Daphne Basangwa
1.041 Mini Project
Spring 2013
May 5, 2013

Analysis of Jazz Band Network

Introduction

My chosen network is a network of Jazz Bands that recorded between the years of 1912 and 1940. The nodes are bands and the edges connect two bands if they have shared at least one musician in their recordings. The network comprises of 198 bands and 2742 links.

Findings of Previous Work

I referred to "Community Structures in Jazz" by Pablo M. Glazer and Leon Danon for research. They studied the community structure of the same network that I study in this paper and their analysis of the network using the Girvan and Newman algorithm yielded four distinct communities based on geographic location (recording city) and race of the musicians. In Gephi, using the modularity algorithm, I also find four distinct communities. Because my network data was limited (no information on node identification), I was unable to verify that the communities obtained were similar to the ones obtained by Glazer and Danon. However, I have taken my findings as representative of the same communities as theirs. Figure 1 shows the community structure of the network showing the four communities obtained.

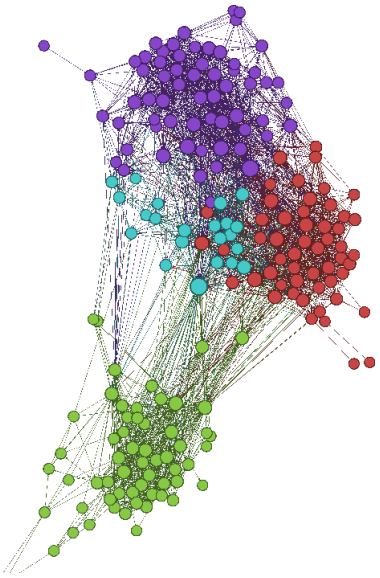

Figure 1: Community Structure of Jazz Band Network

I present below a quantitative analysis of the network. I studied the degree distribution, the nodal clustering coefficient, and the degree of a node's nearest neighbor (k_{nn}) and how that varies with the degree, k, of the node. I also compare the network to a scale-free network of the same size.

Assortativity - k_{nn} vs k

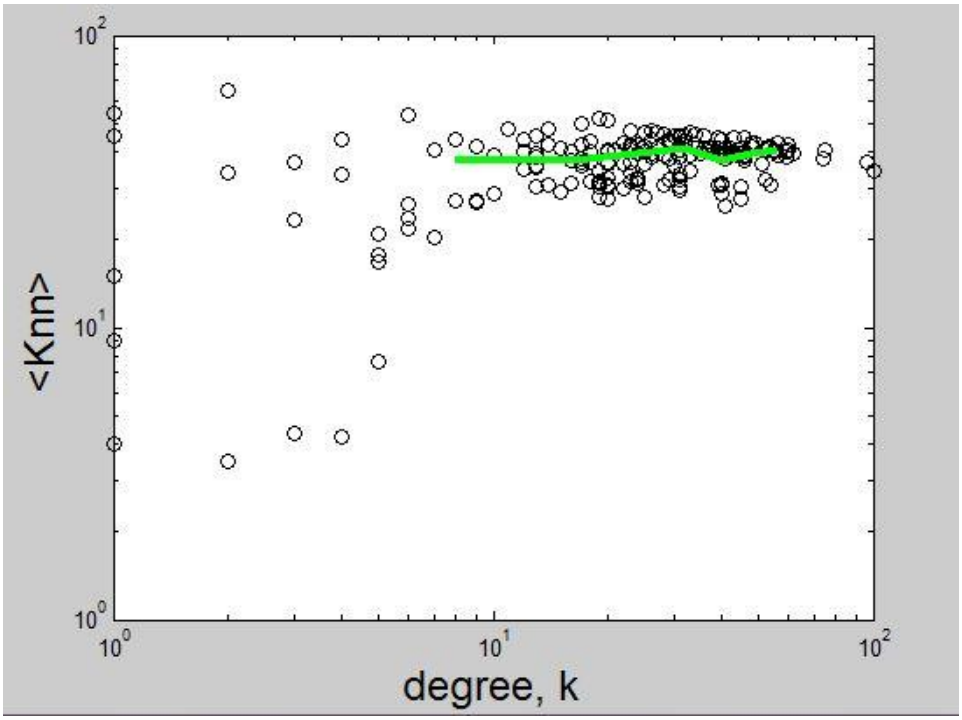

Figure 2A: k_{nn} Vs. k – Empirical Jazz Band Network

Because this band network is built on human acquaintanceship, one would expect it to possess similar quantities to a social network, such as "small world" characteristics or a certain degree of assortativity. Figure 2A shows a plot of the degree of the nearest neighbor of a node (k_{nn}) against its degree, k. The green line shows a binned average of the data. The graph does not indicate that the network is assortative. Instead, it shows k_{nn} fluctuating around an almost constant value. I propose that competition between popular bands hinders the collaboration of their musicians, thereby creating a network that is not assortative. Also, less popular bands would be interested in associations with musicians from more popular bands for the purpose of promoting their own popularity.

