

BT5240: Computational Systems Biology

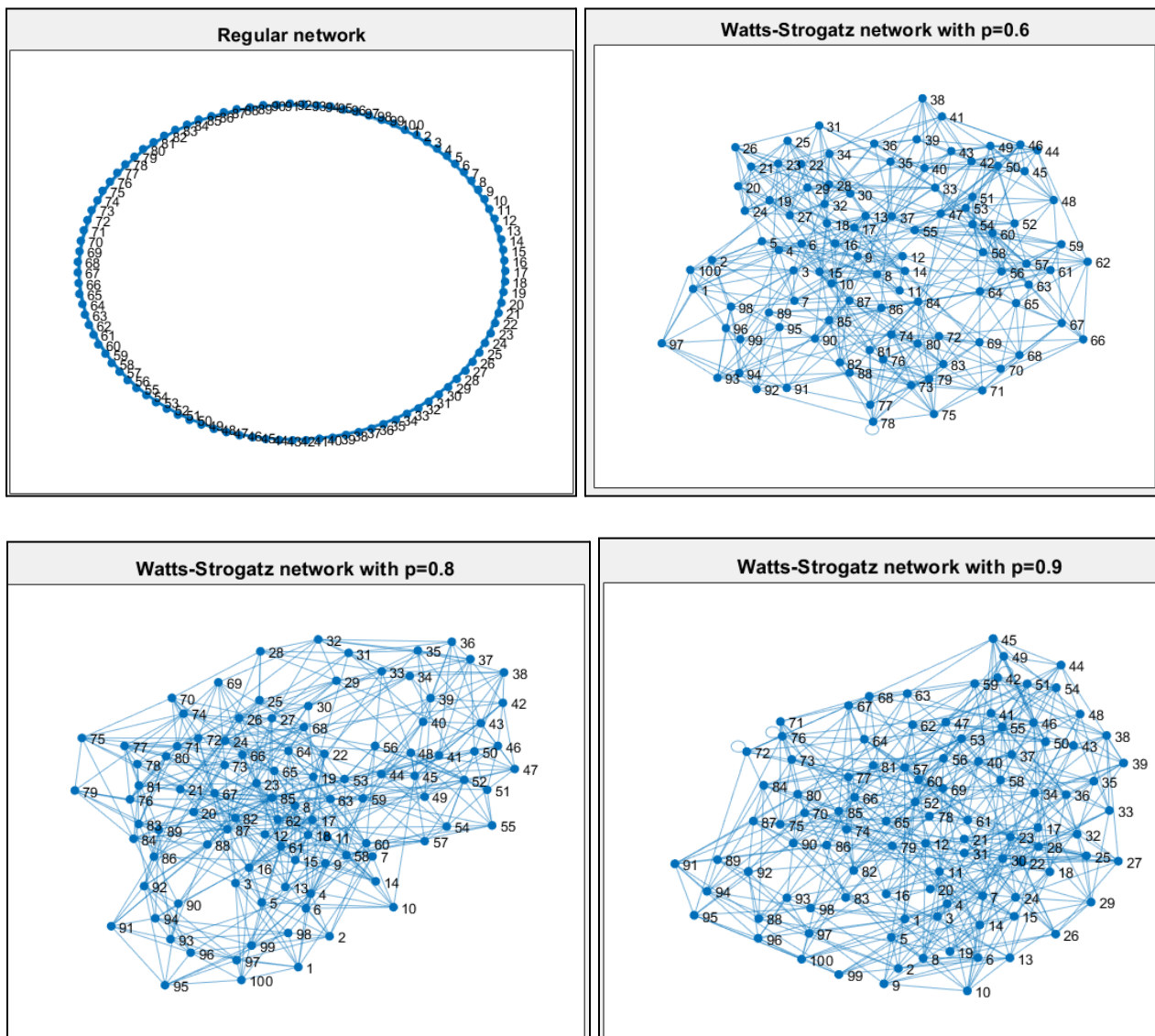
Assignment 2 - Report

By :
Shreya Rajagopalan
BE21B038

Problem 1:

Consider a Watts–Strogatz network with 100 nodes and each node having 10 neighbours. Rewire the network with probabilities (p) = 0.6, 0.8 and 0.9, and you will see that these networks change from the parent network. Calculate the average clustering coefficient and the characteristic path length of the rewired networks. Can these parameters be obtained in any random network of the same nodes and edges? Hint: Generate 100 Erdos–Renyi random networks with identical numbers of nodes and edges. Compare the average clustering coefficient and characteristic path lengths of random graphs and the rewired network.

