

Homework

Networks

11/15/2015

Due: Sunday, November 22.

In this homework, you will explore different community detection algorithms.

Part 1

The igraph package has a few community detection algorithms implemented. You will run these algorithms on Zachary's karate club data set where we know the ground truth and investigate what communities you get out. Visualize communities obtained by different algorithms. Briefly comment on outputs of different algorithms and their ability to uncover the true communities.

By default, algorithms compute communities that give highest modularity. If you are using an algorithm that is hierarchical you can cut the resulting tree to obtain a different number of communities. See the starter script for more details.

Part 2

Load the network wikipedia.gml provided. It is in gml format, which can be imported into `__igraph` using the following command

```
read.graph("wikipedia.gml", format="gml")
```

The vertices of this network are wikipedia pages. The label of each vertex is the title of the wikipedia page.

Now use any community detection algorithm. Do you think the communities found make sense? You can use the vertex labels to check this.

Note that some algorithms work only on undirected graphs. Also note that some algorithms are too slow to be run on big networks. Check the manual to see if you can run the community detection algorithm on the wikipedia graph.