Homework Network Science course - Part1

Cristina Gava

Matr 1153449

November 28, 2018

In this first part of the Homework, I focused on some of the main parameters presented in the course so far, in particular with respect to pdf estimation, assortativity value, clustering measures and ML fitting.

The dataset

The dataset selected was taken from [1]: the network is the transformation of the temporal graph (seen as a sequence of timestamped edges) into a static graph concerning reachability. In such way the new static network has edges representing metaphoric (phone calls or emails) or plain physical proxhimity between two entities. Hence this temporal reachbility graph represents a sequence of contacts obeying time and so the fact that a user could have transmitted a piece of generic information to another user.

First network metrics

First of all, the main metrics of the network have been extracted and are summed up in Table 1. As we can see from the table, the netowrk has a high variance and so a high spread value. This information can also be observed in Figure 1, where the degree pdf spreads significantly along the x-axis, resulting in a high variation in the degree values we can find inside the network. In the same way, also the skewness has a high value, resulting in an asymmetric degree distribution unbalanced towards the lower degrees (reult confirmed also by the low average degree).

Metric	Value
Average degree	29.59
Number of links	1862
Number of nodes	3594
Nodes degree variance	764.83
Nodes degree skewness	2651.44
Inhomogeneity ratio k	25.8477

Table 1: Network metrics concerning nodes and links number and the degree moments.

