Assignment Seven

Sybil Melton

November 5, 2014

CONTENTS

LISTINGS

1 KARATE CLUB D3 GRAPH

Using D3, create a graph of the Karate club before and after the split.

- Weight the edges with the data from: http://vlado.fmf.uni-lj.si/pub/networks/data/ucinet/zachary.dat

- Have the transition from before/after the split occur on a mouse click.

1.1 SOLUTION

I started the solution by chosing the force directed graph, because it shows the"bond strength" by the distance between the nodes . [2] Using the example, I learned the json format required to create the graph from the karate club data. Then I modified my code from the previous assignment, so the node sizes were computed from the node centrality betweenness. The square root of the centrality was taken, but if it was less than five, then five was used for the size so every node could be seen on the graph. The link weights were taken from the edge weights in the data. [1] The resulting diction was written into json files, one for the entire club and one for the split. [4] Listing 1, in section 1.3 is the python code used calculate the betweenness and write the json files for the graphs.

1.2 RESULT

When the page is opened, the entire karate club graph is created from the kc.json data. There are two buttons. When Split-2 is clicked, the club graph splits in two, in according to the data in split.json. All data files used are included in this report. If the "Karate Club" button is pressed, the graph goes back to one. The original script created multiple instances of the graphs, so the remove() method was used to prevent that behavior. [3] Hovering over a node highlights it in red and shows the node id. The resulting page has also been placed at http://www.cs.odu.edu/ smelton/webscience/karateclub.html.

1.3 PYTHON CODE

```python
1  import connected
2  import networkx as nx
3  import betweenness
4  import csv
5  import argparse
6  import json
7  import math
8  from networkx.readwrite import json_graph
9
10 def buildClubGraphs(inFile, n):
11   club = []
12   #weighted edge data retrieved from http://vlado.fmf.uni-lj.si/pub/networks/data/ucinet/zachary.dat
13   #open the csv and put the edge matrix into a list of lists
14 with open(inFile) as i:
15   reader = csv.reader(i)
16   for row in reader:
17     club.append(row)
18   #create a list of tuples, with edges and weights
19   eList = []
20 for i in range(0,len(club[0])):
21   for j in range(0, len(club[0])):
22     if int(club[i][j]) > 0:
23       eList.append((i+1,j+1,int(club[i][j])))
24   #Use list of edges to create the nx graph
25   K = nx.Graph()
26   K.add_weighted_edges_from(eList)
27   # Create the json file for the whole karate club
28   d = json_graph.node_link_data(K)
29   print 'Original graph has ' + str(K.number_of_edges()) + ' edges\n'
30   graphs = list(nx.connected_component_subgraphs(K))
31   removed = []
32   c = nx.betweenness_centrality(K, weight='weight', normalized=False)
33 for i in range(0, len(c)):
34   if c[i+1] > 5:
35     size = math.sqrt(c[i+1])
36   if size >=5:
37     d['nodes'][i]['size'] = size
38   else:
39     d['nodes'][i]['size'] = 5
40
41 while len(graphs) < n:
42   b = nx.edge_betweenness_centrality(K, weight='weight', normalized=False)
43   e = (0, 0)
44   centrality = 0.0
45   for i in b:
46     if b[i] > centrality:
47       centrality = b[i]
48       e = i
49     # check
50     #b.get(item)
51   edges = [e] #put returned tuple into a list, to be used to remove from graph
52   #keep track of removed edges and their weighted edge betweenness centrality
53   removed.append({'edge': e, 'cent': centrality})
54   K.remove_edges_from(edges)
55   graphs = list(nx.connected_component_subgraphs(K))
56
57 cc = nx.betweenness_centrality(K, weight='weight', normalized=False)
58
59 print 'Removed edges to build '+str(n)+' subgraphs\n'+'\n'.join(str(r) for r in removed)+'\n'
60 for i in range(0,len(graphs)):
61   print 'Nodes in graph ' + str(i+1) + ': ' + ','.join(str(g) for g in sorted(graphs[i].nodes()))
62   #
63   dd = json_graph.node_link_data(K) # node-link format to serialize
64   # write json to file
65 for i in range(0, len(cc)):
66   if cc[i+1] > 5:
67     size = math.sqrt(cc[i+1])
68   if size >=5:
69     dd['nodes'][i]['size'] = size
70   else:
71     dd['nodes'][i]['size'] = 5
```

```
72
73 for i in range(0, len(c)):
74   if graphs[0].has_node(dd['nodes'][i]['id']):
75     dd['nodes'][i]['group'] = 1
76   else:
77     dd['nodes'][i]['group'] = 2
78
79
80 json.dump(d, open('kc.json','w'))
81 json.dump(dd, open('split.json','w'))
82
83
84 if __name__ == '__main__':
85   parser = argparse.ArgumentParser(description='Build club graph and split into communities')
86   parser.add_argument('file', help='input edge matrix file, csv format')
87   parser.add_argument('communities', type=int, help='integer, number of communities to split into')
88   args = parser.parse_args()
89   f = args.file
90   n = args.communities
91   buildClubGraphs(f, n)
```

Listing 1: Python graph code

2 TWITTER

2.1 RESULT

REFERENCES

[1] Zachary karate club network dataset – KONECT, October 2014. Accessed: 2014-10-22.

[2] Mike Bostock. Force-directed graph. `http://bl.ocks.org/mbostock/4062045`. Accessed: 2014-11-1.

[3] Sam. How can I remove or replace SVG content? `http://stackoverflow.com/questions/10784018/how-can-i-remove-or-replace-svg-content`. Accessed: 2014-11-1.

[4] tutorialspoint. Python dictionary. `http://www.tutorialspoint.com/python/python_dictionary.htm`. Accessed: 2014-11-1.