Understanding social networks with F#

Evelina Gabasova @evelgab

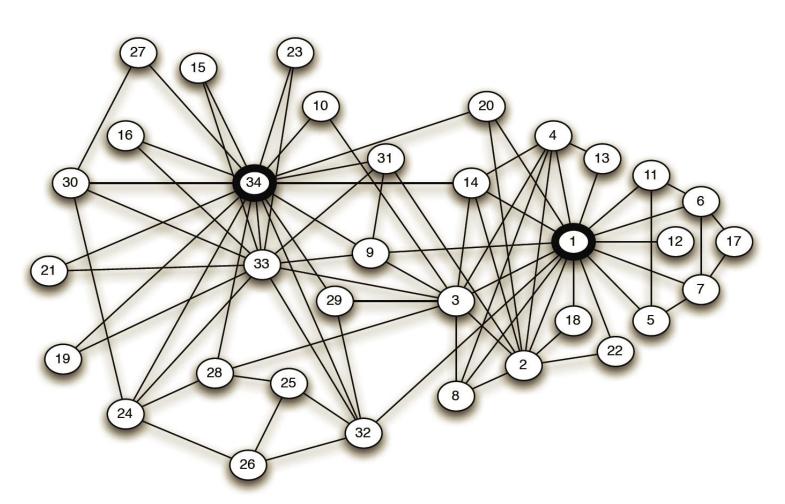
Why network science

- gene interaction networks
- disease spreading
- balancing load in power grids
- brain connections
- software dependency networks
- social networks