Solution:



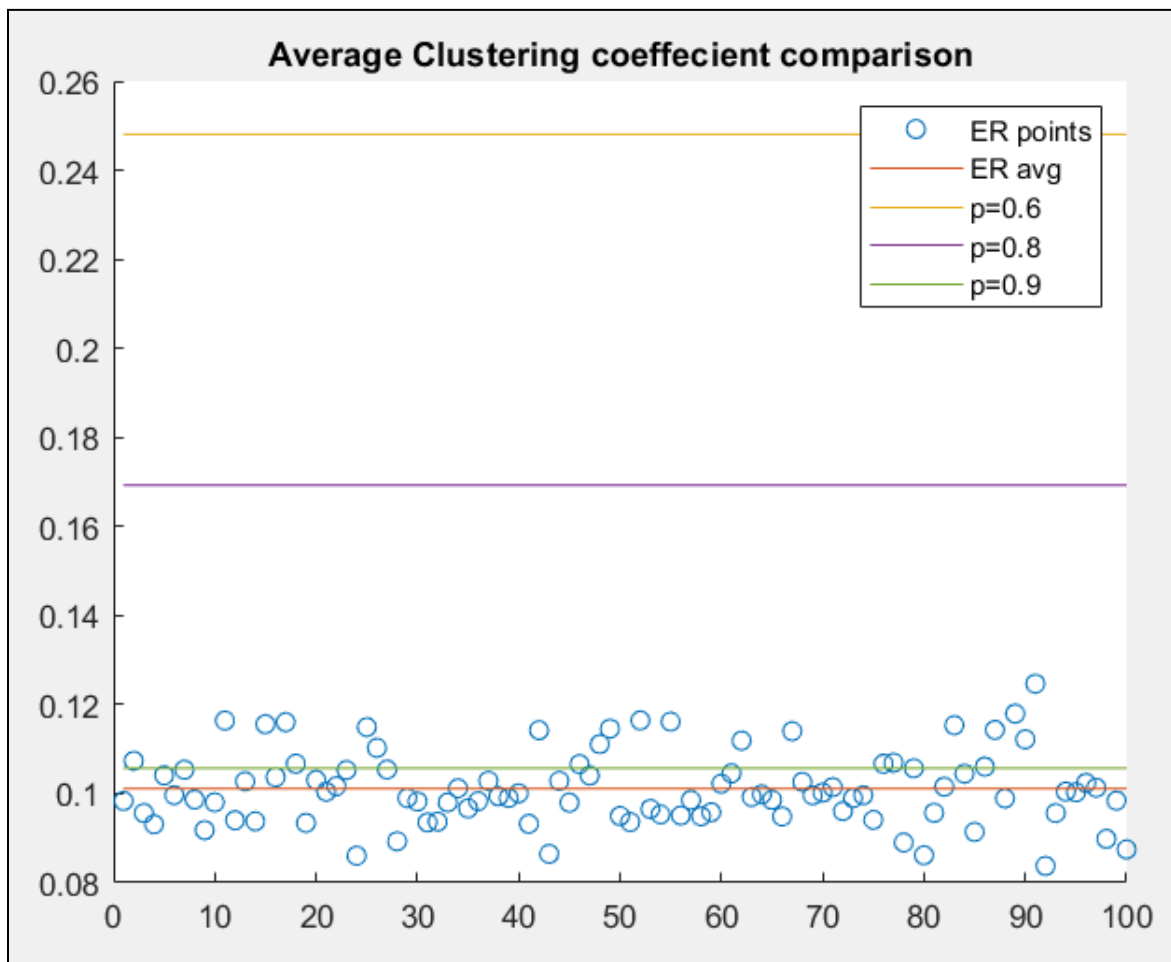
In one iteration, these values were observed

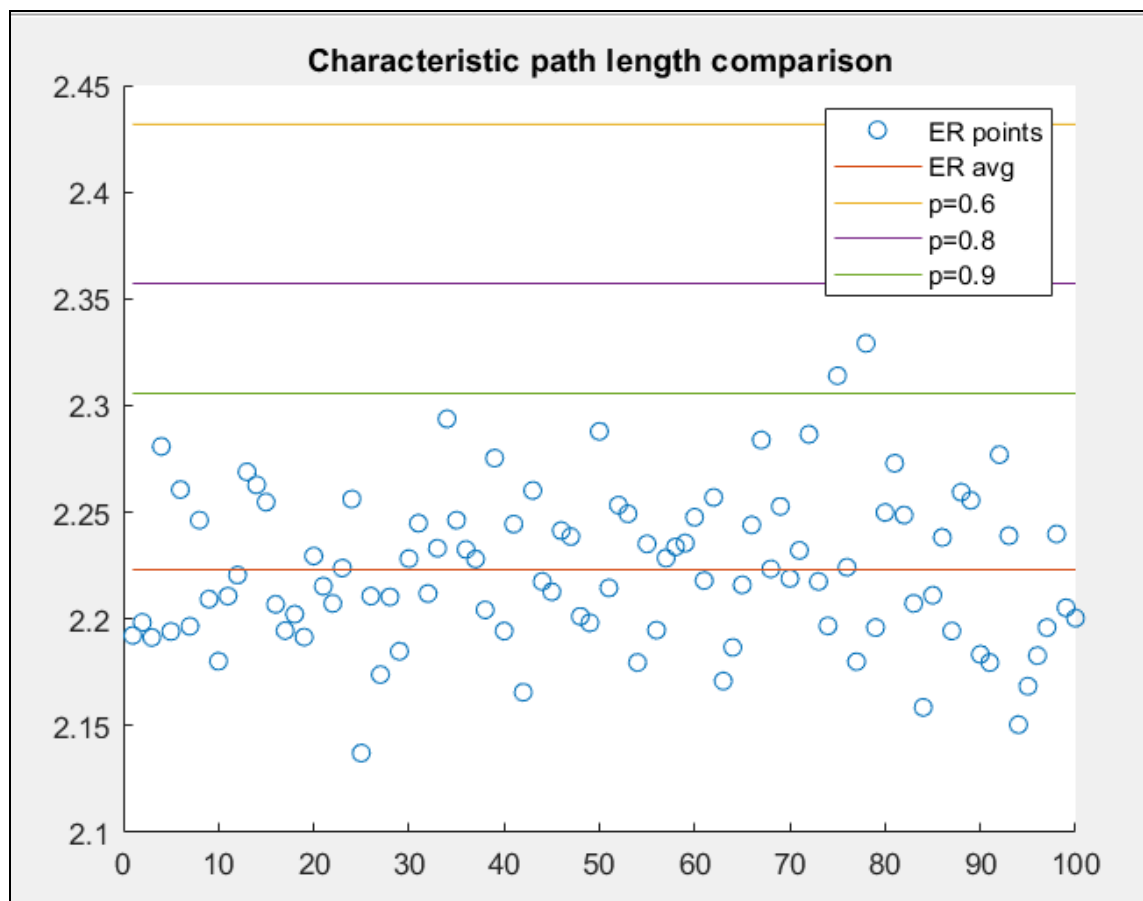
- a. $p = 0.6$:
 - i. Average clustering coefficient = 0.2481
 - ii. Characteristic path length = 2.4317
- b. $p = 0.8$:
 - i. Average clustering coefficient = 0.1693
 - ii. Characteristic path length = 2.3572
- c. $p = 0.9$:
 - i. Average clustering coefficient = 0.1057
 - ii. Characteristic path length = 2.3057

