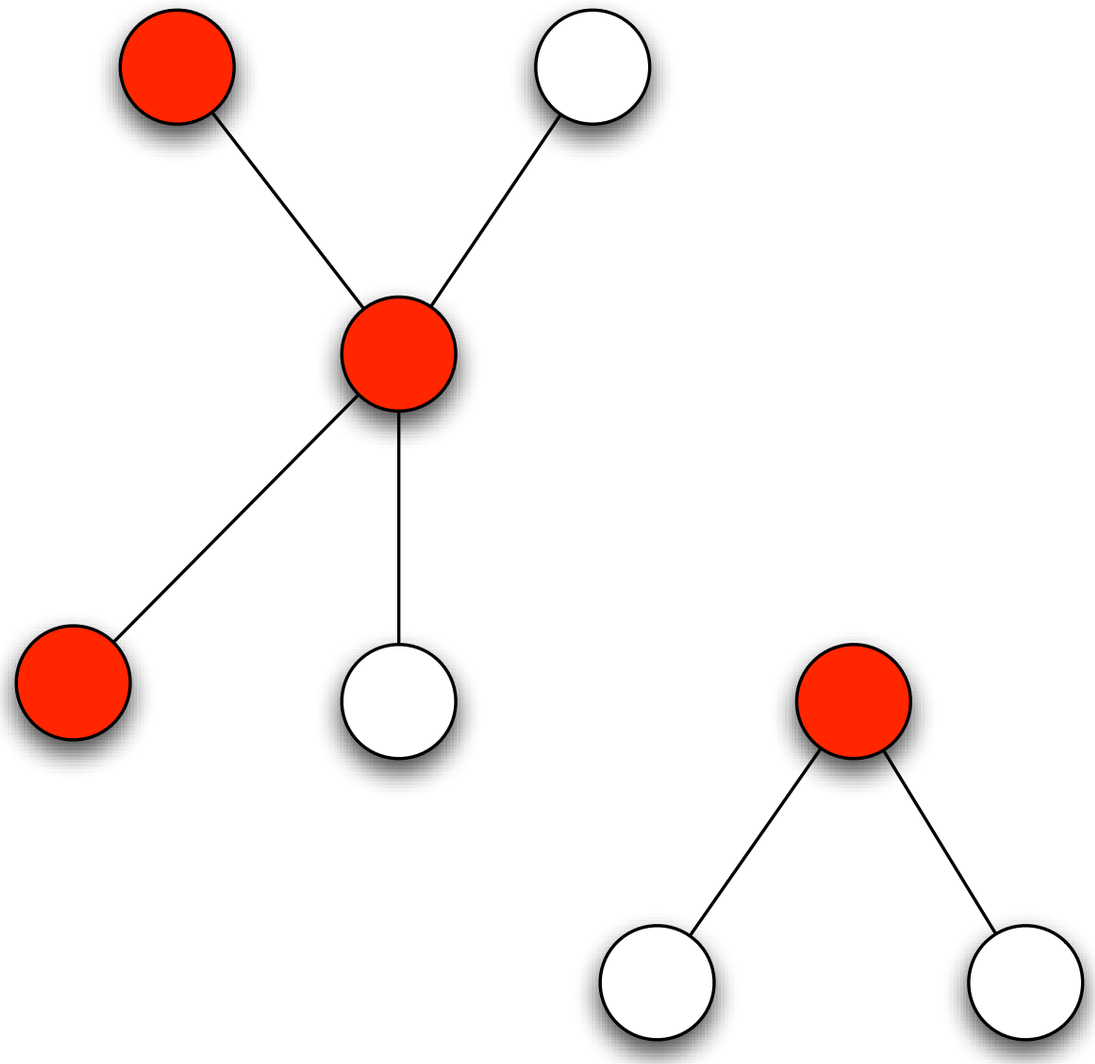
Adar et al., Blogspace



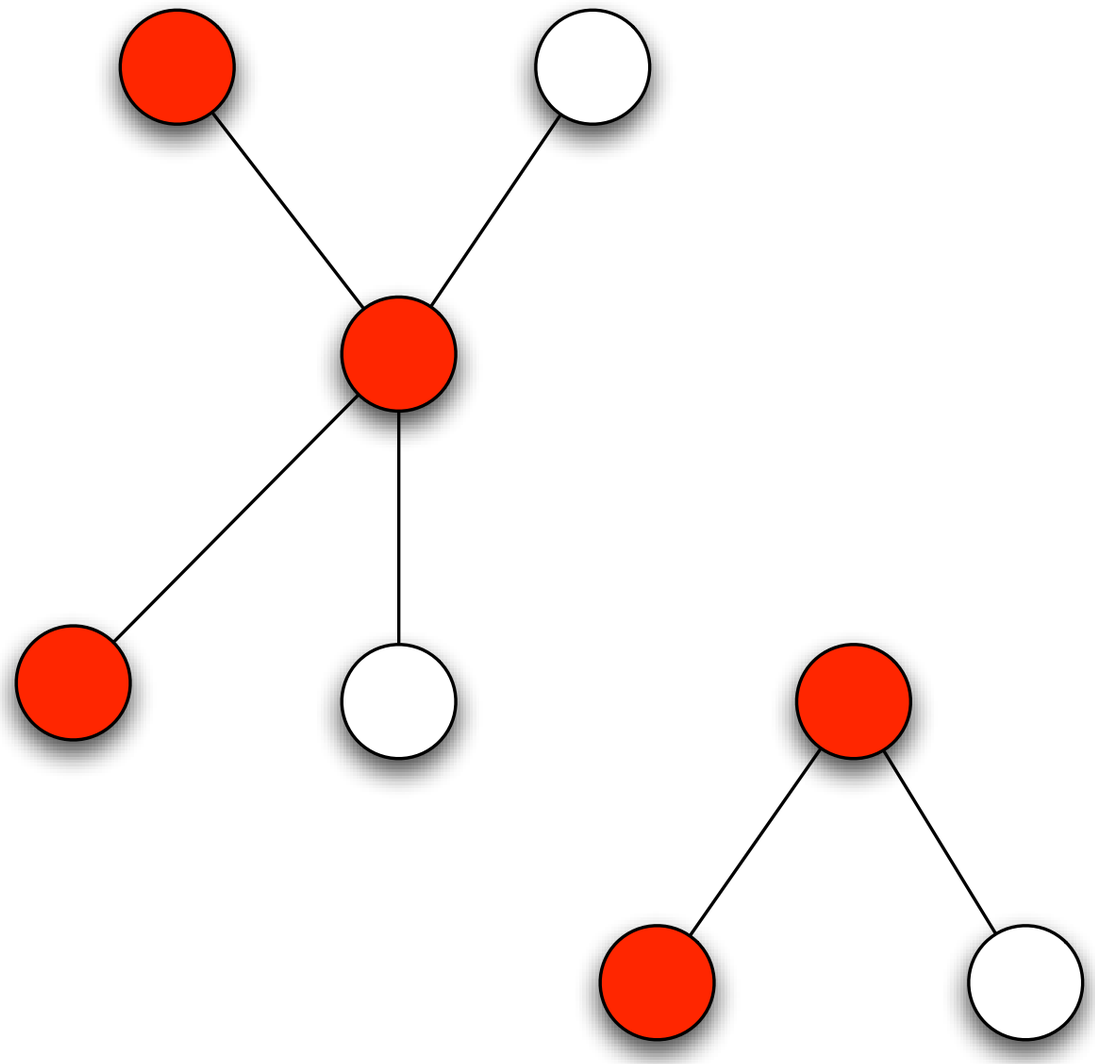
Adar et al., Blogspace



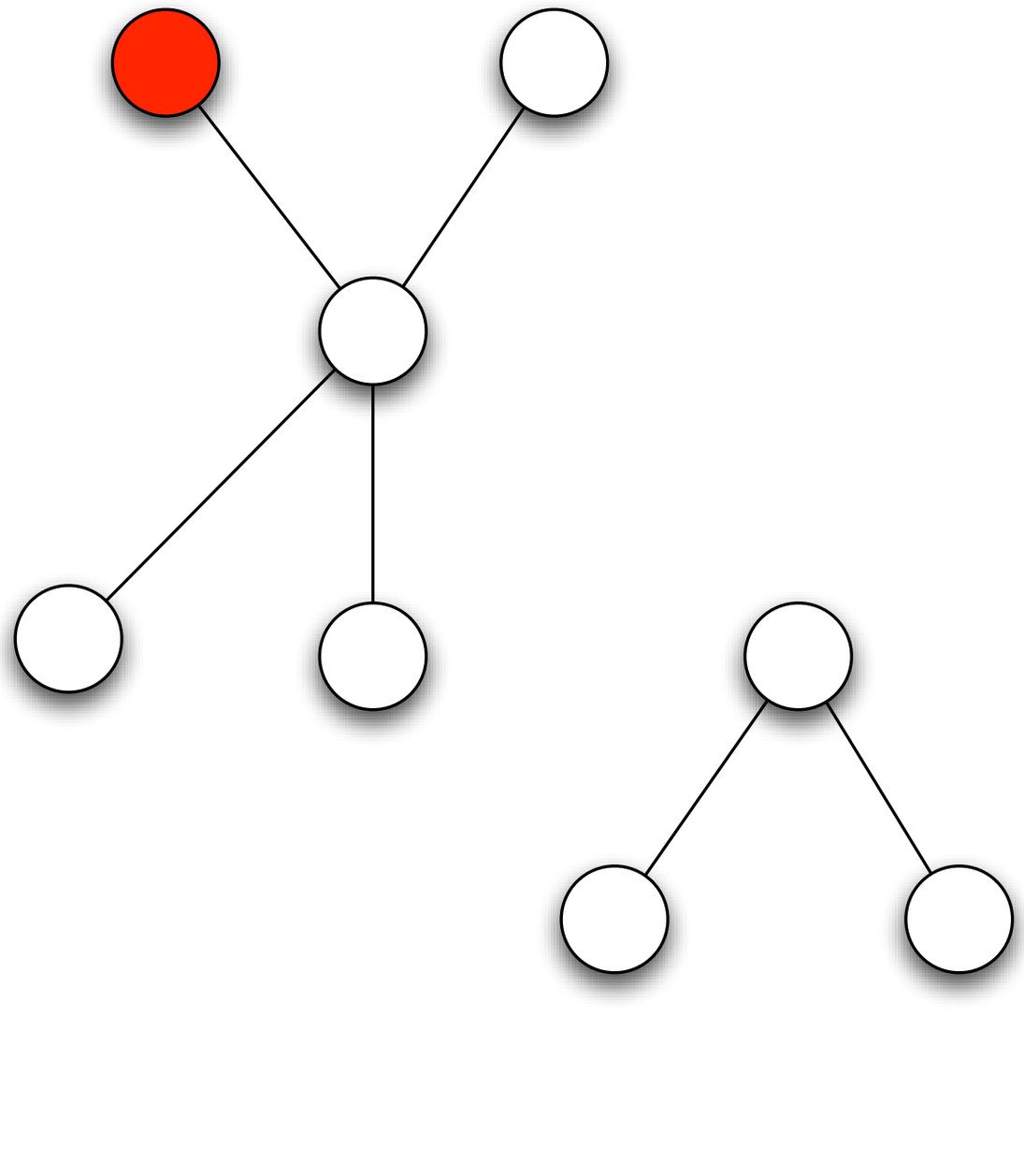
Adar et al., Blogspace



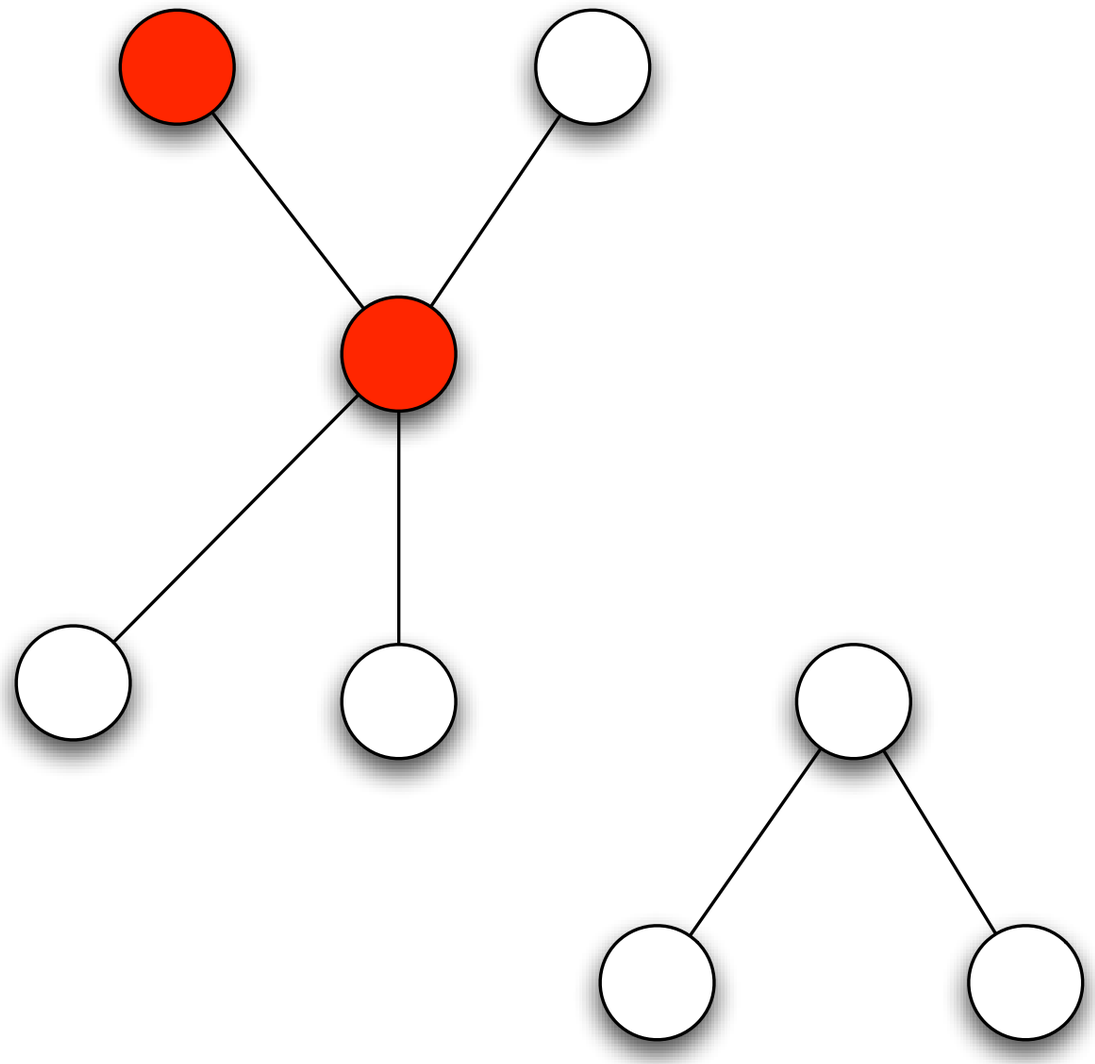
