

First course in Network Science

Sections 4.3-4.5

The content of this presentation is based on the book: A First Course in NETWORK SCIENCE. Filippo Menczer, Santo Fortunato, Clayton A. Davis, ISBN: 9781108471138, Cambridge University Press.

Sarah Alidoost
Analytics SIG, 15/03/2021



Filippo Menczer, Santo Fortunato
and Clayton A. Davis

A First Course in **NETWORK SCIENCE**



► **4 Directions and Weights**

► 4.1 Directed Networks

► 4.2 The Web

► 4.3 PageRank

► 4.4 Weighted Networks

► 4.5 Information and Misinformation

► 4.6 Co-occurrence Networks

► 4.7 Weight Heterogeneity

► 4.8 Summary

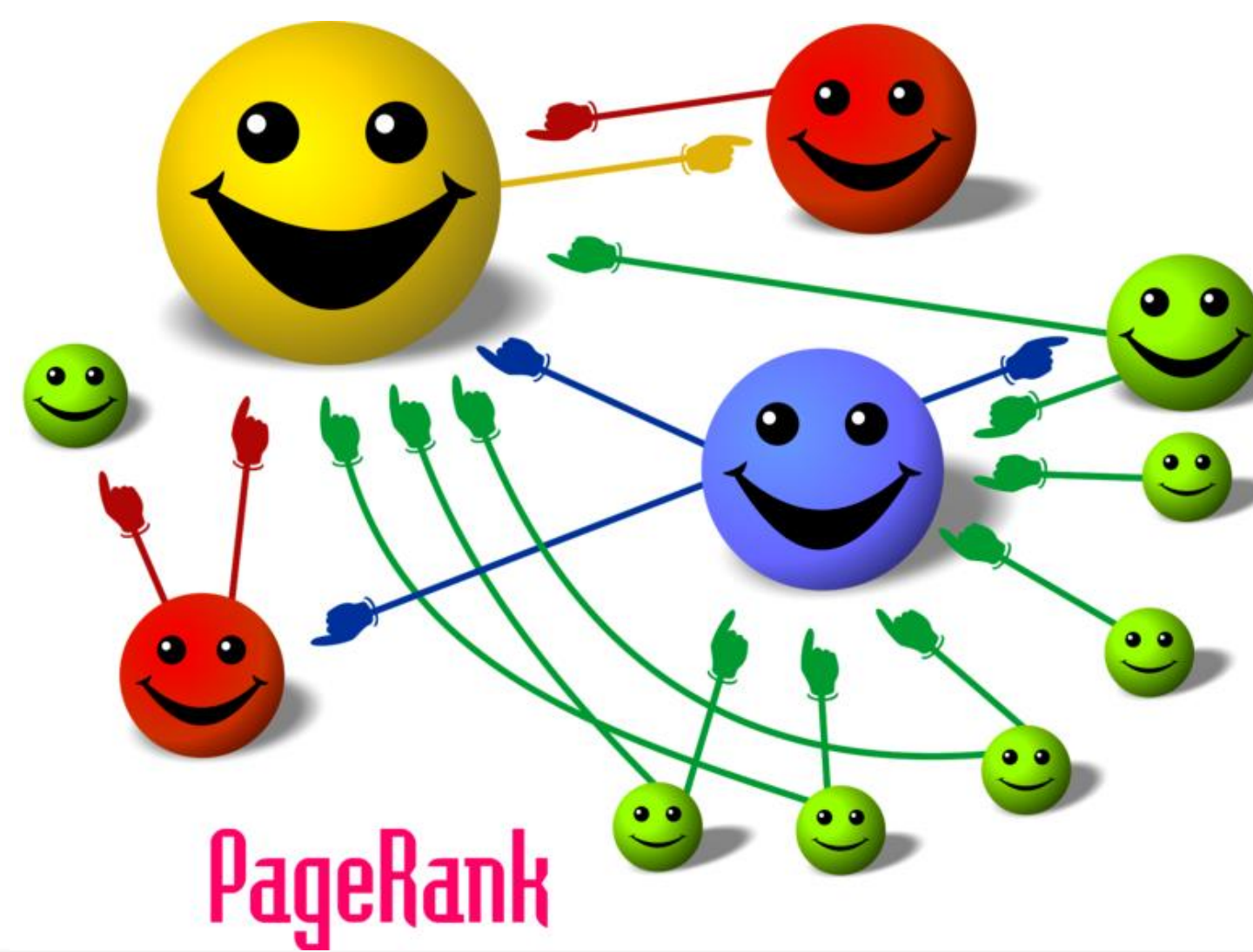
► 4.9 Further Reading

► Exercises



4.3 PageRank / background

- When we submit a query, the search engine can quickly list all the matching pages based on a search index.
- Ranking algorithms are critical components of a search engine.
- Search engines use network centrality measures as ranking criteria.
- Example: **PageRank** introduced by Google in 1998.



<https://en.wikipedia.org/wiki/File:PageRank-hi-res.png> , CC-BY-SA

4.3 PageRank / definition

- PageRank is an algorithm to compute a centrality measure that aims to capture the prestige or importance of each node.
- It is typically used in directed networks.
- Example: a spammer's page created by copying from Wikipedia, and the Wikipedia page.