A comparison of the plot with that of a generated scale free network of the same size (Figure 2B) shows some similarities: both networks have more nodes of lower degree and fewer hubs (nodes of higher degree). However, the empirical jazz band network shows a lot more variability in the node degree.

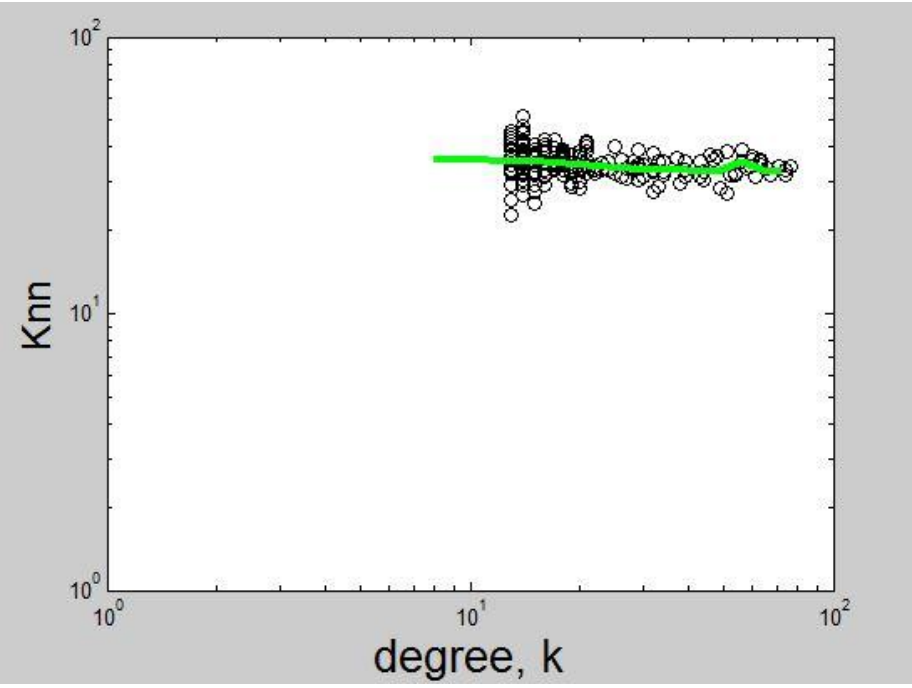

Figure 2B: k_{nn} Vs. k – Scale-Free Network

Clustering Coefficient

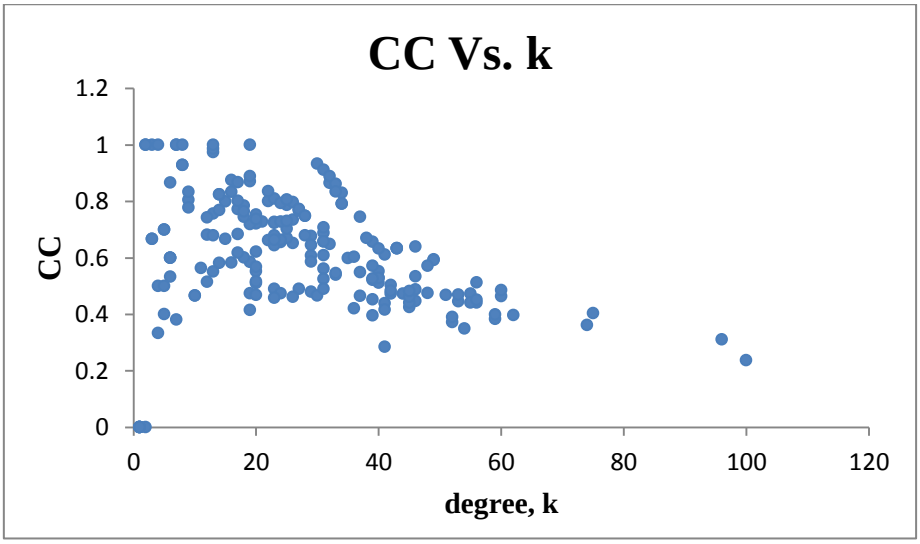

Figure 3: Clustering Coefficient Vs. Degree

Figure 3 shows a plot of the clustering coefficient of a node against the degree of the node for the empirical jazz band network. The network shows negative correlation between clustering coefficient and node degree i.e. the clustering coefficient decreases as the node degree increases. This is similar to the relationship of clustering coefficient and node degree seen in a Barabasi-Albert scale-free network of the same size. Scale free networks tend to have hubs (highly linked nodes) and hubs tend to have lower clustering coefficients compared to other

nodes. This is because the high connectedness of hubs lowers the likelihood of linkages existing between their neighbors. Consider the equation for the local clustering coefficient of a node:

$$C = \frac{2L}{k(k-1)}$$

where L is the number of links that exist between neighbors of the node and k is the degree of the node. For hubs, the denominator is large and tends to lower the value of the clustering coefficient. For less connected nodes, the likelihood of a linkage between neighbors tends to be higher because the degree k of the node is lower.