Adar et al., Blogspace



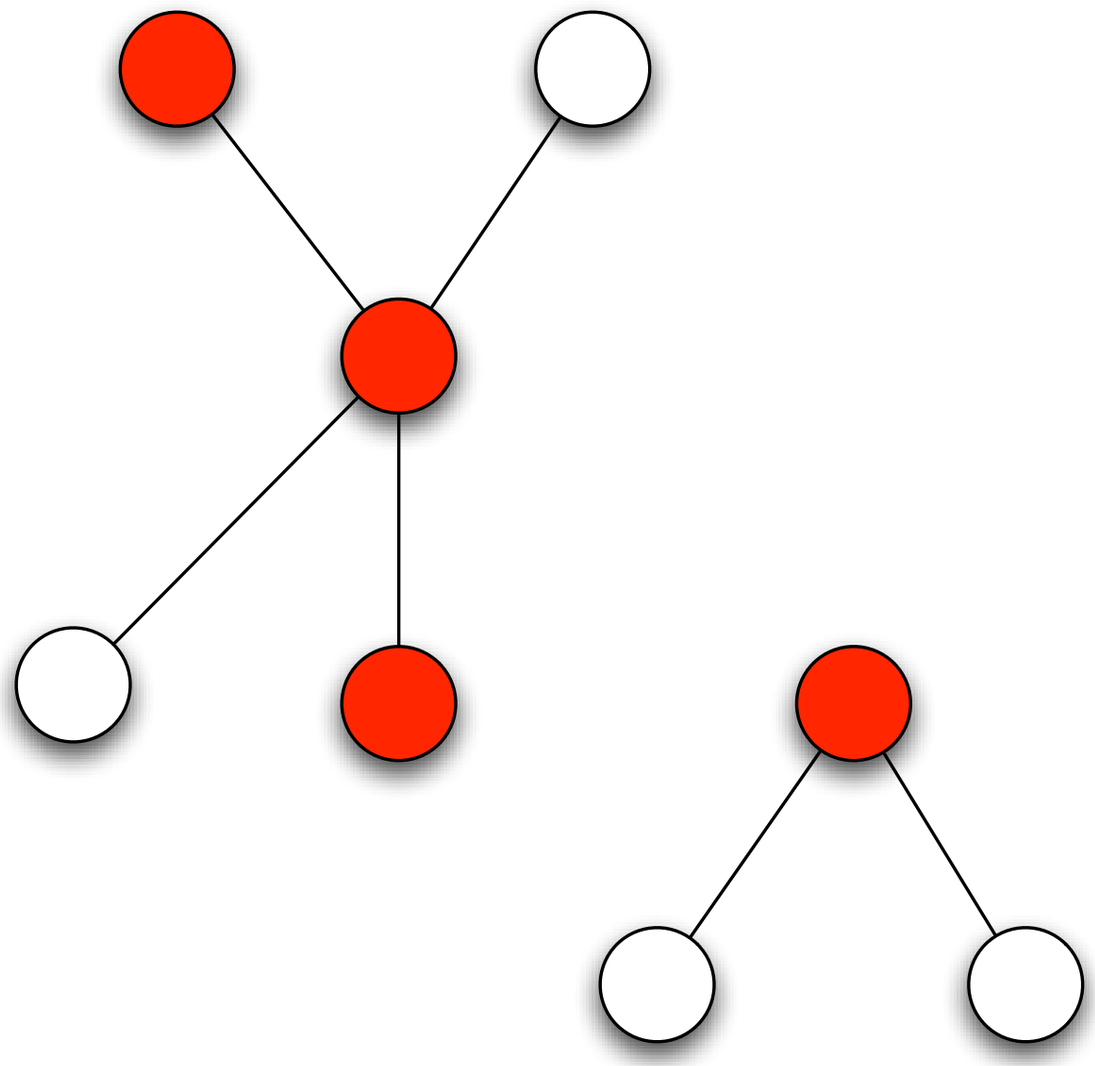
Adar et al., Blogspace



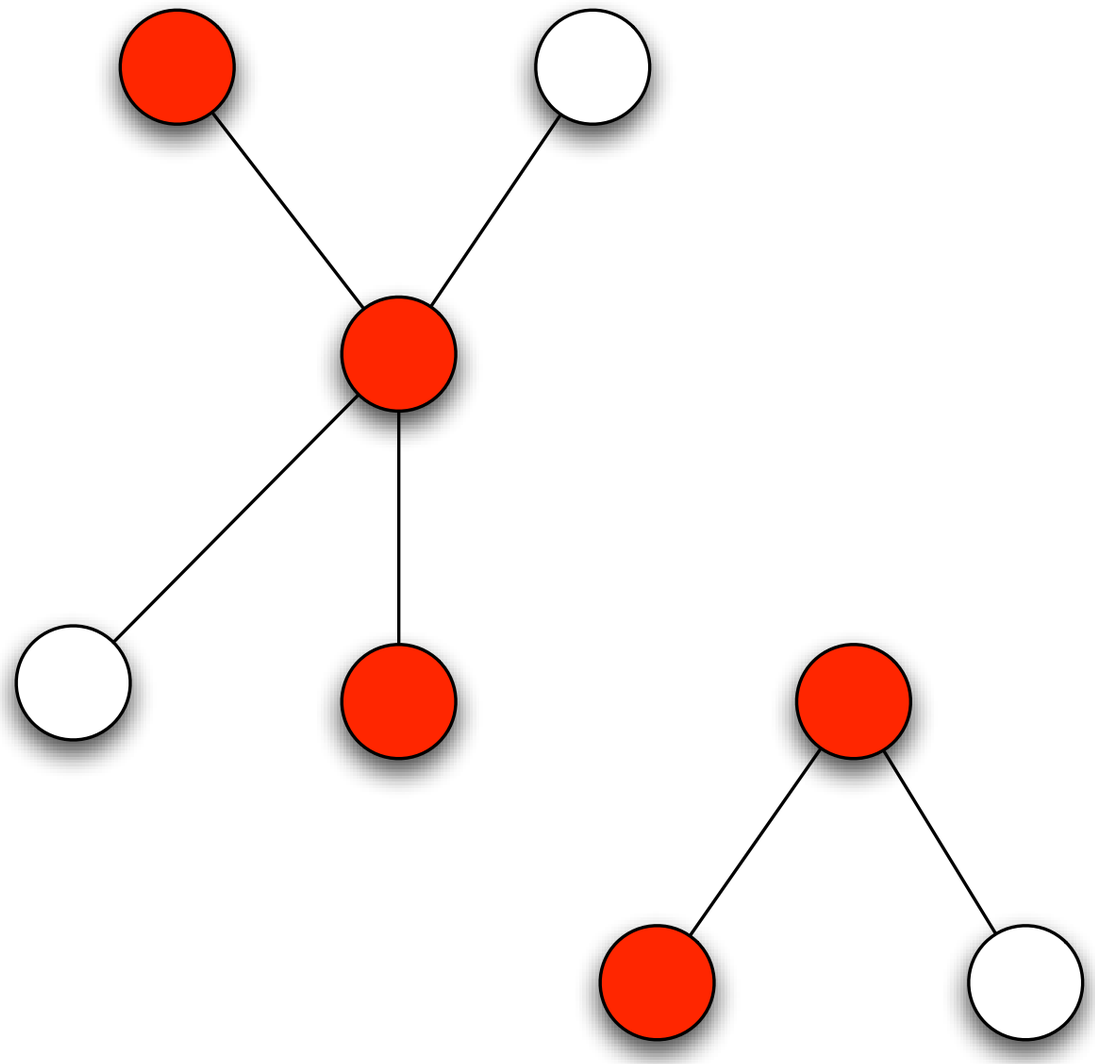
Adar et al., Blogspace



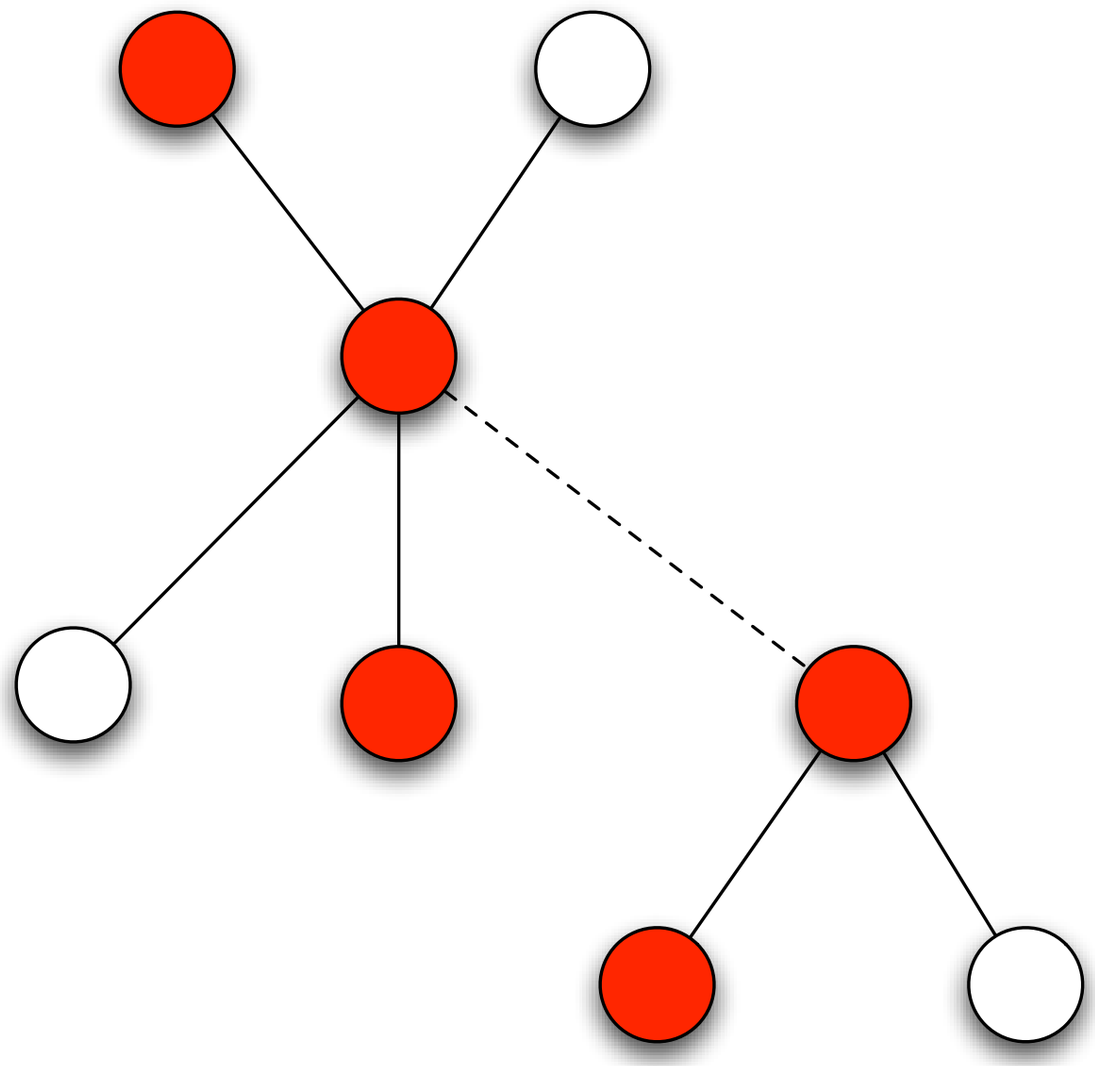
Adar et al., Blogspace