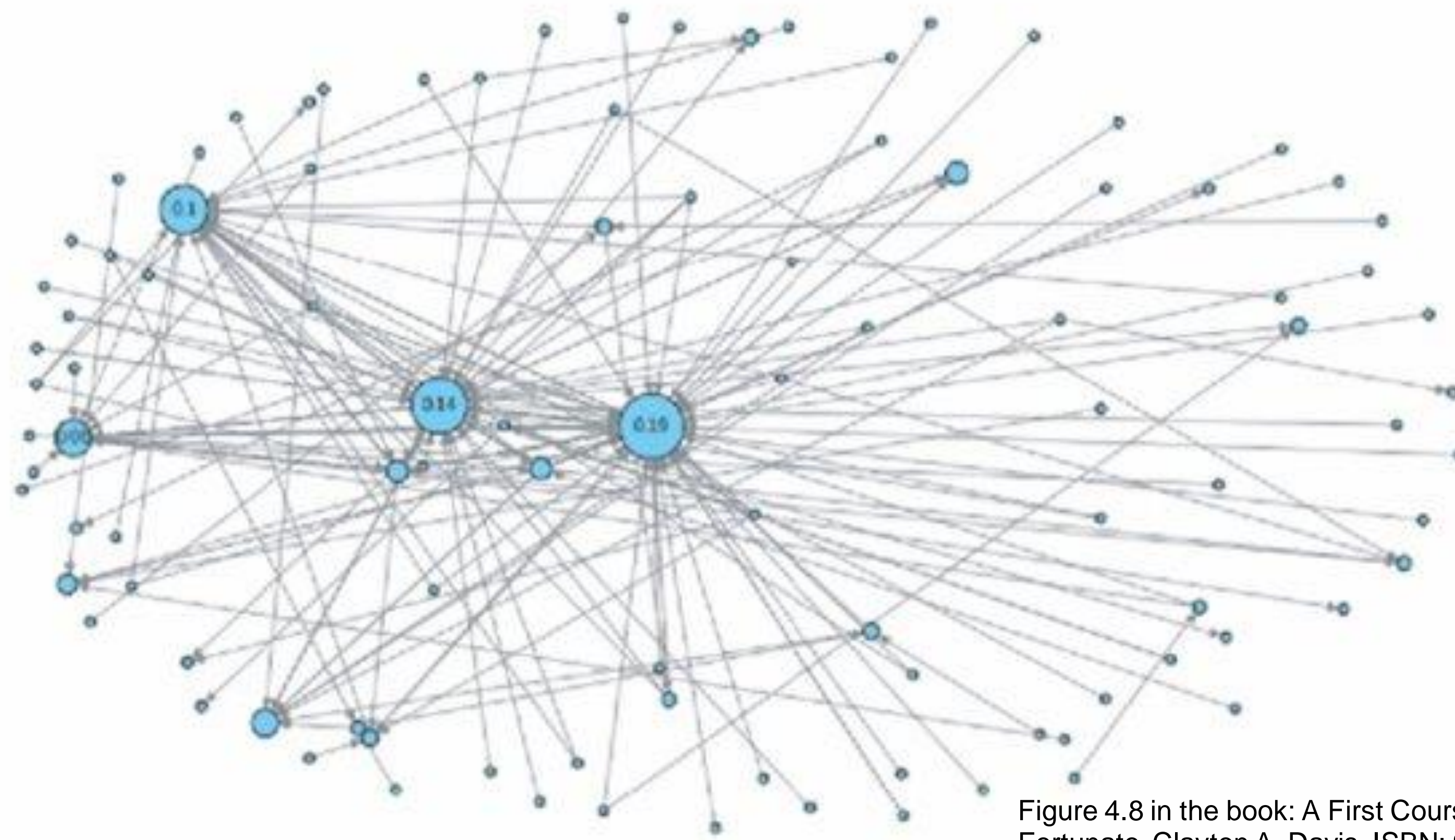


Figure 4.8 in the book: A First Course in NETWORK SCIENCE. Filippo Menczer, Santo Fortunato, Clayton A. Davis, ISBN: 9781108471138, Cambridge University Press.

The size of the nodes is proportional to their PageRank.



4.3 PageRank / searching behaviors

- Random walk or random surf: random link from page to page
- Teleportation or random jump: user starts a new browsing session.

$$R_t(i) = \frac{\alpha}{N} + (1 - \alpha) \sum_{j \in \text{pred}(i)} \frac{R_{t-1}(j)}{k_{\text{out}}(j)}.$$

E.q. 4.1 in the book: A First Course in NETWORK SCIENCE. Filippo Menczer, Santo Fortunato, Clayton A. Davis, ISBN: 9781108471138, Cambridge University Press.

4.3 PageRank / how to compute

It is computed with an interactive approach called the **power** method.

PageRank of
node i at time t

$$R_t(i) = \frac{\alpha}{N} + (1 - \alpha) \sum_{j \in \text{pred}(i)} \frac{R_{t-1}(j)}{k_{\text{out}}(j)}.$$

teleportation to node i

How one can
traverse one
of the links
to i during random
walk

4.3 PageRank / how to compute

$$R_t(i) = \frac{\alpha}{N} + (1 - \alpha) \sum_{j \in \text{pred}(i)} \frac{R_{t-1}(j)}{k_{\text{out}}(j)}.$$

4.3 PageRank / how to compute

- Interactive demonstration of PageRank using NetLogo (Appendix B.1)
- The *pagerank()* function from NetworkX package:

```
PR = nx.pagerank(D) # D is a DiGraph
```



Figure B.1 in the book: A First Course in NETWORK SCIENCE. Filippo Menczer, Santo Fortunato, Clayton A. Davis, ISBN: 9781108471138, Cambridge University Press.