Yes, these parameters can be obtained in any random network of the same nodes and edges.

In the code, the matrix ER has the average clustering coefficient and characteristic path length values for 100 ER networks in columns 1 and 2 respectively.

In comparison to the Erdos-Reyni random networks





The general trend observed is that the clustering coefficient and characteristic path length approach that of random networks as the rewiring probability is taken closer to 1.

Problem 2:

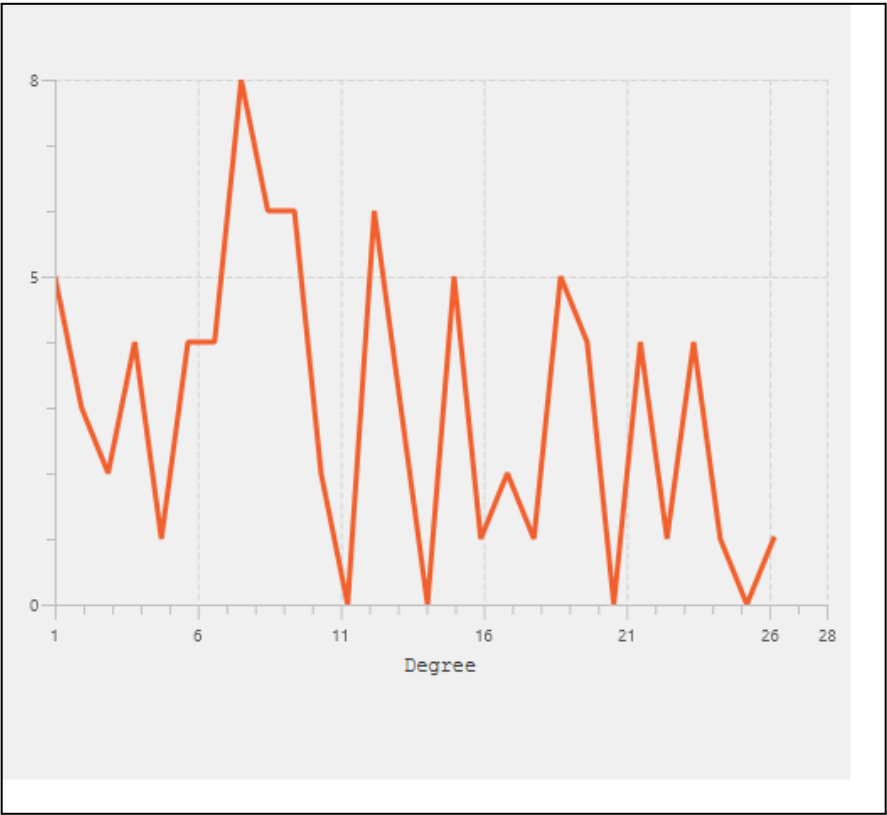
Exploring microbial interactions is necessary to understand microbial community structure, regulation, and their maintenance in the built environments. Here we provided three microbial networks obtained from three different built-environments of interest: a. Hospital b. Office and c. Metro stations. The networks are provided as the edge list, along with the normalised correlation as edge weight. Use Cytoscape to calculate the following statistics for the individual networks, compare and comment on your observations

1. Plot of the degree distribution
2. Plot of the average clustering coefficient as a function of the node degree
3. The following statistics
 - (a) Average node degree
 - (b) Average local clustering coefficient
 - (c) Characteristic path length
4. Perform differential network analysis to show the unique interactions in each environment and how hub nodes affect that
5. Comment on any observations due to biased network size.