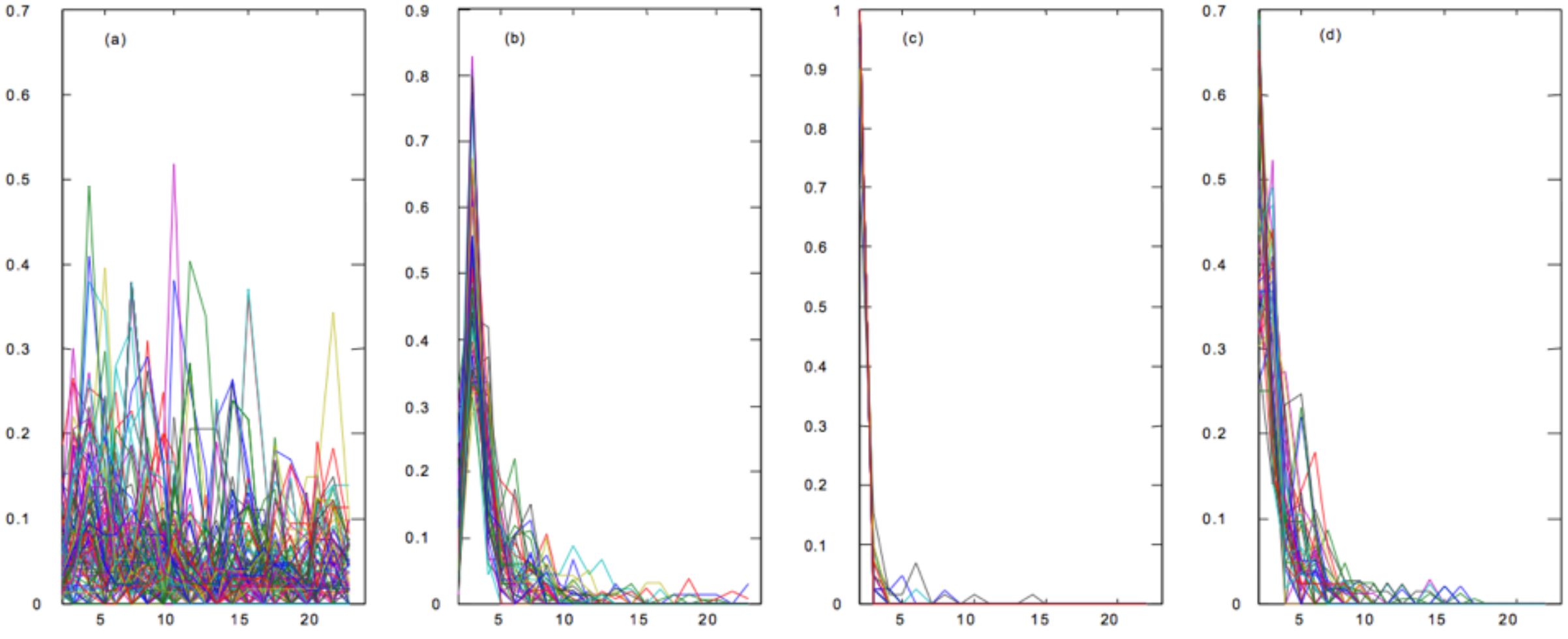
Adar et al., Blogspace



Adar et al., Blogspace



Adar et al., Blogspace



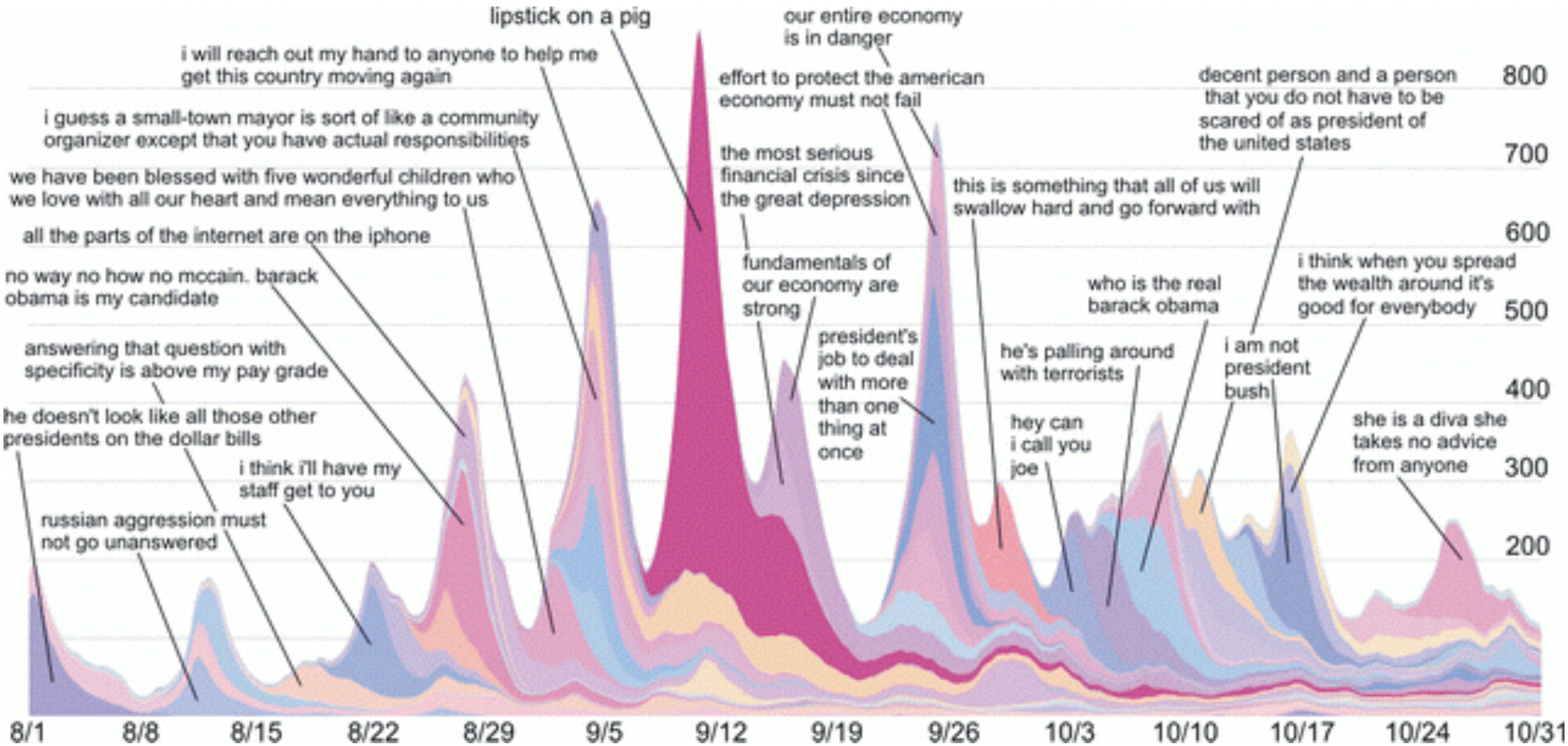
apple friendster

nytimes

slashdot

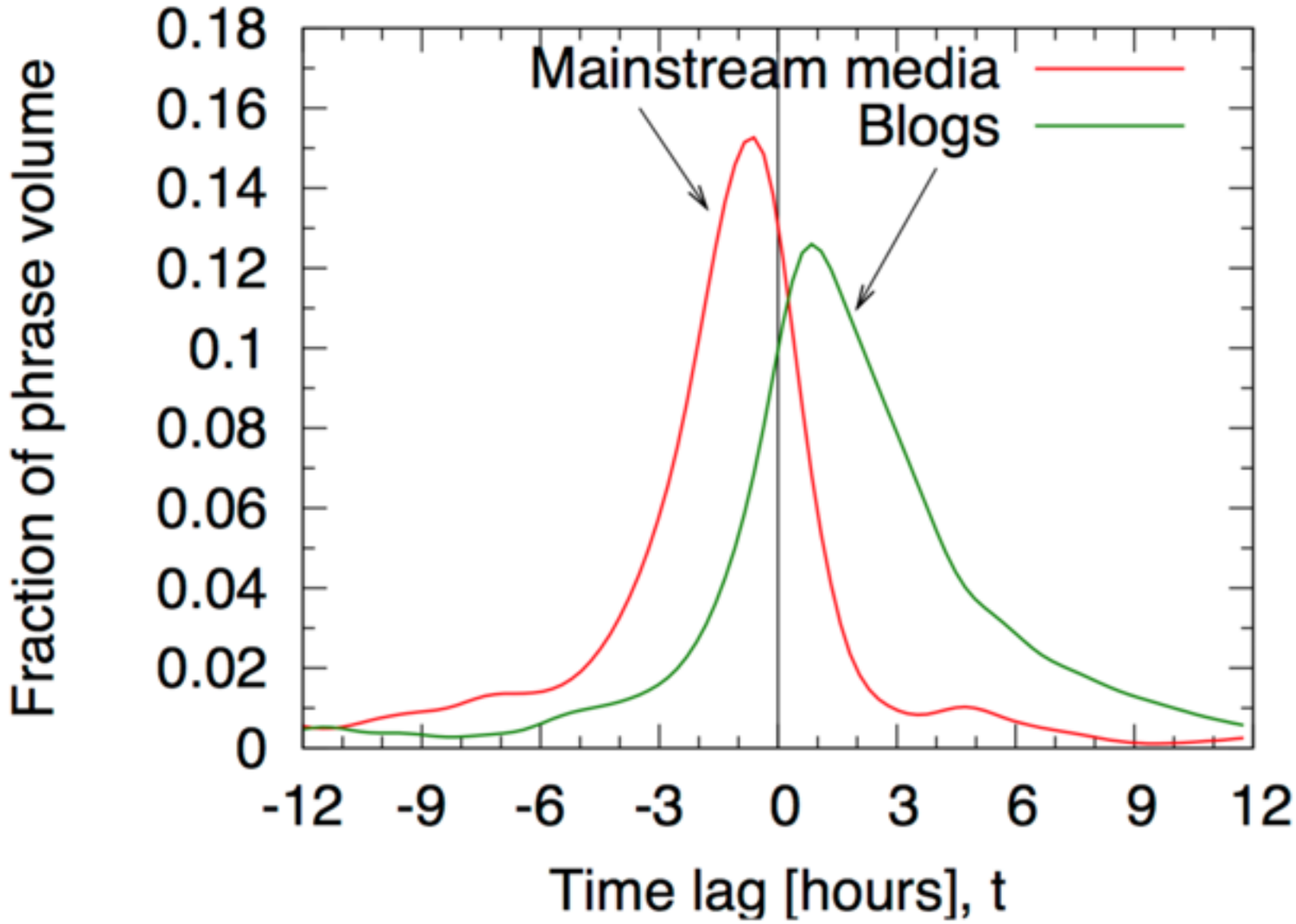