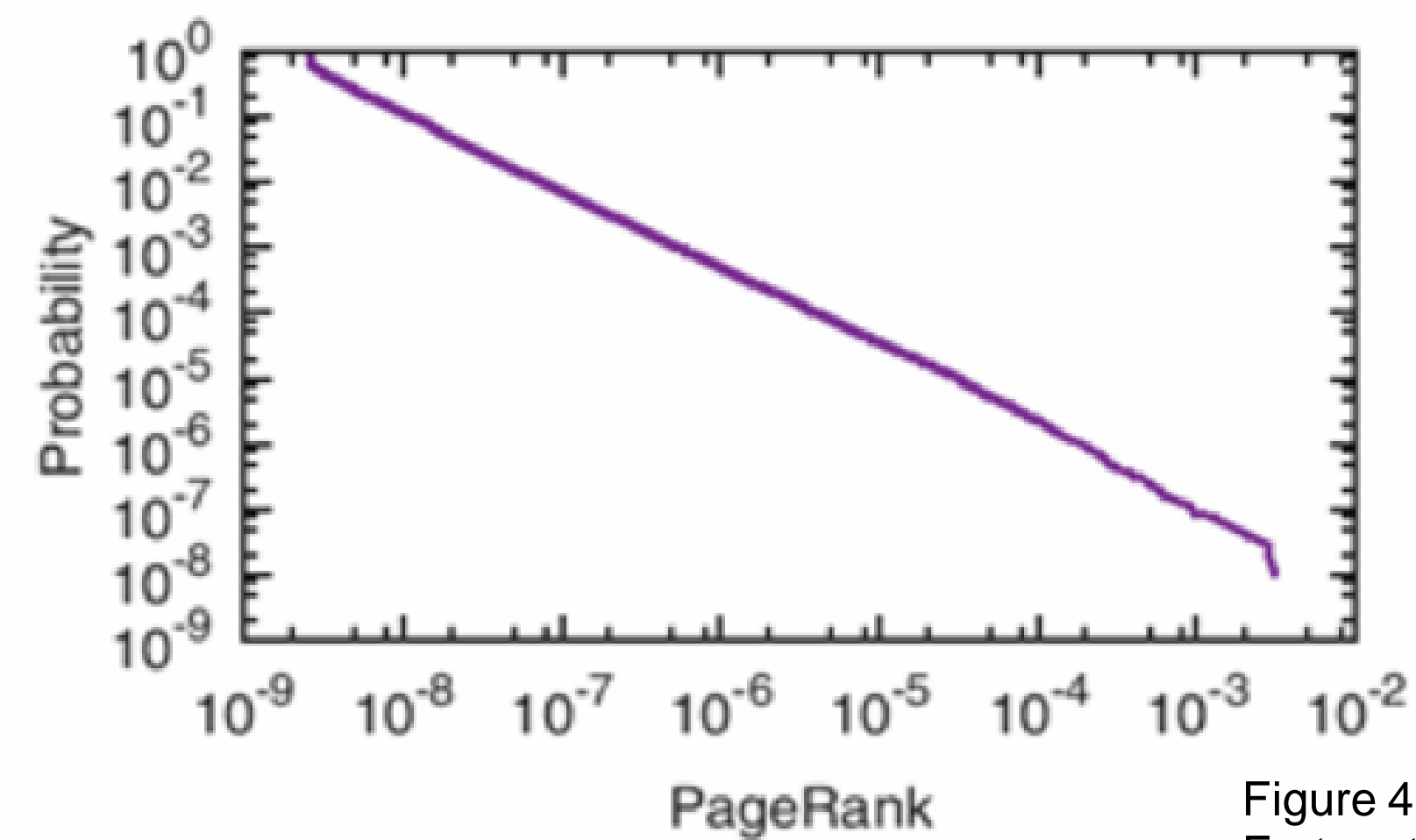
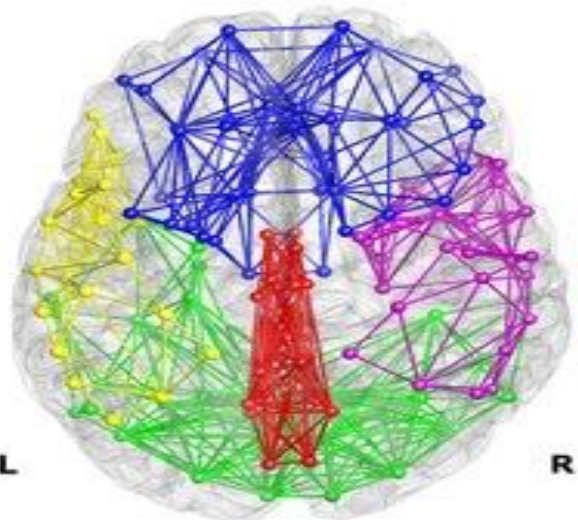
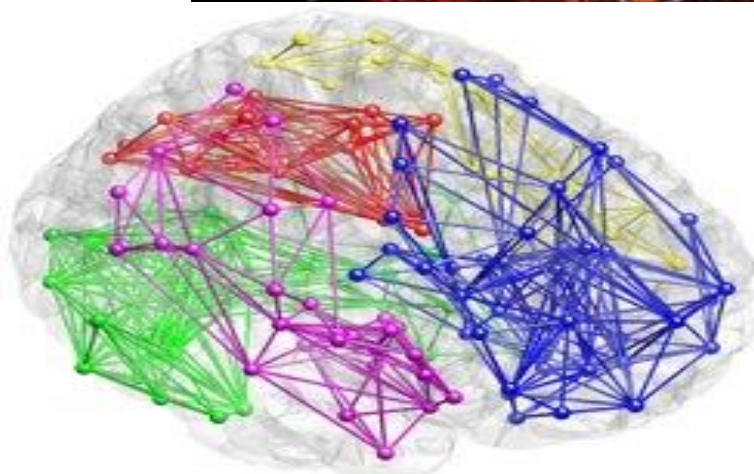