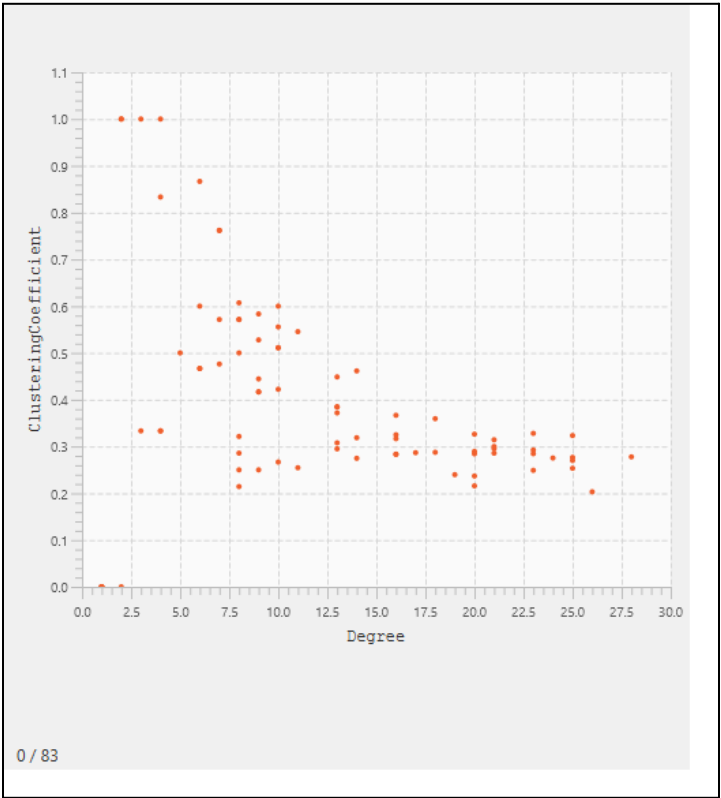
Solution:

a. Hospital

1.Degree Distribution



2. Average clustering coefficient vs Degree



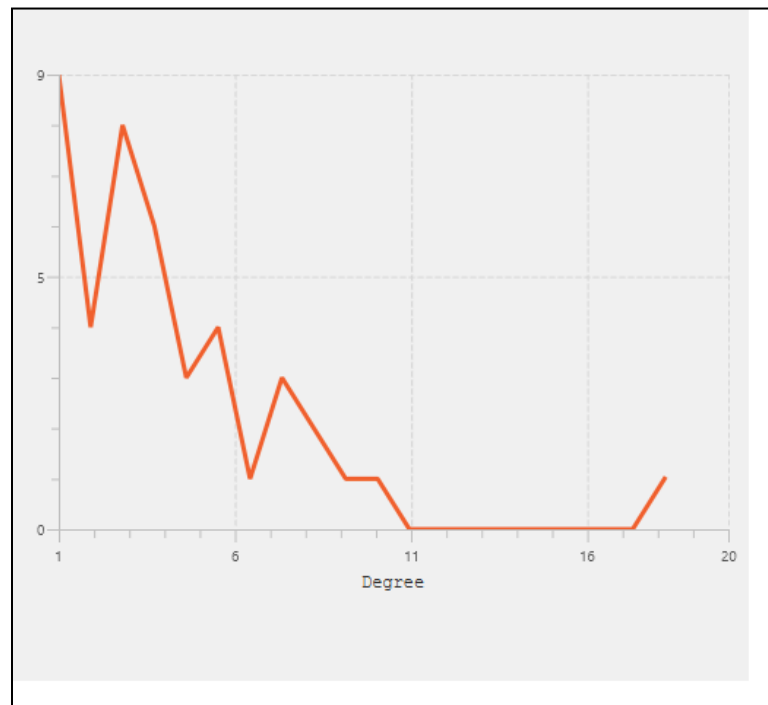
3. Statistics

Hospital (undirected)	
Summary Statistics	
Number of nodes	83
Number of edges	508
Avg. number of neighbors	12.241
Network diameter	4
Network radius	3
Characteristic path length	2.147
Clustering coefficient	0.392
Network density	0.149
Network heterogeneity	0.599
Network centralization	0.197
Connected components	1
Analysis time (sec)	0.114