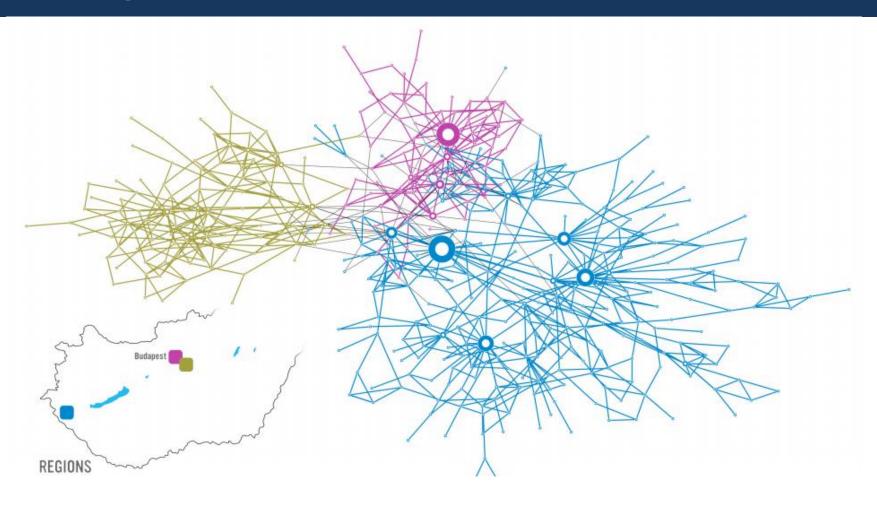
•

Social network analysis

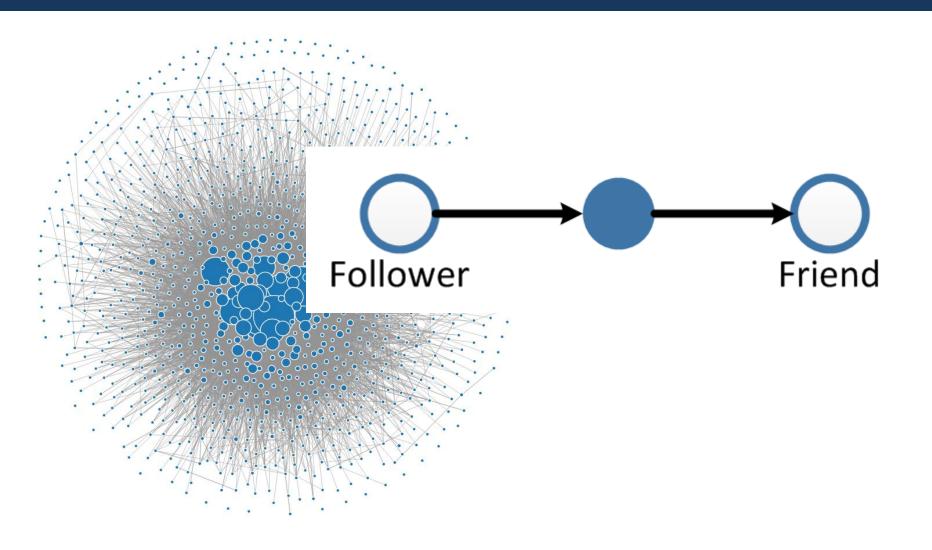
Karate club network



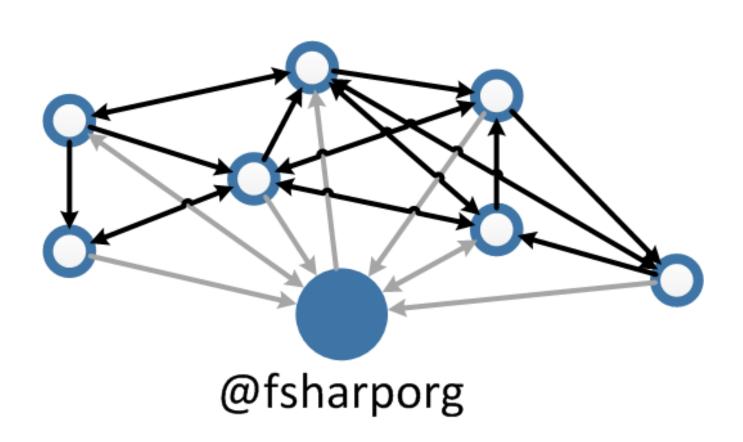
Insights from social networks



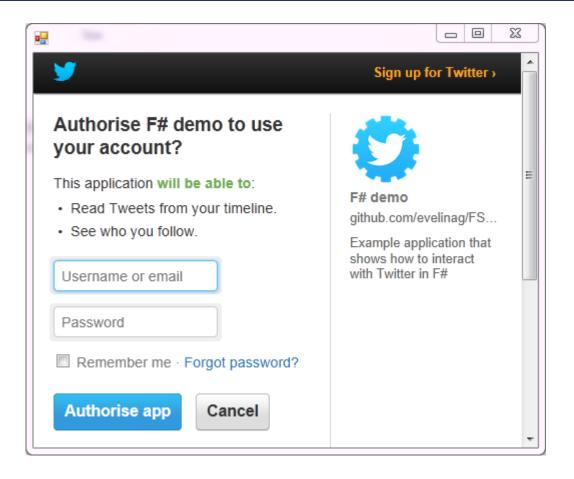
Twitter network



How large is your ego?



Connecting to twitter





Interactions on Twitter

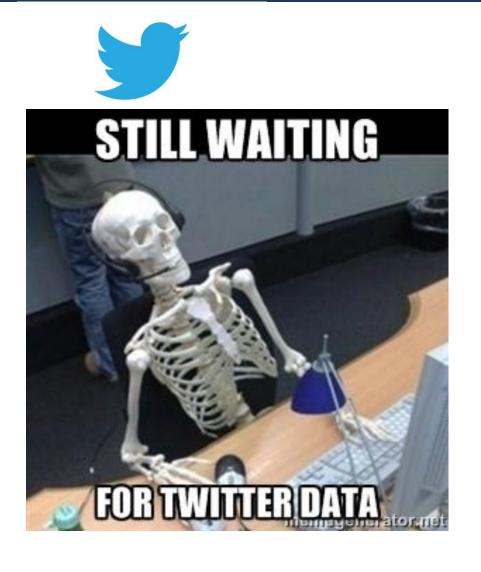
Are numbers of

Retweets and Favorites

correlated?



Downloading data

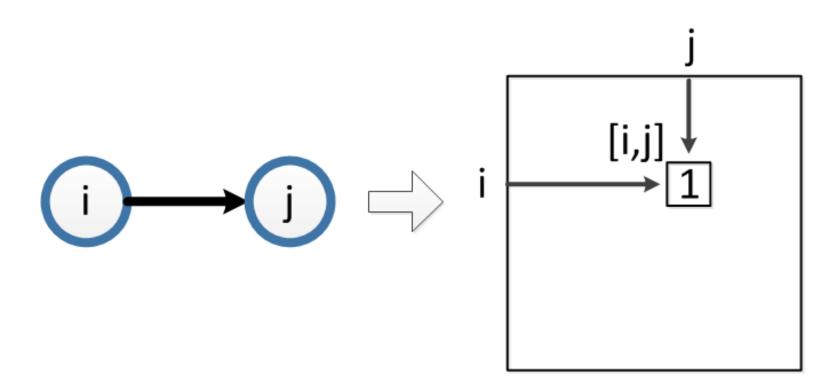


- 1) List of nodes
- Connections between nodes

Twitter API allows only 15 requests every 15 minutes to list connections.