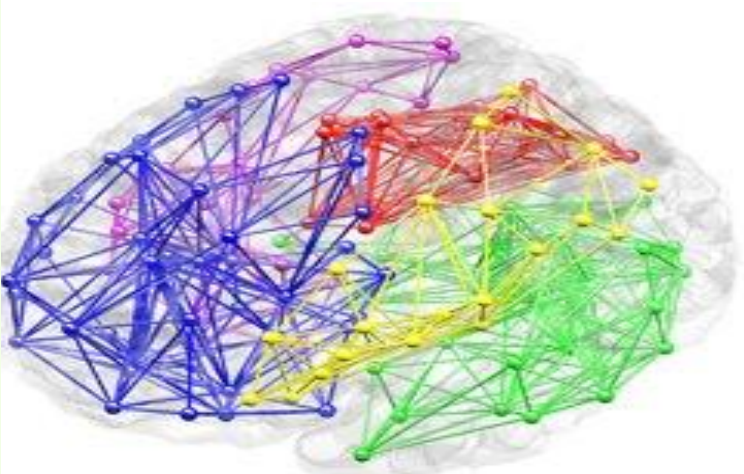
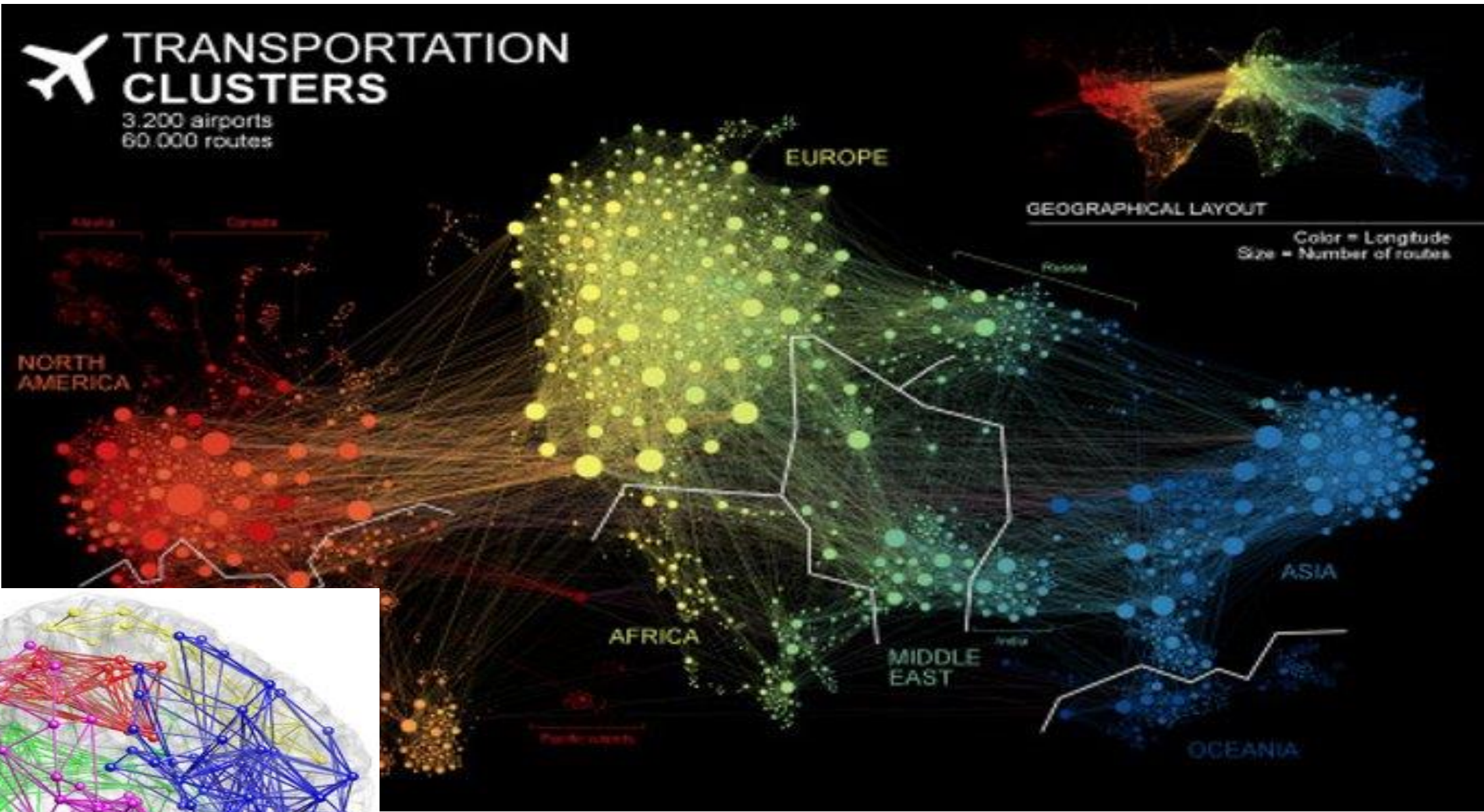
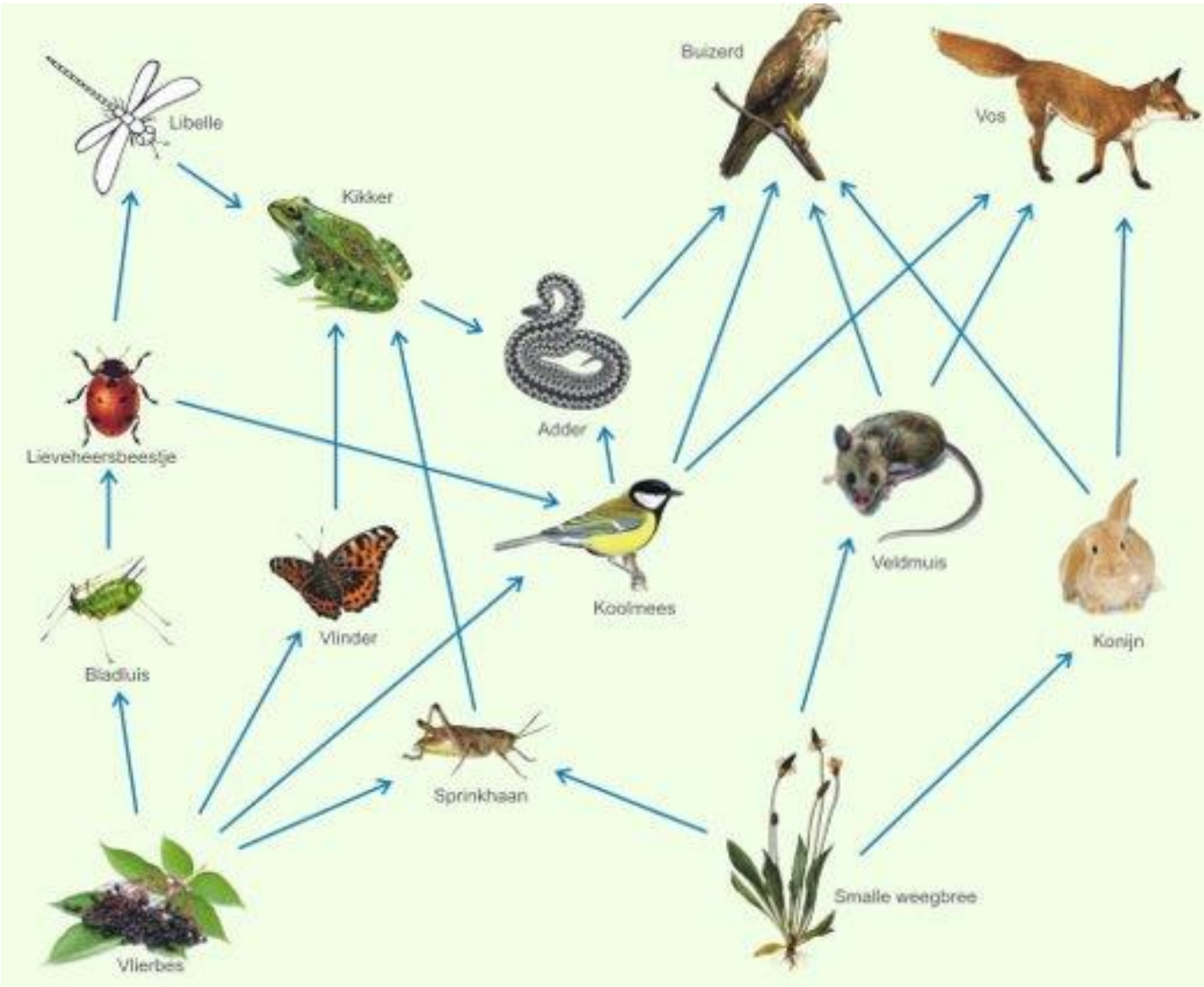