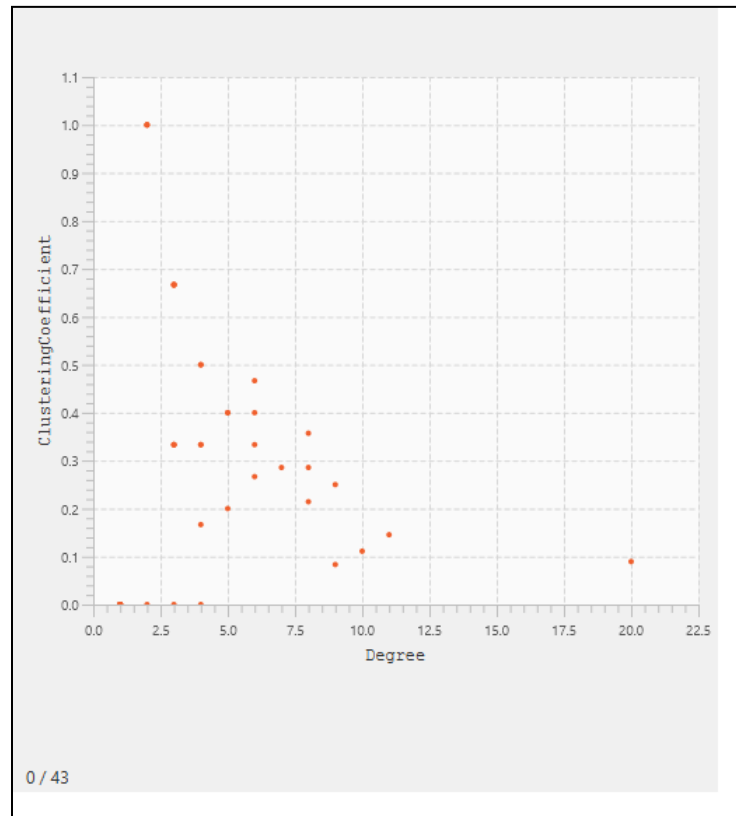
- a. 12.241
- b. 0.392
- c. 2.147

b. Metro

1. Degree distribution



2. Average clustering coefficient vs degree



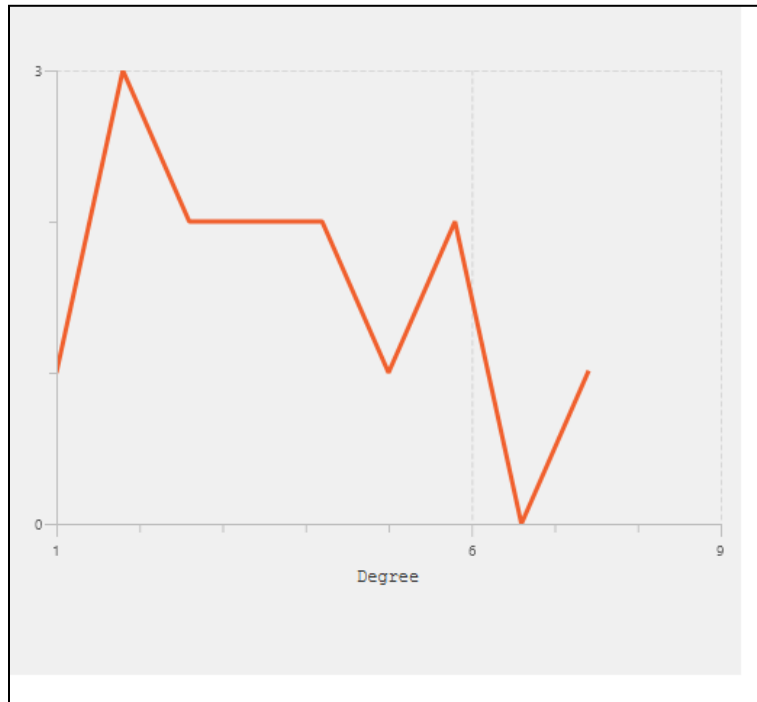
3. Statistics

Metro (undirected)	
Summary Statistics	
Number of nodes	43
Number of edges	97
Avg. number of neighbors	4.512
Network diameter	5
Network radius	3
Characteristic path length	2.555
Clustering coefficient	0.297
Network density	0.107
Network heterogeneity	0.800
Network centralization	0.387
Connected components	1
Analysis time (sec)	0.005