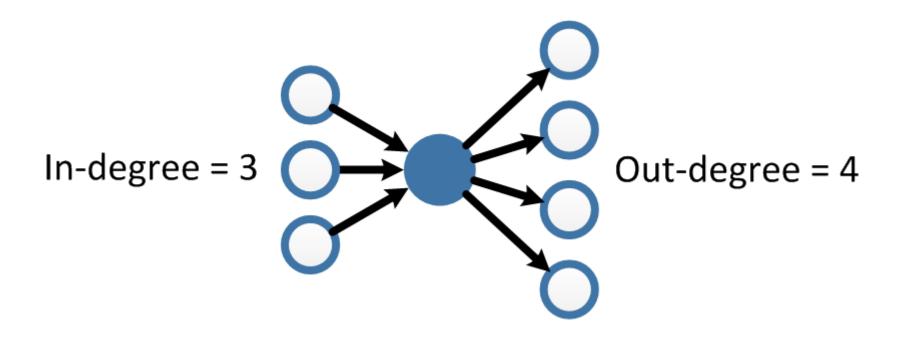


Adjacency matrix

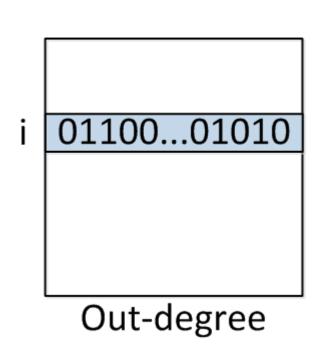


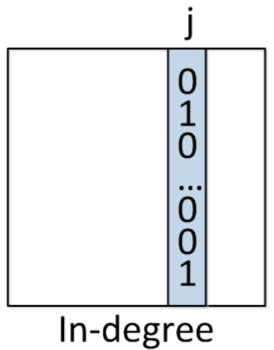


Degrees



Degrees

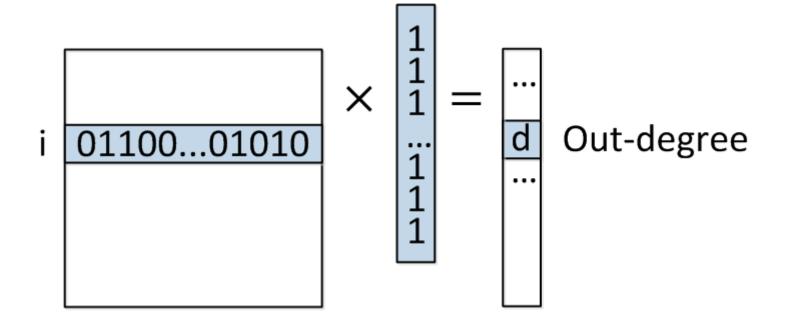






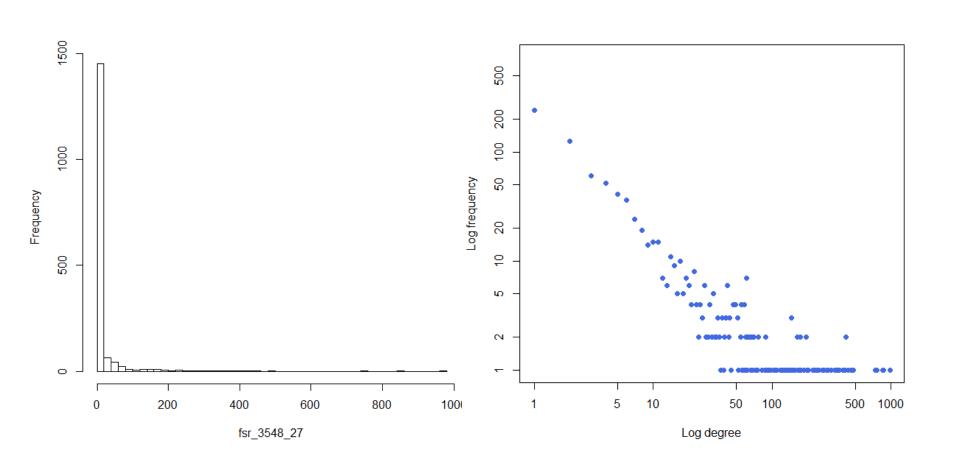


Degrees





Degree distribution

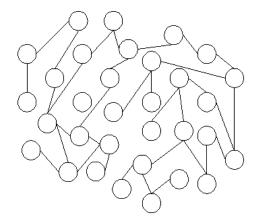


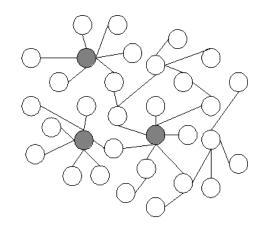
Scale-free networks

Power law

