Universität Zürich[UZH]

Blockchain & Distributed Ledger Technologies

UZH Blockchain Center

Centrality Measures

Network Science '21: Assignment 3

Prof. Dr Claudio J. Tessone, Dr Carlo Campajola

Blockchain & Distributed Ledger Technologies Group

Objectives

1. Gain intuition on the different centrality measures and their mutual relation
2. Explore the effects of randomisation on nodes properties in different scenarios
3. Understand the role of the parameters in PageRank algorithm

A03.1 Centrality

A03.1 Centrality correlations

Task: *Learn about centrality measures and how they correlate in different networks.*

For the given networks compute:

- the degree k_i, closeness c_i, betweenness b_i and eigenvector e_i centrality for each node;
- then do a scatter plot for each pair of centralities (6 plots total);
- compute Pearson's, Spearman's and Kendall's correlation coefficients for each pair and note them on the scatter plots;
- briefly explain, for the Jazz collaborations data, what each of the centrality measures means in practice.

A03.1 Hints

- Centrality measures are easily available from `networkx`;
- Correlation coefficients can be imported from the `scipy.stats` module.

A03.1 Correlation coefficients

- Pearson's r correlation: $r(X,Y) = \frac{\mathrm{Cov}(X,Y)}{\sigma_X \sigma_Y}$
- Spearman's ρ: $\rho(X,Y) = r(\mathrm{rank}(X), \mathrm{rank}(Y))$
  where $\mathrm{rank}(X)$ is the sorting order ("ranking") of X
- Kendall's τ: $\tau(X,Y) = \frac{1}{\binom{N}{2}} \sum_{i<j} \mathrm{sign}(X_i - X_j)\mathrm{sign}(Y_i - Y_j)$

  intuitively $\frac{\text{\# concordant ranking pairs - \# discordant ranking pairs}}{\text{\# pairs}}$

A03.2 Centrality and randomisation

Task: *Explore the effect of rewiring randomisation on centrality metrics.*

- For each dataset $\mathcal{G}$, randomise it to obtain a network $\mathcal{G}_{rnd}$ (same procedure as previous assignments).
- Do a scatter plot of each centrality before vs. after randomisation (e.g. k_i^{rnd} vs k_i) and compute the corresponding correlation coefficient.
- Why do you get that result for degree centrality?

A03.3 PageRank

A03.3 Page Rank

In the dataset "Florida_foodweb.graphml" each node is a species, and a directed link exists between i and j if i feeds from j

> ***Task:*** *Rank the nodes' importance by means of the PageRank algortihm.*

- Use different values of α, e.g.
  $\alpha = 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.85, 0.9, 0.95, 0.99$
- Compute Spearman's rank correlation ρ between the PageRanks for $\alpha_0 = 0.85$ and all other αs, then plot them as a scatter plot with α on the x-axis and $\rho(\alpha)$ on the y-axis.

A03.1-2 Datasets provided

- 9-11 terrorist network: Nodes represent individuals and Edges represent their known social associations, centered around the hijackers that carried out the September 11th, 2001 terrorist attacks **[1]**
- Jazz collaboration network: Nodes represent jazz musicians and Edges represent collaborations in bands that performed between 1912 and 1940 **[2]**
- Political blogs network: Nodes represent political weblogs from before the 2004 election and Edges represent hyperlinks among them **[3]**

A03.1-2 Datasets provided

[1] V. Krebs, ”Mapping networks of terrorist cells.” Connections 24, 43-52 (2002)
[2] P. Gleiser and L. Danon, ”Community Structure in Jazz.” Advances in Complex Systems 6(4), 565-573 (2003)
[3] L. A. Adamic and N. Glance, ”The political blogosphere and the 2004 U.S. election: divided they blog.” Proc. 3rd Internat. Workshop on Link Discovery (LinkKDD), 36-43 (2005)

Universität Zürich[UZH]

Blockchain & Distributed Ledger Technologies

UZH Blockchain Center

Prof. Dr Claudio J. Tessone, Dr Carlo Campajola

Blockchain & Distributed Ledger Technologies Group

tessone@ifi.uzh.ch

http://www.blockchain.uzh.ch