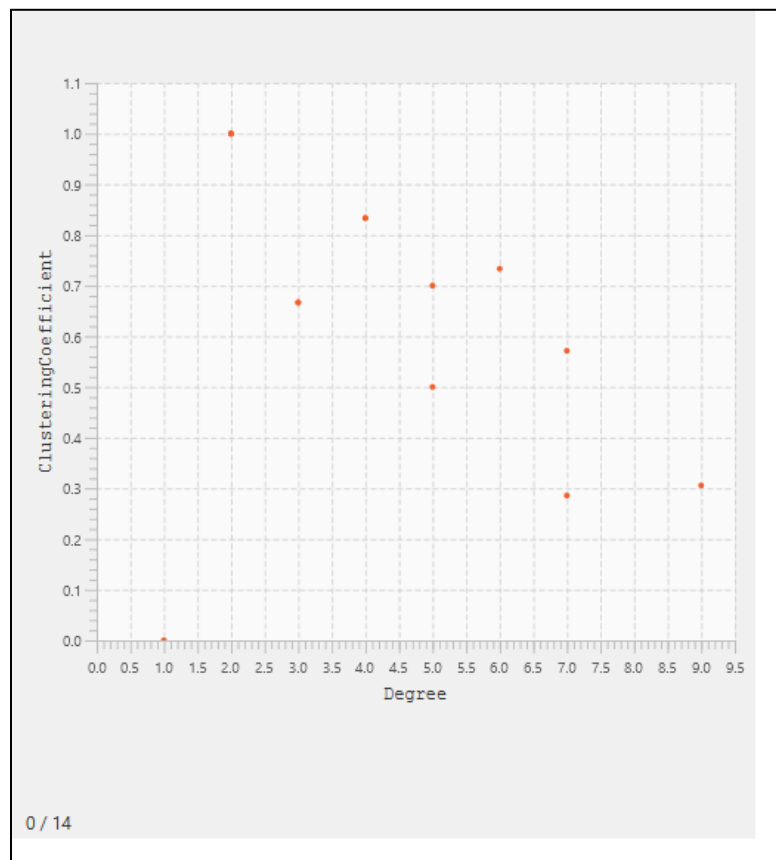
- a. 4.512
- b. 0.297
- c. 2.555

c. Office

1. Degree distribution



2. Average clustering coefficient vs degree

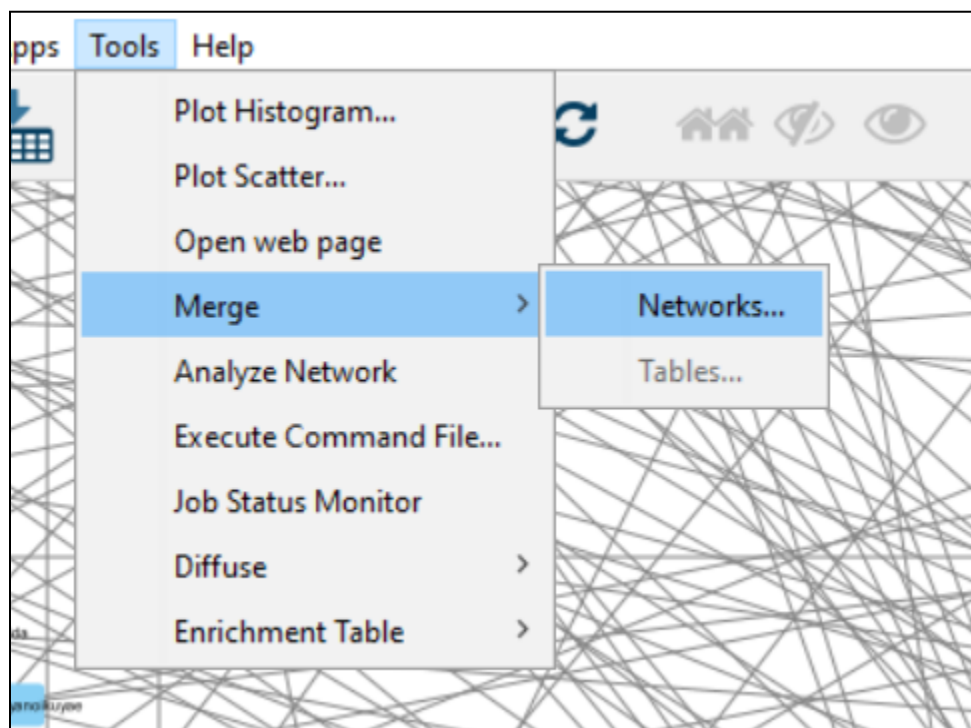


3. Statistics

Office (undirected)	
Summary Statistics	
Number of nodes	14
Number of edges	30
Avg. number of neighbors	4.286
Network diameter	4
Network radius	2
Characteristic path length	1.923
Clustering coefficient	0.650
Network density	0.330
Network heterogeneity	0.525
Network centralization	0.423
Connected components	1
Analysis time (sec)	0.001

- a. 4.286
- b. 0.650
- c. 1.923

4. Differential Network Analysis



Advanced Network Merge

Union Intersection **Difference**

☒ Only remove nodes if all their edges are being subtracted, too
☐ Remove all nodes that are in the 2nd network

Available Networks:

- Metro-Office
- Office unique**
- Office-Hospital
- Reference Network Ideker 2011

Networks to Merge:

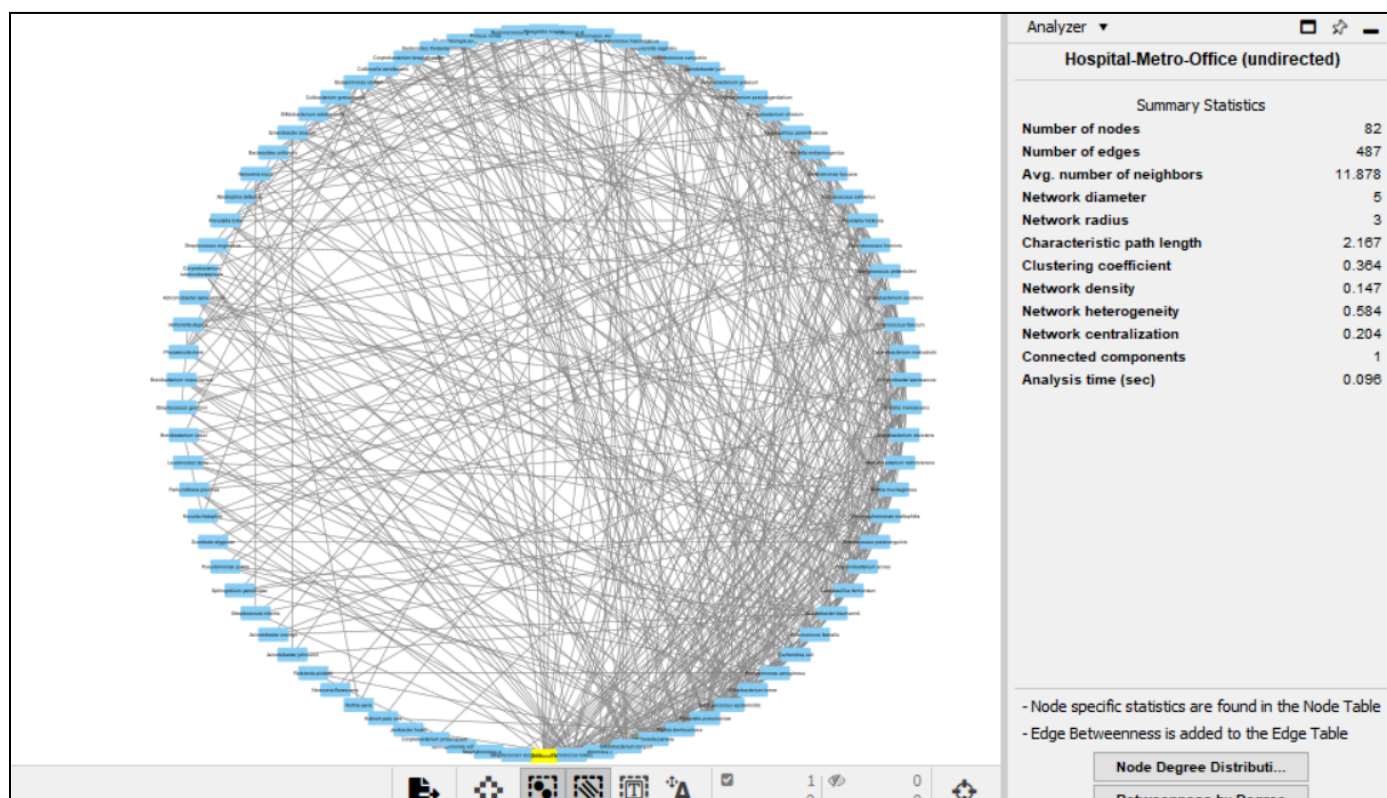
- Hospital-Metro
- Office

> Advanced Options

Merge Cancel

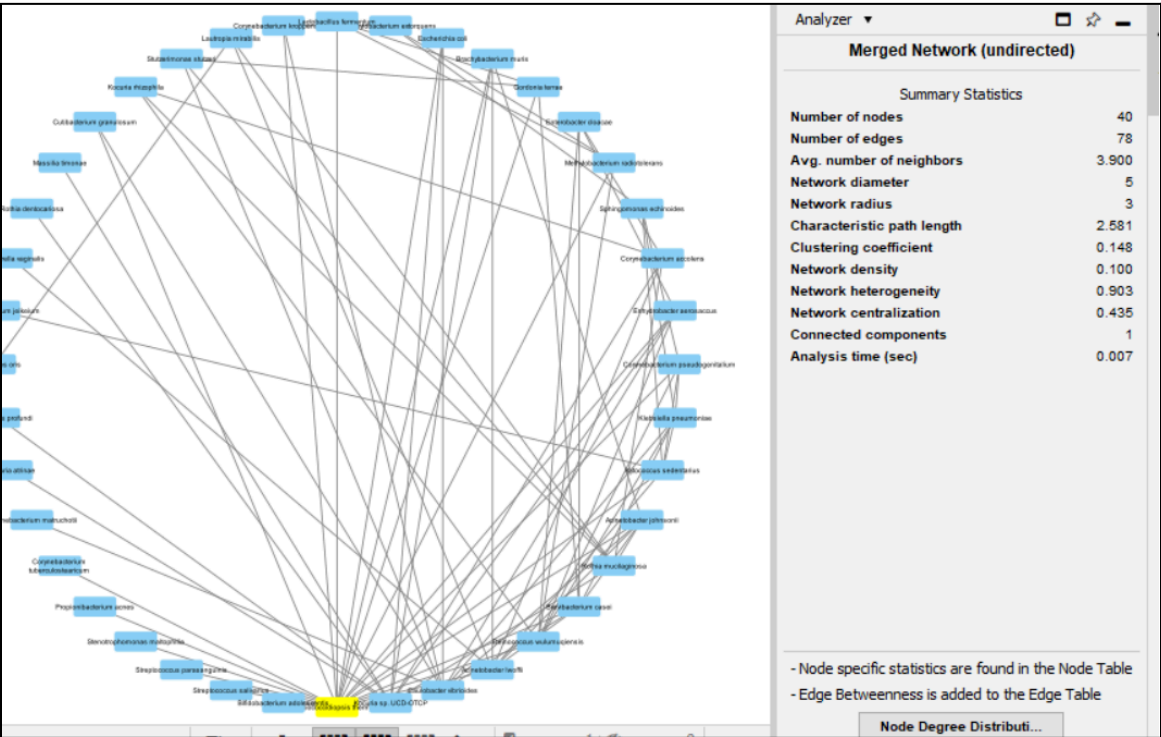
a. Hospital Unique interactions

I.e., Hospital - Metro - Office



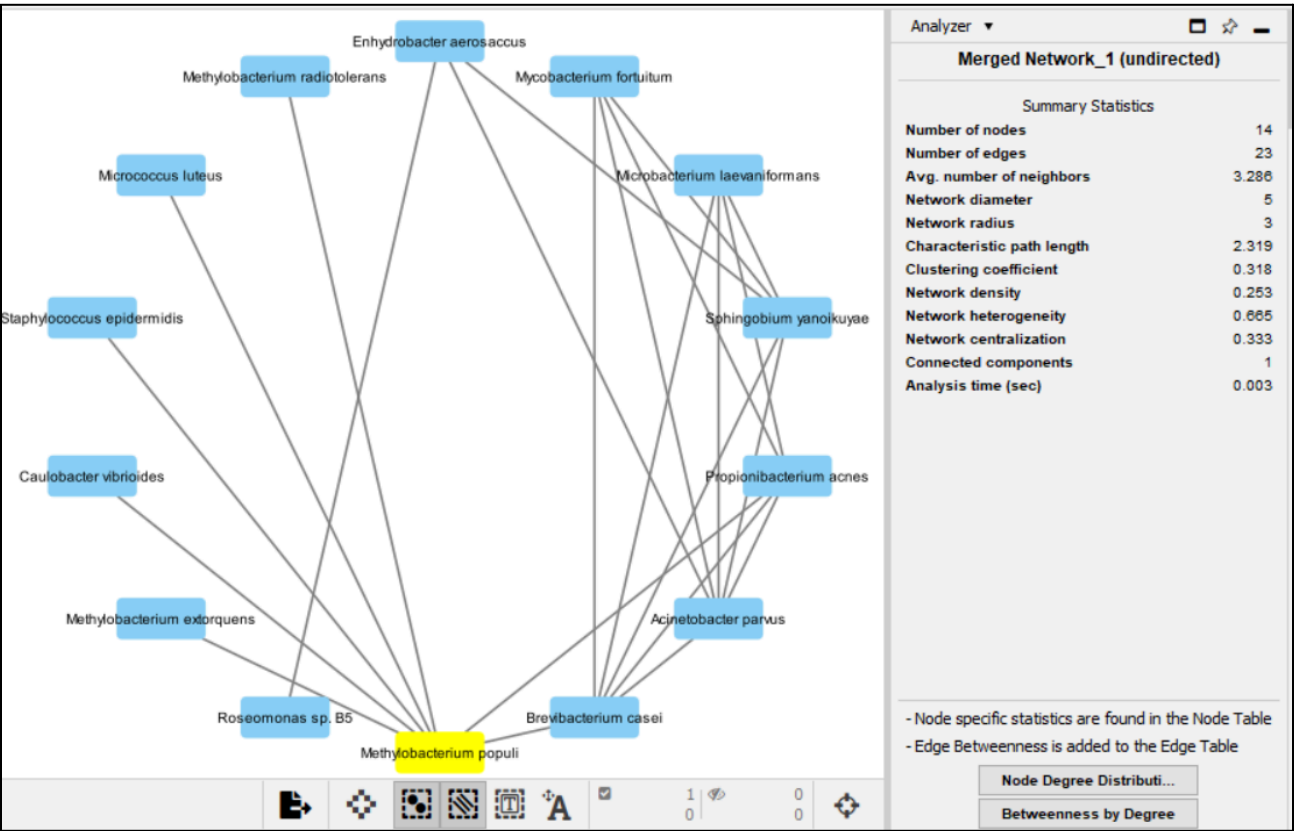
b. Metro Unique interactions

I.e., Metro - Office - Hospital



c. Office Unique interactions

I.e., Office - Hospital - Metro



The nodes with the highest degrees (hubs) have been highlighted in these unique interaction networks. In this case, *Propionibacterium acnes* was a hub with the highest degree in the Metro network but it is not a unique interaction as its degree is not preserved in the differentially analysed network. This gives us the idea that hubs usually have many interactions possibly spanning more than one network. Thus, in a case of targeted attack, these nodes must be attacked first.

5. Network size

The hospital microbial interaction (H) is a large network, followed by the Metro microbial interaction (M) and Office microbial interaction (O) with respect to the number of nodes and edges. The effects of network size can be clearly observed from the network statistics. Average degree of H is much higher than that of O and M, whereas average clustering coefficient and characteristic path lengths don't vary too much.

Due to the large size of H, when differential analysis is performed, it is observed that most of the interactions are preserved, and the average degree is largely unaffected. For the smaller networks, even removing a few interactions can cause the average degree to drop significantly.