$$P(d) \sim d^{-\gamma}$$

- Networks growing over time with preferential attachment
- Hubs
- Robustness





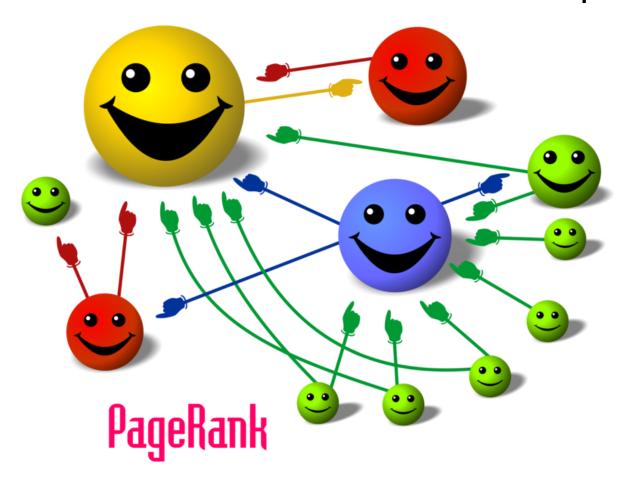
Your friends have more friends than you have.

Hubs in networks

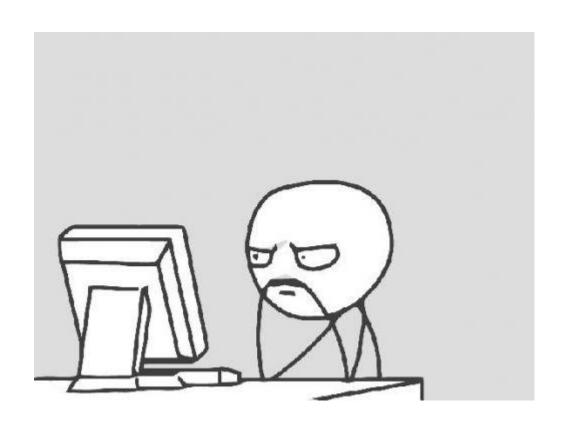
How to identify most important nodes in a network?

Centrality with PageRank

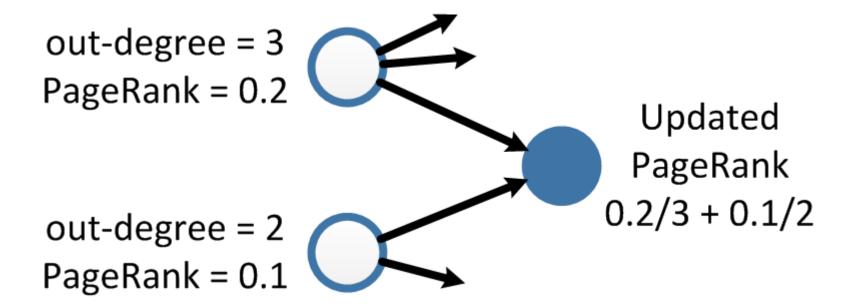
Your followers are not created equal.