Other Physical Properties

Figure 4 below shows a log-log plot of the degree distribution of the empirical data on the left and the same plot for a scale-free network of the same size on the right. The red lines indicate binned averages of the data points. I found that the empirical network shows power-law properties with an exponent of about -0.5.

The network shows a wide range in degree distribution, much wider than the degree distribution given by a Random graph and Small world network of the same size. In the jazz band network, the most connected node has a degree (k_{max}) of 100, which is greater than half the size of the network. This shows a strong deviation from the small world networks, which do not tend to possess hubs.
The average degree k_{av} of the network is comparable to the average degree given by a scale free network of the same number of nodes. The Maximum degree k_{max} is higher, as shown in table 1 below.

Network	Kmax	Kav	Nodes
Jazz Bands	100	27.6	198
Scale Free	76	25.1	198

Table 1: Comparison between Jazz Network and Scale-Free Network

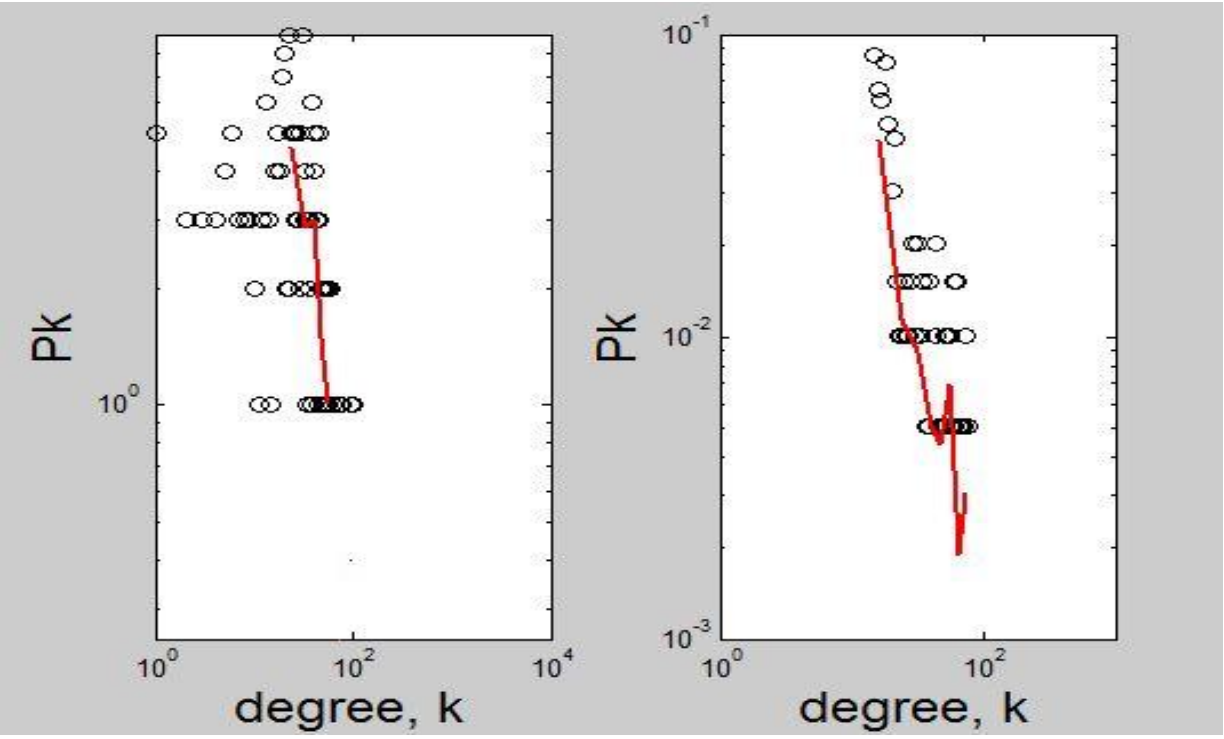

Figure 4: Degree Distributions

Future Work and Conclusion

I have found that the jazz band network follows a power law distribution with an exponent of about -0.5. In the future, it would be interesting to have more detailed information of the nodes such as band names or weights of the connections (based on how many common musicians the bands shared) to enable a more detailed analysis of the network.

References

A. Arenas, L. Danon, A. D´ıaz-Guilera, P.M. Gleiser, & R. Guimer`a. *Community analysis in social networks.* The European Physics Journal B.

P. M. Gleiser & Leon Danon. *Community Strucure in Jazz.* World Scientific Publishing Company. October 2003.