Figure 4.9 in the book: A First Course in NETWORK SCIENCE. Filippo Menczer, Santo Fortunato, Clayton A. Davis, ISBN: 9781108471138, Cambridge University Press.

4.3 PageRank vs in-degree

- The distribution of PageRank is similar to the distribution of in-degree on the Web.
- Among two pages with the same in-degree, the one linked by pages with higher PageRank wins the game.
- Search Engine Optimization (SEO) helps websites improve their search ranking.

4.4 Weighted Networks / examples



- Modularity structure
- prefrontal
 - medial parietal
 - temporo-occipital
 - centro-temporal (L)
 - centro-temporal (R)



WIKIPEDIA



Node centrality measures are **strength, in-strength, out-strength.**



4.5 Information and Misinformation / diffusion networks

- Diffusion network is a case study to explore the features of weighted directed networks.
- Related terms: trees, cascade trees (star trees), forest (set of trees), multilayer networks, and temporal networks (section 2 in the book).

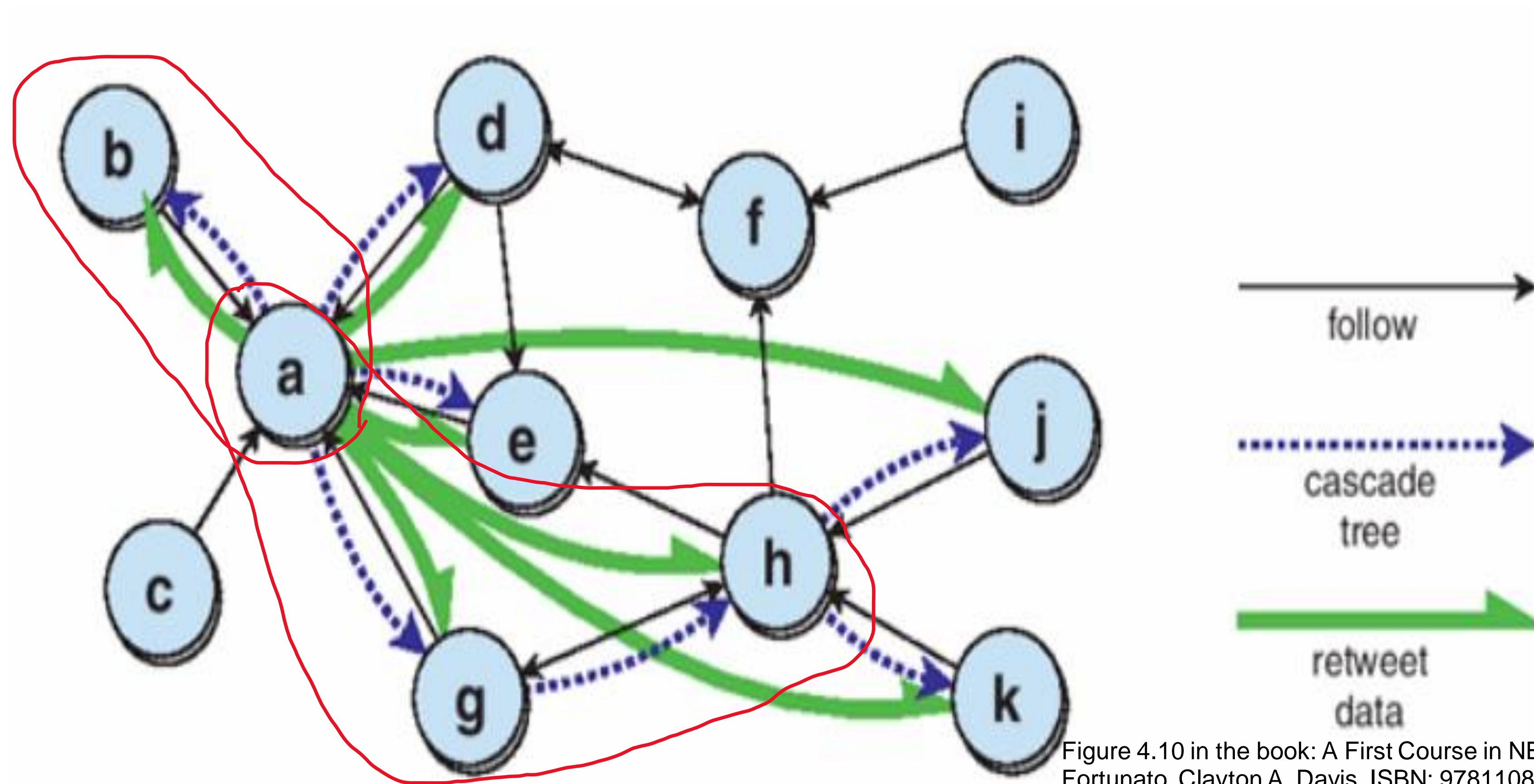
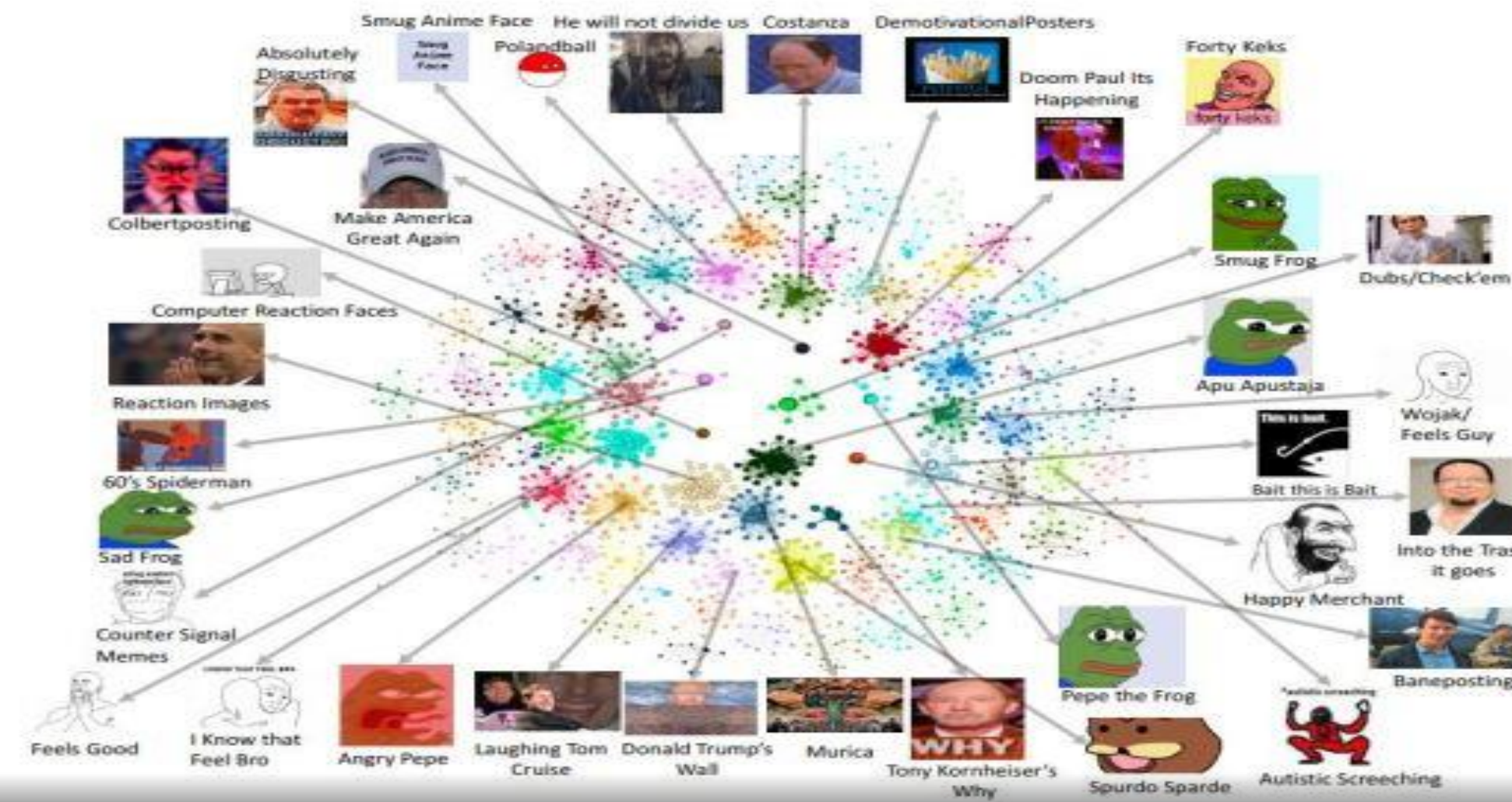


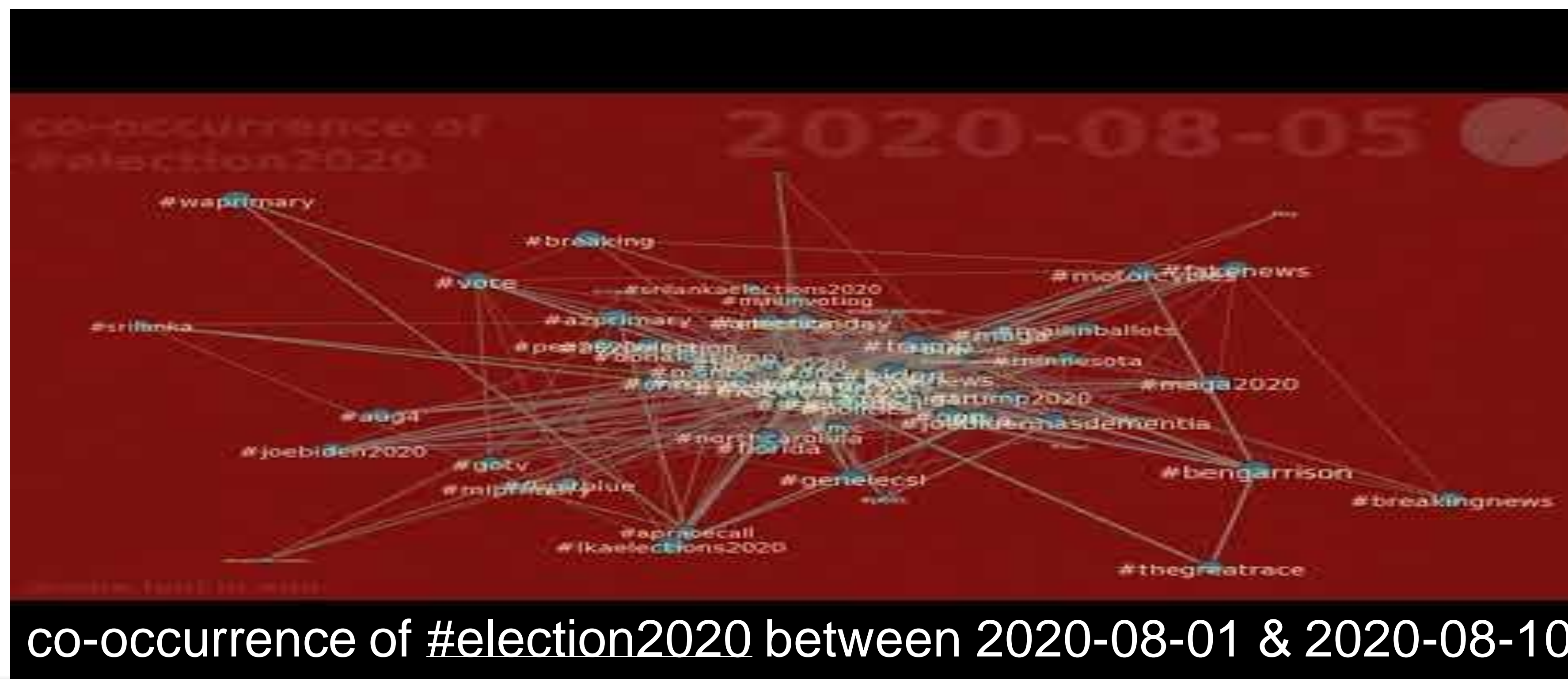
Figure 4.10 in the book: A First Course in NETWORK SCIENCE. Filippo Menczer, Santo Fortunato, Clayton A. Davis, ISBN: 9781108471138, Cambridge University Press.