wired/culture

Meme tracking



J. Leskovec et al. (2009), "Meme-tracking and the Dynamics of the News Cycle"

Meme tracking



J. Leskovec et al. (2009), "Meme-tracking and the Dynamics of the News Cycle"

Meme tracking

Rank	Lag [h]	Reported	Site
1	-26.5	42	hotair.com
2	-23	33	talkingpointsmemo.com
4	-19.5	56	politicalticker.blogs.cnn.com
5	-18	73	huffingtonpost.com
6	-17	49	digg.com
7	-16	89	breitbart.com
8	-15	31	thepoliticalcarnival.blogspot.com
9	-15	32	talkleft.com
10	-14.5	34	dailykos.com
16	-14	54	blogs.abcnews.com
30	-11	32	uk.reuters.com
34	-11	72	cnn.com
40	-10.5	78	washingtonpost.com
48	-10	53	online.wsj.com
49	-10	54	ap.org

Table 1: How quickly different media sites report a phrase.

J. Leskovec et al. (2009), "Meme-tracking and the Dynamics of the News Cycle"

In practice

- Influencers on Twitter (RTs)
- Influencers on Twitter (action)
- Viral marketing
- Inferring social network structure

- Nodes
- Edges
- Information

Facebook contagion study

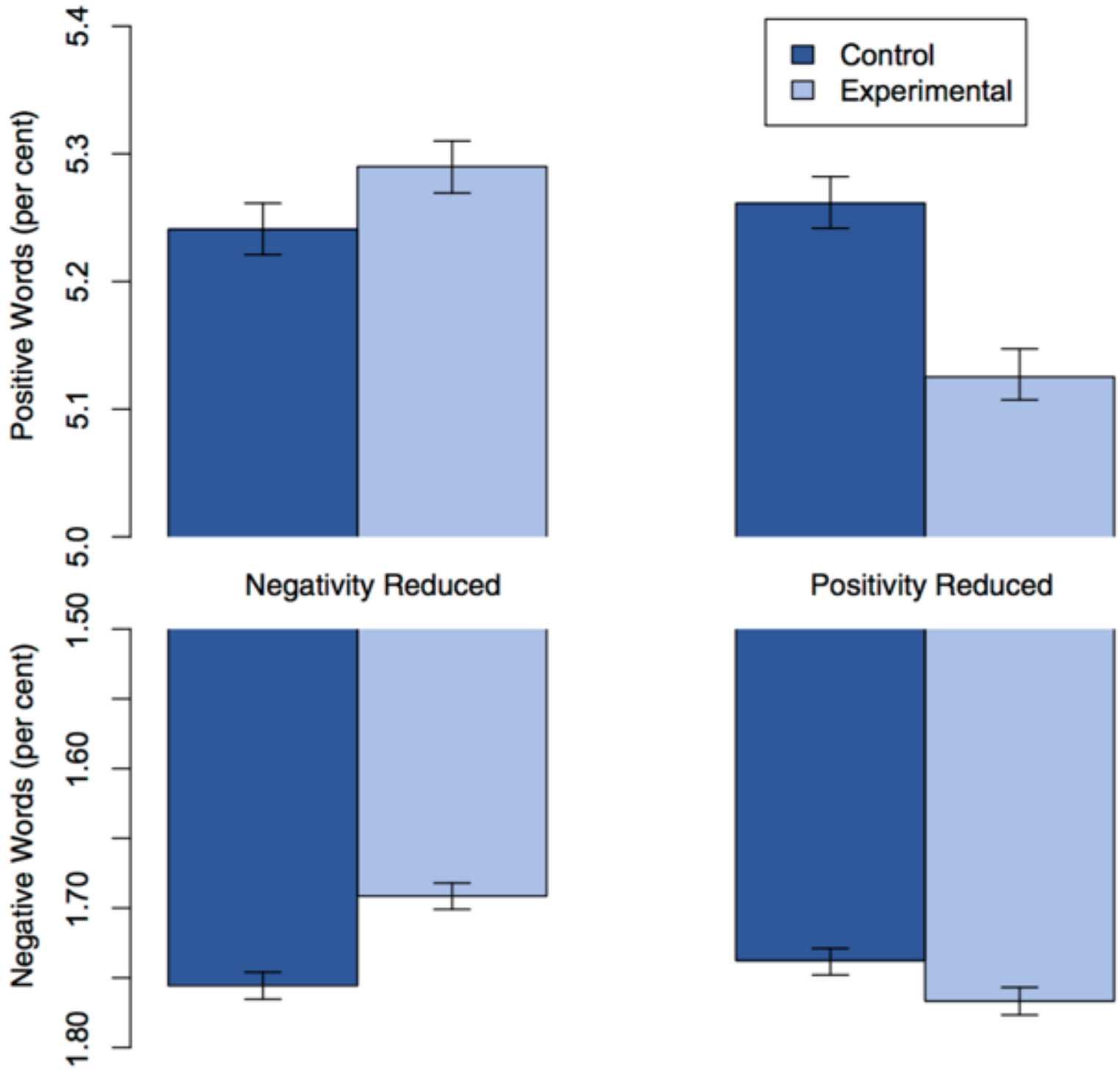


Fig. 1. Mean number of positive (*Upper*) and negative (*Lower*) emotion words (percent) generated people, by condition. Bars represent standard errors.