CS-E5740 Complex Networks,
Answers to exercise set 7

Alex Herrero Pons, Student number: 918697

November 9, 2020

Problem 1

a) We can see all three distributions in the same plot using loglog-scal in Figure 1.
A we can easily observe all of them have heavier tails than the Gaussian distribution.

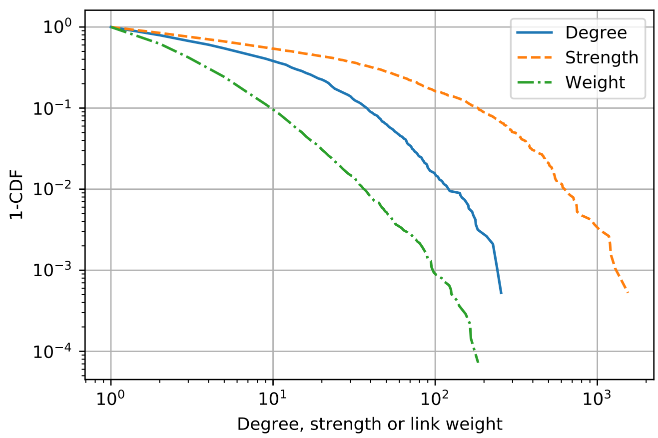

Figure 1: 1-CDF of the node degree k, node strength s and link weight w distributions.

From the figure we can also extract some estimations for the 90th percentiles when 1-CDF = 0.1. The estimations are shown in Table 1.

Distribution	P_{90}
Degree	50
Strength	200
Weight	10

Table 1: Estimations of the 90th percentiles (P_{90}) of the different distributions.

b) We can observe the average link weight $\langle W \rangle = \frac{s}{k}$ per node for the data and the respectively bin-average versions in Figures 2 and 3.

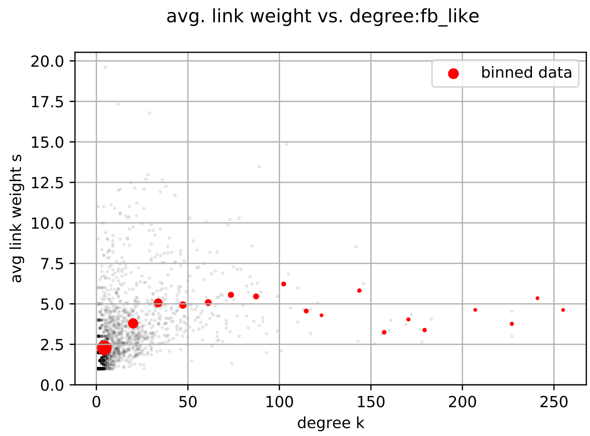

Figure 2: All the data points of $\langle w \rangle$ as a function of k with linear x-axes with the bin-averaged version.

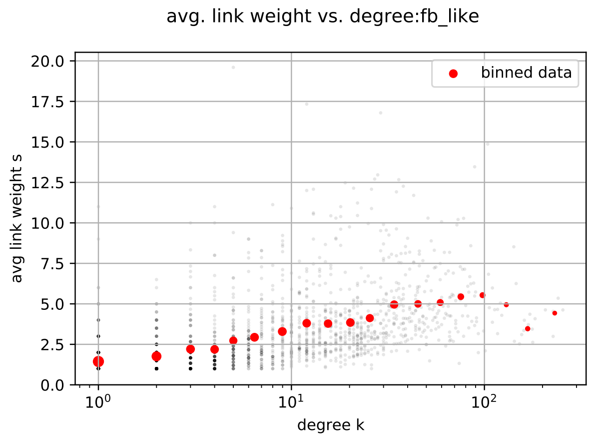

Figure 3: All the data points of $\langle w \rangle$ as a function of k with logarithmic x-axes with the bin-averaged version.

We can say that the logarithmic approach (Figure 3) suits better for presenting $\langle W \rangle$ as a function of k because there are some observable clusters in the different values of the degree and the binned data fits perfectly with this clusters.

c) The overlap for every link as a function of link weight created in this exercise is shown in Figure 4. In the same figure it's also shown the bin-averaged version using 30 bins. Observing the bins we can see a subtle linear trend that spreads with high weights.

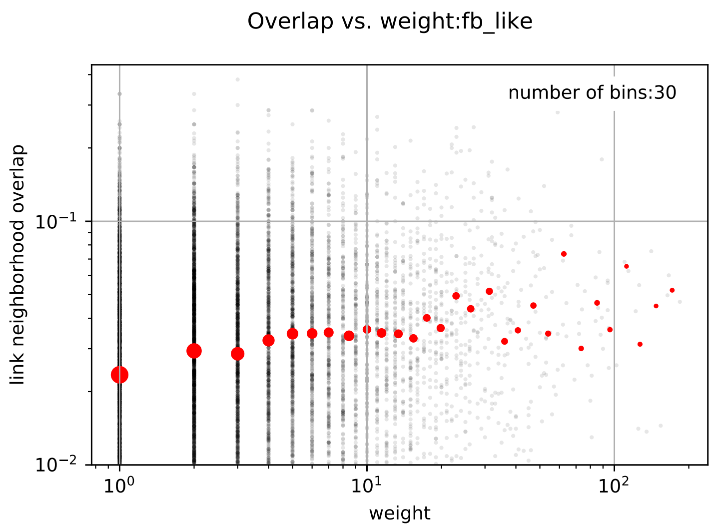

Figure 4: Overlap for every link as a function of link weight.

Problem 2

a) The results of the computed properties of the network are:

- Number of nodes $N = 279$.
- Number of links $L = 2088$.
- Network density $D = 0.05384079832907867$.
- Network diameter $d = 4$.
- Average clustering coefficient $C = 0.6465167472774311$.

b) The Figure 5 shows the network represented on top of the map of USA.

Figure 5: Visualization of the network with all links on top of the map of USA.

c) The visualization of the minimum and maximum spanning trees are shown in Figures 6 and 7 respectively.

Figure 6: Visualization of the minimum spanning tree of the network on top of the map of USA.

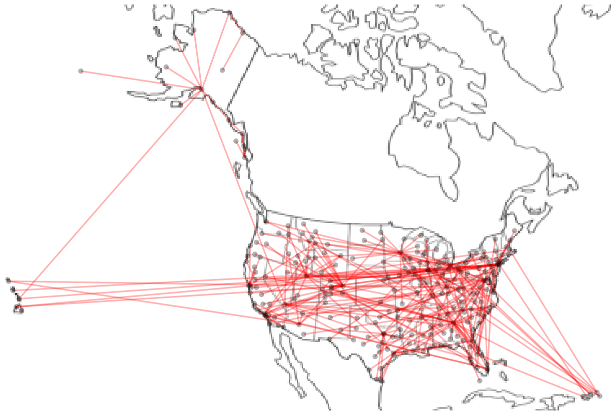

Figure 7: Visualization of the maximum spanning tree of the network on top of the map of USA.

To understand the overall organization of the air traffic in the US it's better to use the maximum spanning tree (Figure 7) because it shows the edges with more weight so it's representing a larger number of flights.

d) The thresholded network with only the $M = N - 1$ strongest links is represented in the Figure 8.

Figure 8: Visualization of the network with the M strongest links on top of the map of USA.

This thresholded network has 278 edges and 97 of them are in common with the maximum spanning tree. That makes a 34.89% of coincidence. With low coincidence and comparing figures 7 and 8 we can say that the simple thresholding does not yield a similar network as the maximal spanning tree.