7. Closeness and Betweenness Centrality

Bookmark this page

Exercises due Oct 20, 2021 17:29 IST Completed

Citation Network

1/1 point (graded)

Consider the citation graph where there is a directed edge from paper i to j if i has cited j as a reference. Assume that there may or may not be papers that cite one another. Among the following centrality measures, which one is most appropriate as a measure of importance for each paper?

- ○ Eigenvector centrality
- ◉ Page-rank centrality
- ○ Katz centrality

✔

Solution:

Page-rank centrality is the most appropriate. It overcomes all of the issues associated with eigenvector and Katz centrality measures. A paper that is cited by a very important paper such as a survey paper should not get very high importance just for the fact that it was cited by a survey paper. Further, page-rank centrality works just fine even if the citation graph turns out to be a DAG.

Submit You have used 1 of 2 attempts

ℹ Answers are displayed within the problem

Twitter

1/2 points (graded)

Consider a Twitter dataset that consists of the following data: users $i = 1, \ldots, n$, list of followers f_i for each user i.

1. What is the most appropriate graph type that could model the dataset?

- ◉ Simple, directed graph
- ○ Simple, undirected graph
- ○ Simple, bipartite graph
- ○ Simple, weighted (directed) graph
- ○ Multigraph

✔

2. Which centrality measure of the ones studied so far is most suited to capture the importance of the users in the network?

- ○ Eigenvector centrality
- ○ Page-rank centrality ✔
- ◉ Katz centrality

✖

Solution:

1. The user graph with follows represented as directed edges is a simple, directed graph.

2. Similar to the citation network, page-rank centrality is the most appropriate measure of centrality here.

Submit You have used 2 of 2 attempts

Answers are displayed within the problem

Closeness Centrality and Betweenness Centrality

The **closeness centrality** of a node i is the reciprocal of average distance of the node to every other node. Recall that d_{ij} is the shortest path length (smallest weight path, in weighted graphs) between node i and j. The closeness centrality of node i is given as

$$C_i = \left(\frac{1}{n-1} \sum_{j \neq i} d_{ij} \right)^{-1}$$

The measure of “high" or “low" value of closeness centrality is relative. A high closeness centrality C_i, relative to the closeness centrality of the other nodes, indicates that node i is relatively close to all the other nodes.

The **betweenness centrality** measures the extent to which a node lies on paths between other nodes.

$$B_i = \sum_{s \neq i \neq t} \frac{n_{st}^i}{g_{st}},$$

where n_{st}^i is the number of shortest paths between s and t that pass through i and g_{st} is the total number of shortest paths between s and t. Note that this considers both orderings of each pair of nodes, so for undirected graphs, a path counts twice (as it counts both for n_{st}^i and for n_{ts}^i).

Closeness Centrality in a Circle Graph

1/1 point (graded)

Consider the circle graph of n nodes: the nodes all have degree 2 and are linked to form a circle. Assume that n is odd and that $n \geq 3$. Compute the closeness centrality of the nodes.

4/(n+1) ✔ **Answer:** 4/(n+1)

Solution:

For any node, given that n is odd,

- Two nodes are at distance 1.
- Two nodes are at distance 2.
- So on...
- Two nodes are at distance $(n-1)/2$.

Therefore, the closeness centrality is equal to

$$\left(\frac{1}{\quad} \right)^{-1} \qquad \left(\frac{1}{\left[\left(\qquad \right) \right.} \right)^{-1}$$

$$\left(\frac{1}{n-1}\sum_{j\neq i} d_{ij}\right) = \left(\frac{1}{n-1}\left[2\left(1+2+\cdots+\frac{n-1}{2}\right)\right]\right)$$

$$= \left(\frac{1}{n-1}\left[2\frac{(n+1)(n-1)}{8}\right]\right)^{-1}$$

$$= \left(\frac{n+1}{4}\right)^{-1}$$

$$= \frac{4}{n+1}$$

Submit You have used 2 of 3 attempts

Answers are displayed within the problem

Betweenness Centrality in a Line Graph

1/1 point (graded)

Consider the undirected line graph on n nodes: the nodes form a straight line. Let the nodes be numbered $1, 2, \ldots, n$. Compute the betweenness centrality of node i (use the formula defined on this page, rather than the one in the lecture video):

2*(i-1)*(n-i) ✔ **Answer:** 2*(i-1)*(n-i)

Solution:

Let the node indicies increase from left to right in the graph. The only paths that pass through node i are those that start on the left and end to the right of i, or vice versa. There are $(i-1)$ nodes to the left of node i, and $(n-i)$ nodes to the right of node i. Thus, there are $(i-1)(n-i)$ paths that pass through node i (from left to right). Now, note that there are the same number of paths that pass from right to left.

So the betweenness centrality is $2(i-1)*(n-i)$.

Submit You have used 2 of 3 attempts

Answers are displayed within the problem

Discussion

Topic: Module 3: Network Analysis:Graph Centrality Measures / 7. Closeness and Betweenness Centrality

Hide Discussion

Add a Post

All Posts

when will eigenvector and katz centrality be advantageous compared with page-rank?

discussion posted 2 months ago by **zcypaul**

The same as the question title. I just wonder if we should always use page-rank centrality? What's the limitation of page-rank?

This post is visible to everyone.

Add a Response

1 response

vllgalys (Staff)
2 months ago - endorsed 2 months ago by **lam_trinh** (Community TA)

Eigenvector is generally advantageous in undirected graphs. Remember that the distinction between Katz and Page-Rank is that Page-Rank attributes a portion of the score of a network depending on how many degrees are connected. This is definitely advantageous for traversing the internet, but there are many settings (e.g. traffic design) where the number of connections does not dilute the quality of any one edge.

Add a comment

Showing all responses

Add a response:

Preview

Submit

‹ Previous Next ›