Follower and retweet networks on Twitter.



4.5 Information and Misinformation / information diffusion networks

- People are nodes and pieces of information that are passed from a person to person are links.
- Example, Twitter diffusion networks are obtained by aggregating cascade forests across time and across many popular hashtags.



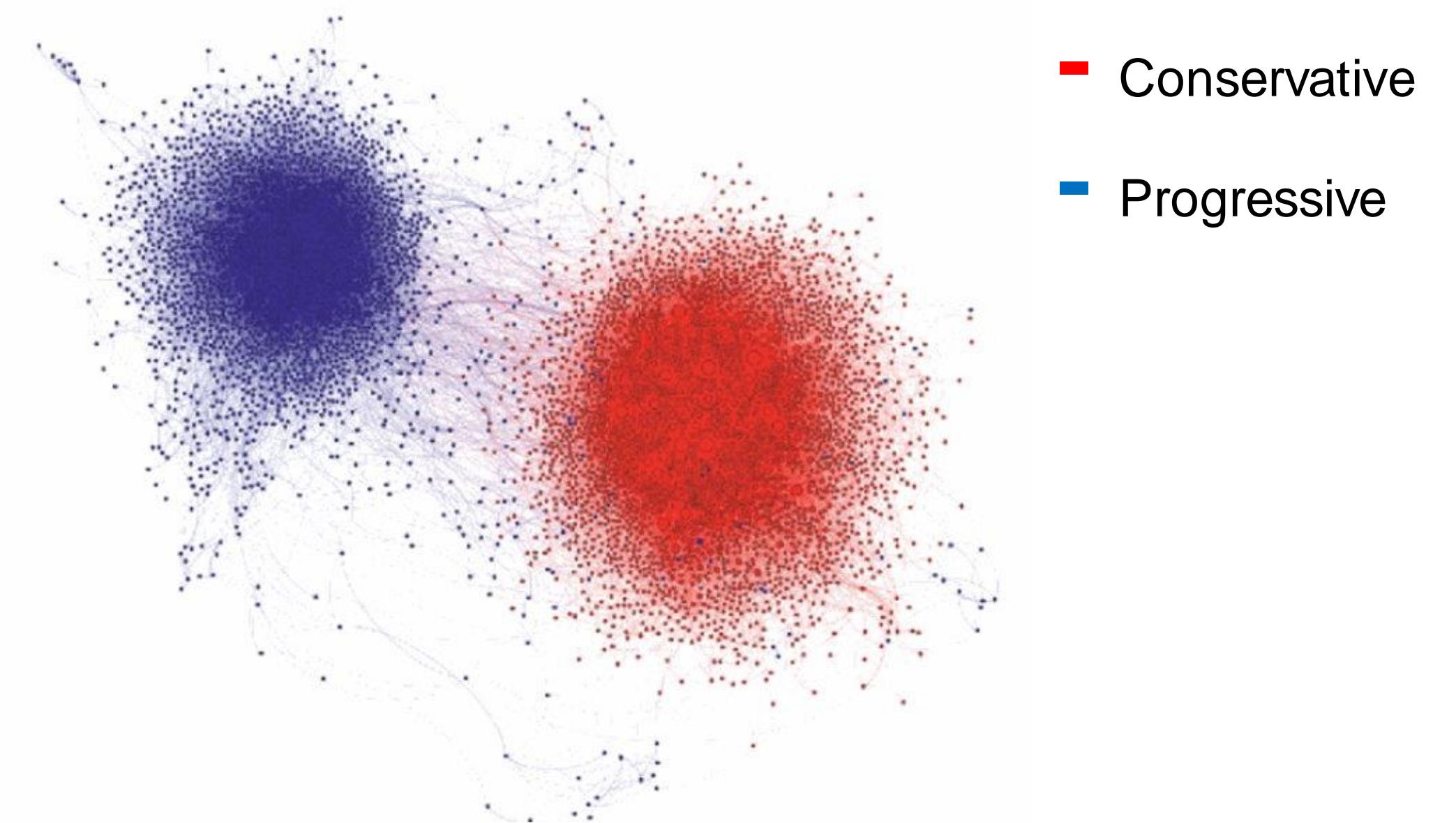
4.5 Information and Misinformation / diffusion networks / examples

Explore interactive diffusion networks from Twitter at

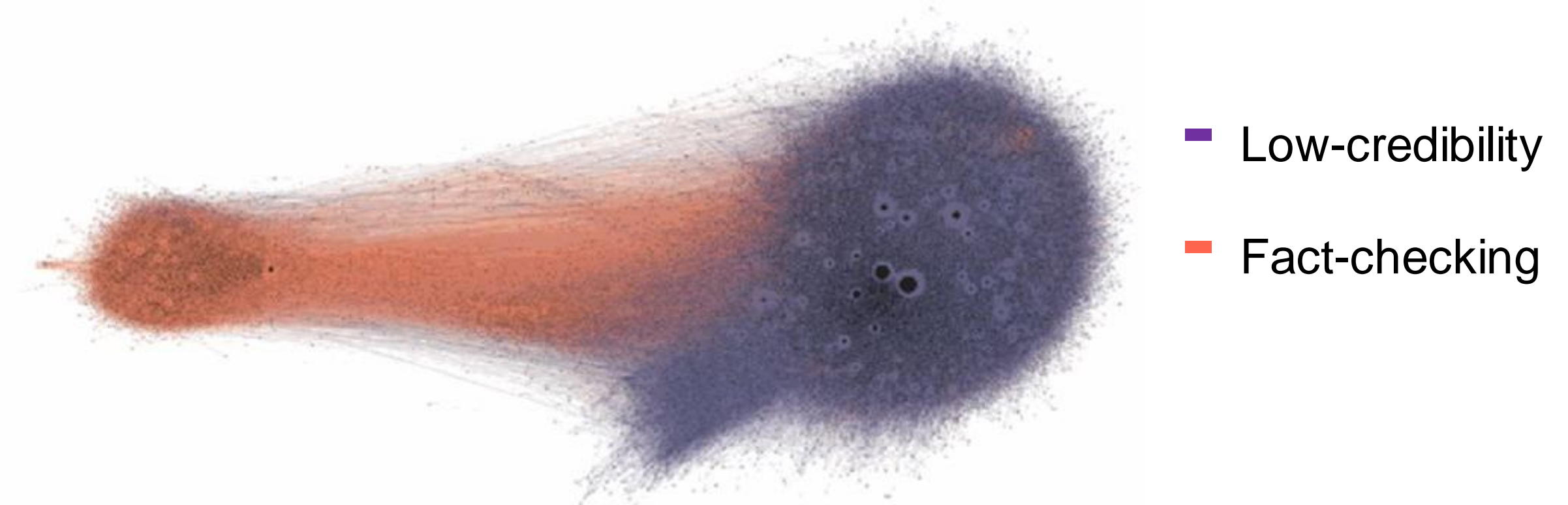
<http://osome.iuni.iu.edu/tools/movies/> (INDIANA UNIVERSITY OBSERVATORY ON SOCIAL MEDIA)

4.5 Information and Misinformation / features of diffusion networks

Echo chambers: is the spread uniform across the network or concentrated with dense clusters of network?



A user is mostly exposed to opinions reinforcing their own.

