Network Measures

Saptarshi Pyne
Assistant Professor
Department of Computer Science and Engineering
Indian Institute of Technology Jodhpur, Rajasthan, India 342030

CSL7390 Social Network Analysis Lectures 3-8
January 10-24th, 2024

What we discussed in the last class

- Graph theoretical terminologies and algorithms

Centrality

- Given a vertex, how important or ‘central’ is it to the whole network?

Degree centrality

- Idea: How many neighbours a vertex has?

- Undi
  - Degree
- Di
  - In-degree
  - Out-degree

Eigenvector centrality [Bonacich, 1987]

- Additionally: What kind of neighbours a vertex has? (quantity or quality or both)

- Undi
  - The leading eigenvector
- Di
  - Right eigenvector

PageRank centrality [Brin and Page, 1998]

- US Patent, 1998:
  https://patents.google.com/patent/US6285999

- Paper:
  Brin and Page (1998). The anatomy of a large-scale hypertextual web search engine. Computer networks and ISDN systems, 30(1-7), 107-117.

An application of PageRank centrality

Q. How does the Google Search work?

Step 1: Google crawls all the webpages on the Internet and assign a set of text indices (such as ‘Computer Science’, ‘AI’) to each webpage.

Step 2: When a user enters a search query, it matches the keywords in the search query with the text indices. Thus, Google finds a subset of webpages relevant to the query.

An application of PageRank centrality (contd.)

Step 3: Google then ranks the selected webpages using the PageRank algorithm. Google also separately ranks the selected webpages with a few more algorithms. The final ranks of the webpages are decided by taking a consensus of all the algorithms.

Google has been the #1 search engine in the world

- not because it finds unique search results given a search query
- but because it ranks the search results correctly. It does so by calculating the importance of each webpage correctly w.r.t. a given search query.

Geodesic path(s) from one vertex to another

The shortest path(s) between two given vertices is called their geodesic path(s).

The length of the geodesic path is called their **geodesic distance**.

If there do not exist any paths between the concerned vertices, then the geodesic distance is infinity.

Mean farness and closeness of a vertex

Mean farness of other vertices from the i^{th} vertex

$$\ell_i = \frac{1}{n-1} \sum_{j(\neq i)} d_{ij}$$

Mean closeness of the i^{th} vertex

$$C_i = \frac{1}{\ell_i}$$

Closeness centrality

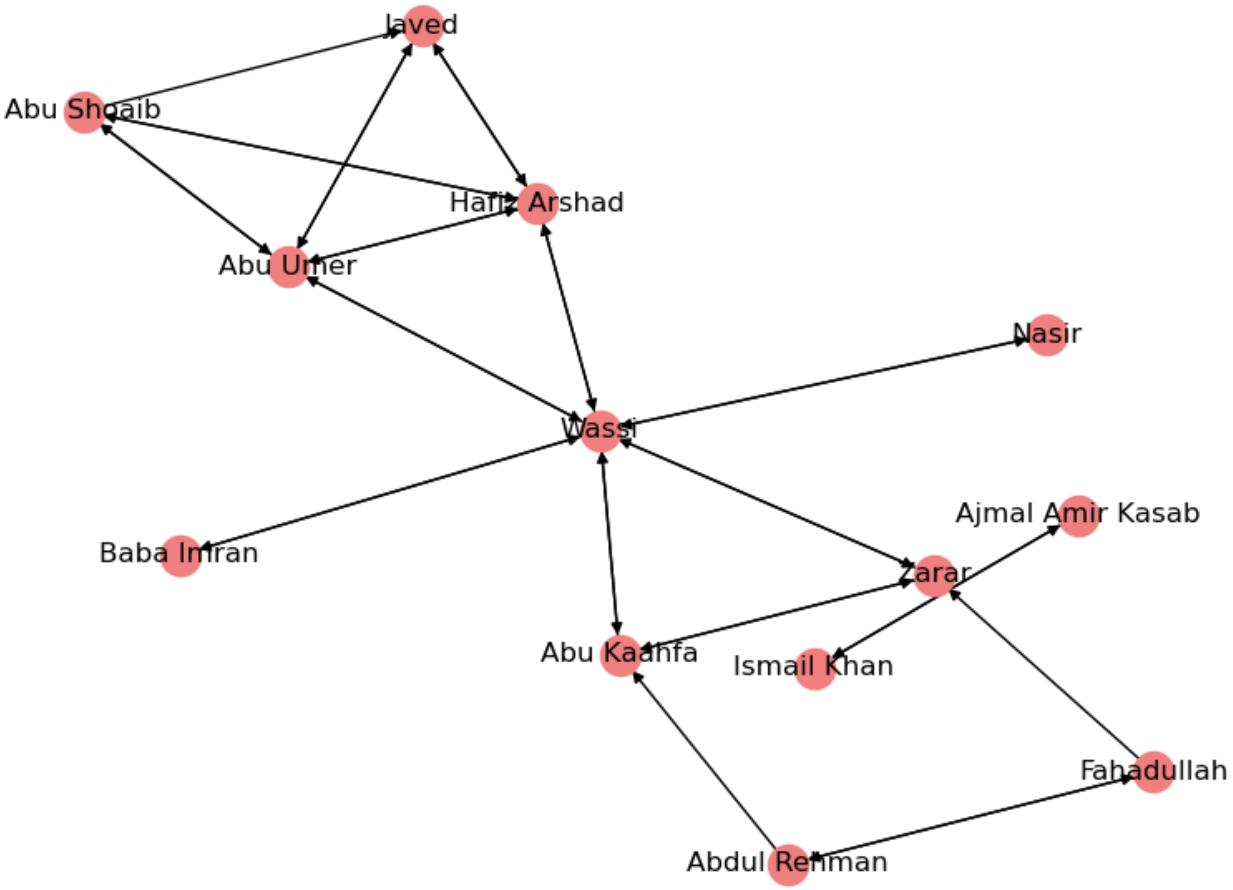
Limitation of the mean closeness:

When a network has multiple components, the mean closeness of every single vertex will be zero.

Resolution: **Closeness centrality**

$$C_i' = \frac{1}{n-1} \sum_{j(\neq i)} \frac{1}{d_{ij}}.$$

Analysis of the phone calls between the Mumbai-attacks terrorists and their handlers (26/11/2008)



Who called whom during the attacks

Thoughts

- What is not a social network?
- What is not a society?

References

- Pages 168-193, ‘Networks’ by Mark Newman, Oxford University Press, 1st edition, 2010.

- https://u.osu.edu/nix.39/2021/01/22/breaking-down-the-eigenvector-centrality-measure/

Thank you