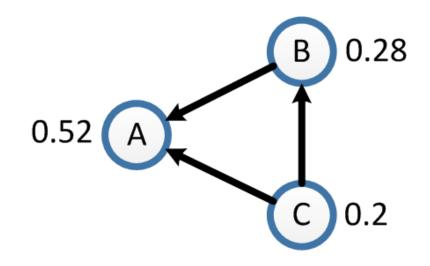
Random surfer model

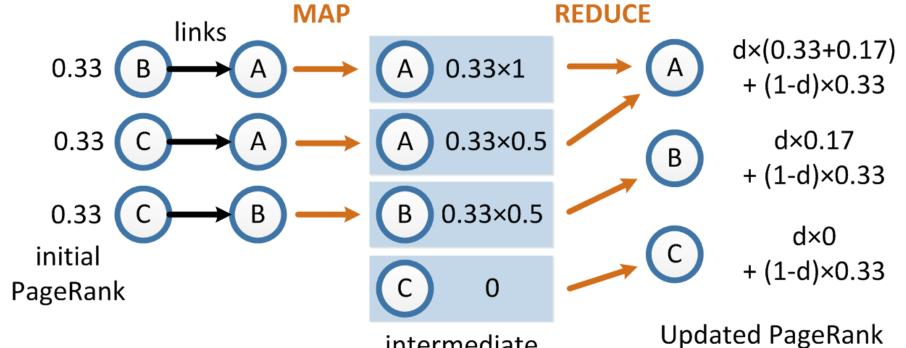


Centrality with PageRank



+ random jumps





intermediate

key/value pairs



PageRank changes

February

- 1. migueldeicaza (0.033130)
- 2. dsyme (0.032783)
- 3. tomaspetricek (0.027756)
- 4. LincolnAtkinson (0.021993)
- 5. VisualFSharp (0.020233)
- 6. c4fsharp (0.019720)
- 7. rickasaurus (0.019189)
- 8. ptrelford (0.018099)
- 9. 1tgr (0.016525)
- 10. sforkmann (0.014970)

September

- 1. dsyme (0.028640)
- 2. migueldeicaza (0.024808)
- 3. VisualFSharp (0.024479)
- 4. tomaspetricek (0.021066)
- 5. c4fsharp (0.019612)
- 6. rickasaurus (0.014272)
- 7. sforkmann (0.013471)
- 8. 1tgr (0.012768)
- 9. ptrelford (0.012669)
- 10. FSPowerTools (0.012113)

PageRank changes

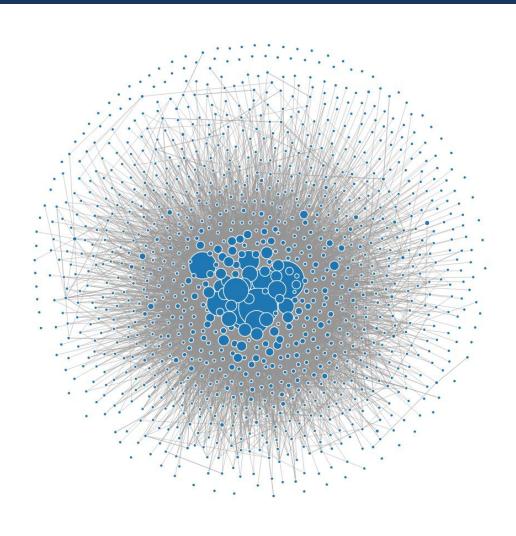
February

- 1. migueldeicaza (0.033130)
- 2. dsyme (0.032783)
- 3. tomaspetricek (0.027756)
- 4. LincolnAtkinson (0.021993)
- 5. VisualFSharp (0.020233)
- 6. c4fsharp (0.019720)
- 7. rickasaurus (0.019189)
- 8. ptrelford (0.018099)
- 9. 1tgr (0.016525)
- 10. sforkmann (0.014970)

September

- 1. dsyme (0.028640)
- 2. migueldeicaza (0.024808)
- 3. VisualFSharp (0.024479)
- 4. tomaspetricek (0.021066)
- 5. c4fsharp (0.019612)
- 6. rickasaurus (0.014272)
- 7. sforkmann (0.013471)
- 8. 1tgr (0.012768)
- 9. ptrelford (0.012669)
- 10. FSPowerTools (0.012113)

Visualization with D3.js



Go play with data!

Thank you

@evelgab evelina@evelinag.com

fsharp.org

F# eXchange 2015
17 April, London