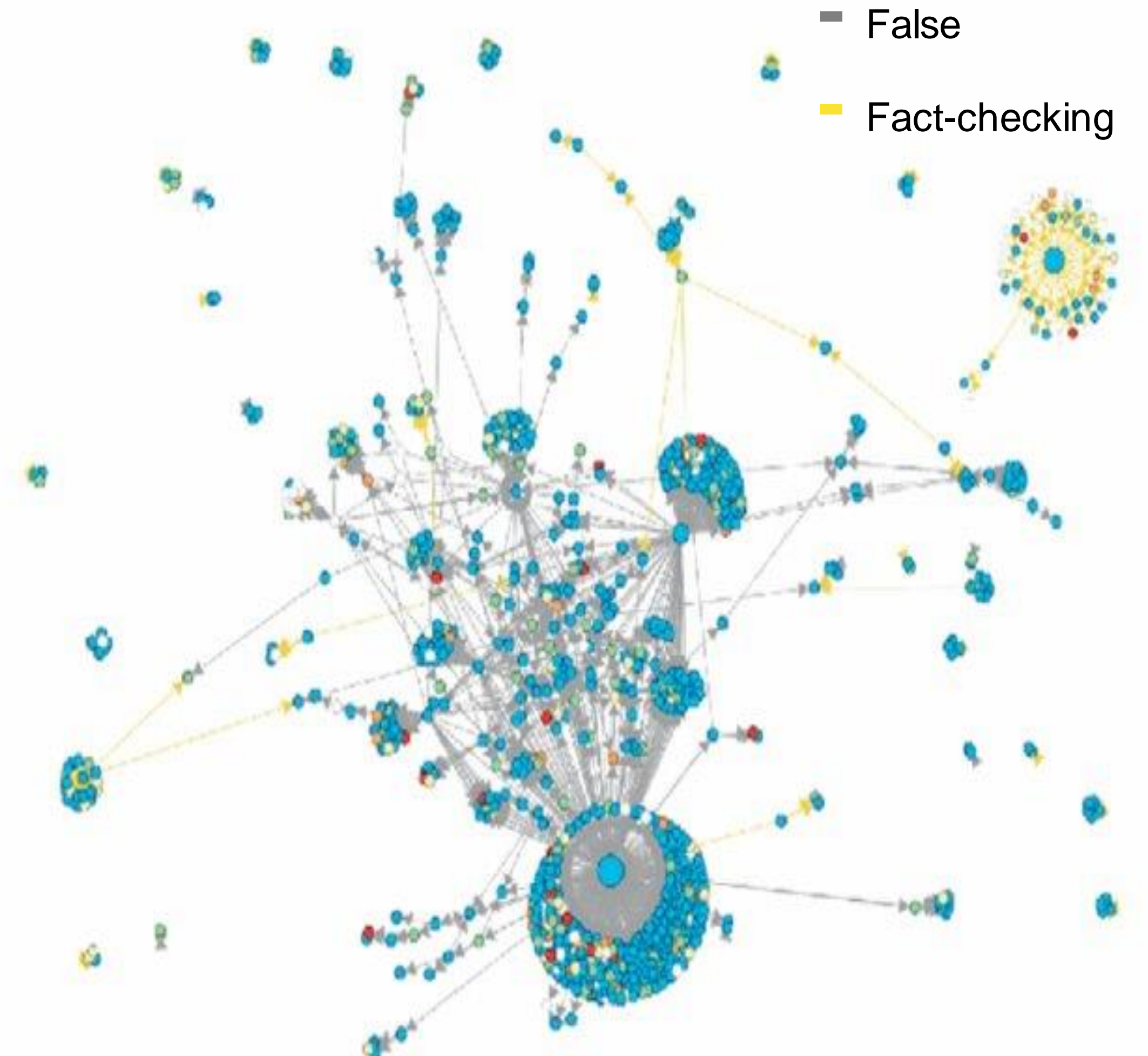


People who are vulnerable to political misinformation.

Figures 0.3 and 4.11 in the book: A First Course in NETWORK SCIENCE. Filippo Menczer, Santo Fortunato, Clayton A. Davis, ISBN: 9781108471138, Cambridge University Press.

4.5 Information and Misinformation / features of diffusion networks

Virality: It can be illustrated by the number of users exposed to a message.

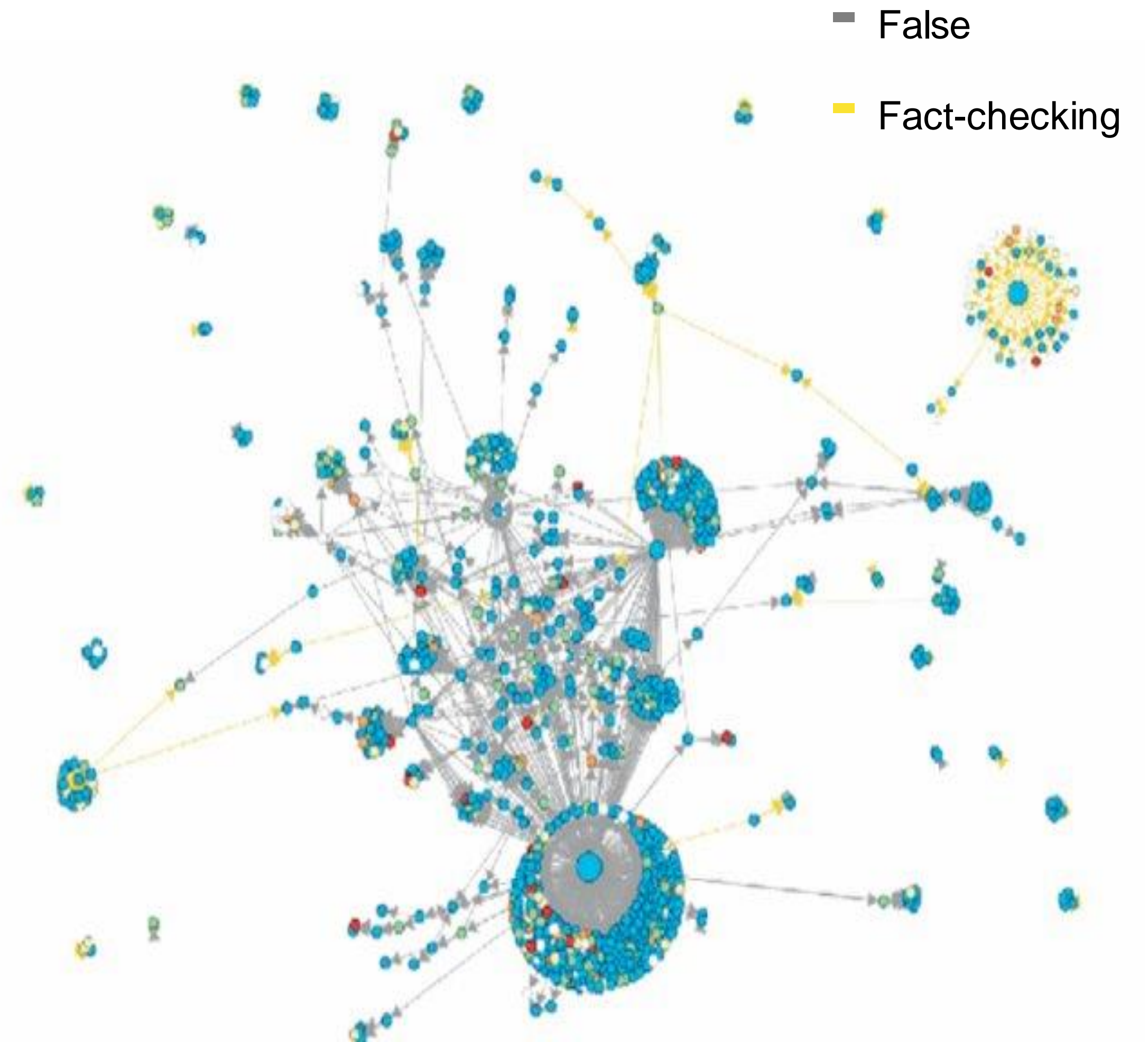


The spread of misinformation is more viral than actual news.

Figures 4.12 in the book: A First Course in NETWORK SCIENCE. Filippo Menczer, Santo Fortunato, Clayton A. Davis, ISBN: 9781108471138, Cambridge University Press.

4.5 Information and Misinformation / features of diffusion networks

Influence: a high value of out-strength.

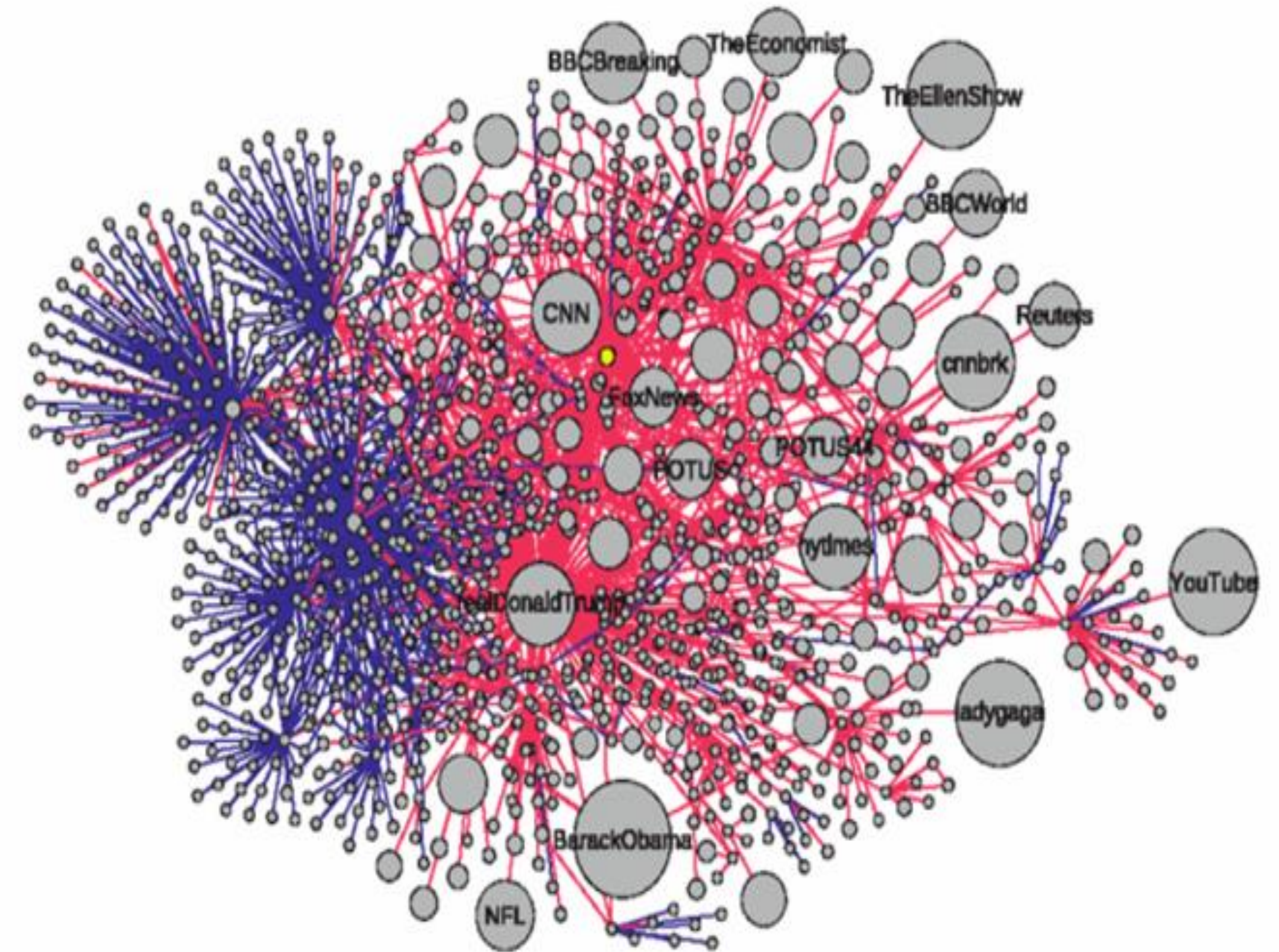


Node size is proportional to out-strength.

4.5 Information and Misinformation / social bots

Bots can gain significant influence.

- links, the article spread by retweet or quoted tweet
- links, the article spread by replies or mentions.



Node size represents the number of followers of an account. Yellow node is a bot.



Tools (beta)

Misinformation Tools

Tools created by the OSoMe team to help combat the spread of misinformation.

Hoaxy



Visualize the spread of claims and fact checking.

Botometer



Check how bot-like a Twitter user behaves.

Fakey



Play this game to learn to recognize fake news on your social feed.

BotSlayer



Set up your own system to detect coordination and bot amplification on Twitter.

EchoDemo



Simulation demonstrating how two basic mechanisms of social media can lead to polarized social networks.

Bot Electioneering Volume



Visualize the activity of likely bots on Twitter around the 2018 US midterm elections.

CoVaxxy



Visualizing the relationship between COVID-19 vaccine adoption and online (mis)information.

Explore Misinformation tools at <http://osome.iuni.iu.edu/tools/>
(INDIANA UNIVERSITY OBSERVATORY ON SOCIAL MEDIA)



Visualize the spread of claims and fact checking.

Search by: Twitter Articles Language: Any ▼

Show: ☐ Recent ☐ Popular ☒ Mixed

Search Or Import Data ?

Trending News

WISN.COM | More than 1,800 COVID-19 vaccines wasted across Wisconsin - WISN Milwaukee [↗](#)
[Search Title](#) [Search Link](#)

CNBC.COM | Bitcoin surpasses \$60,000 in record high as rally accelerates - CNBC [↗](#)
[Search Title](#) [Search Link](#)

CNN.COM | Here is what the potential impeachment of Gov. Andrew Cuomo could look like [↗](#)
[Search Title](#) [Search Link](#)

Popular Claims

DAILYWIRE.COM | Gina Carano To Produce And Star In Upcoming Film For The Daily Wire [↗](#)
[Search Title](#)

BABYLONBEE.COM | Gina Carano Rehired By Disney After She Identifies As An Abusive Male... [↗](#)
[Search Title](#)

BABYLONBEE.COM | Texas Removes Mask Mandate To Scare All The Californians Away [↗](#)
[Search Title](#)

Popular Fact-Checks

TRUTHORFICTION.COM | Law Professor Refutes Trump Attorneys' Attempt to Cite Him in Impeachment Hearing [↗](#)
[Search Title](#)

TRUTHORFICTION.COM | 'My Brain Refuses to Believe There Are 4 People in This Photo' [↗](#)
[Search Title](#)

TRUTHORFICTION.COM | Far-Right U.S. Legislators Using Disinformation to Gain, Strengthen Power [↗](#)
[Search Title](#)

Explore the role of bots <https://hoaxy.osome.iu.edu/>.

Summary:

- Search engines use network **centrality** measures as ranking criteria.
- In weighted networks, node centrality measures are **strength, in-strength, out-strength**.
- Features of diffusion networks such as **echo chambers, virality, and influence** are used in investigating online misinformation.

Exercise: Chapter 4 tutorial notebook.

<https://github.com/CambridgeUniversityPress/FirstCourseNetworkScience/blob/master/tutorials/Chapter%204%20Tutorial.ipynb>