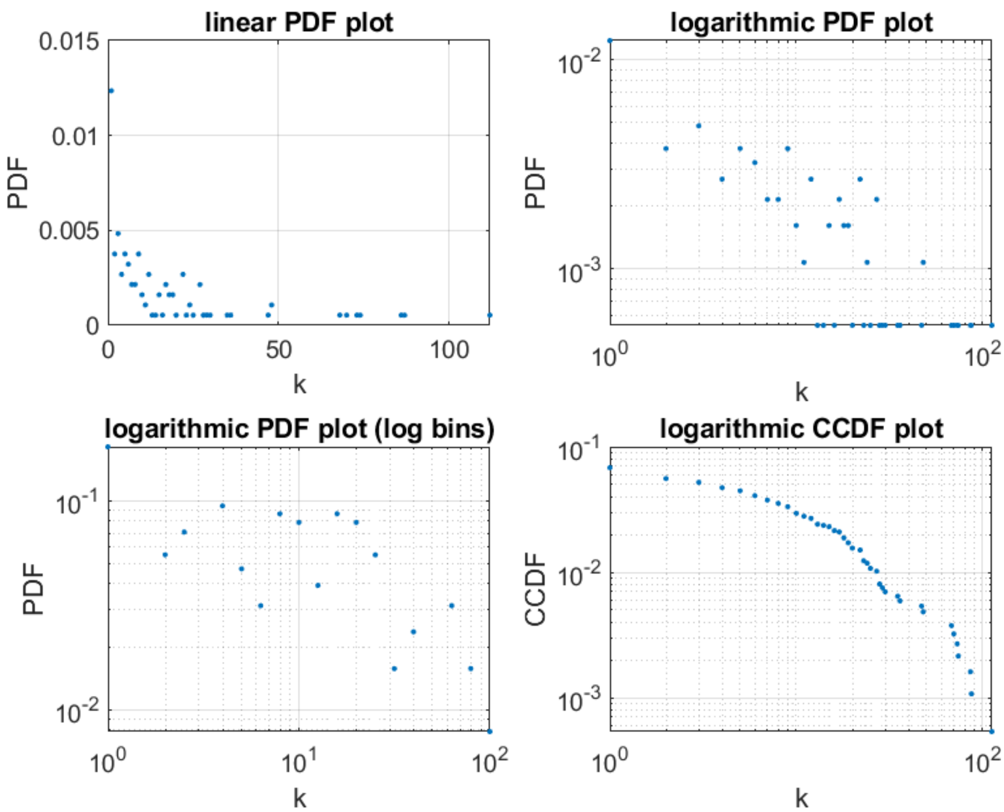

Figure 1: Plots of the various pdf functions: the normal one and the CCDF

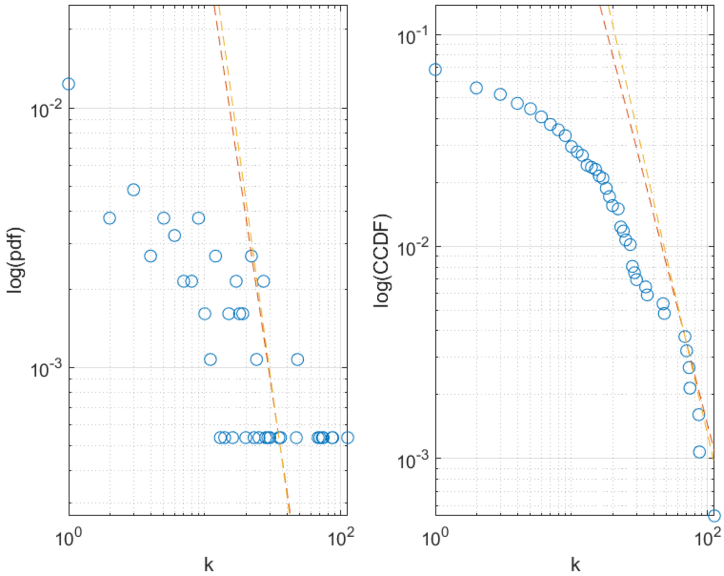

Figure 2: Maximul likelyhood estimation

Power law exponent estimate

For what concerns the power law exponent, Figure 3 and Figure 2 show the γ exponent estimation through ML fitting. The resulting value of γ is 3.7736 which does not go far from the classic hypothesis of $\gamma = 3.5$ and tells us that the network behaves as a common random scale-free network.

Clustering coefficient

Another interesting metric to look at is the clustering coefficient both for the general network and the connected-only component (Table 2): the first value shows how the general network does not have a strong tendence to cluster, but looking at the biggest connected component we see that there is a high clustering behaviour. The behaviour is confirmed by Figure 6 and Figure 5.

Assortativity values

The resulting assortativity average values is -0.46, which means that the network has disassortative properties. The result is confirmed also by the neutral behaviour that emerged in

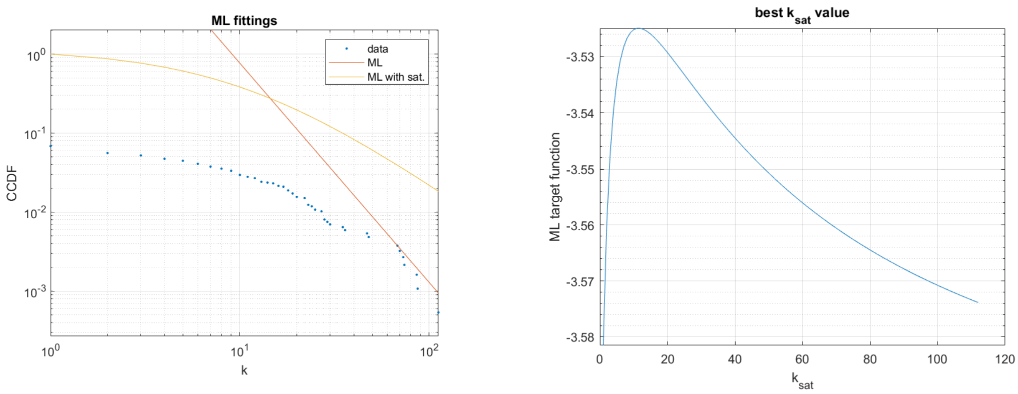

Figure 3: Maximum likelyhood fitting and target function estimation

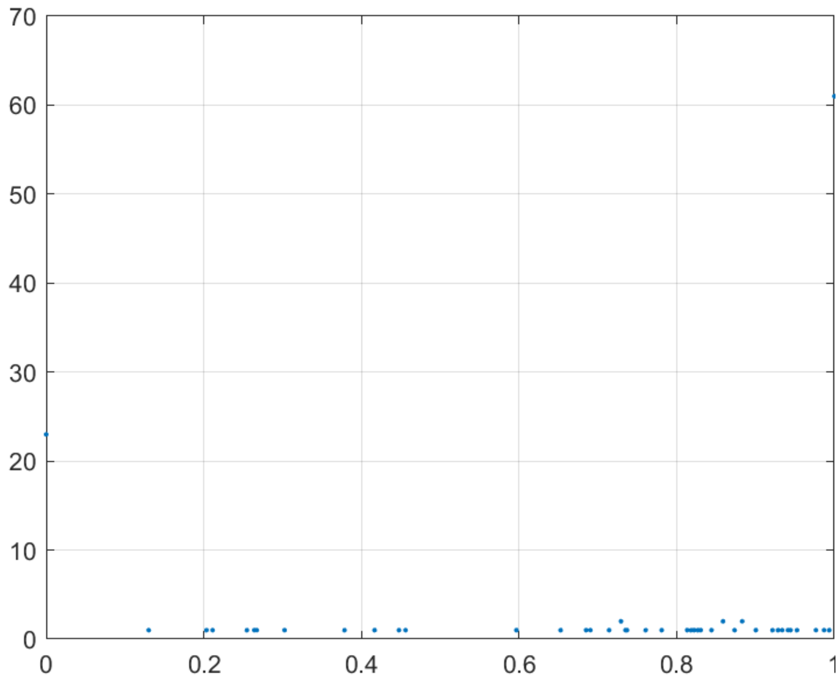

Figure 4: Clustering coefficient pdf

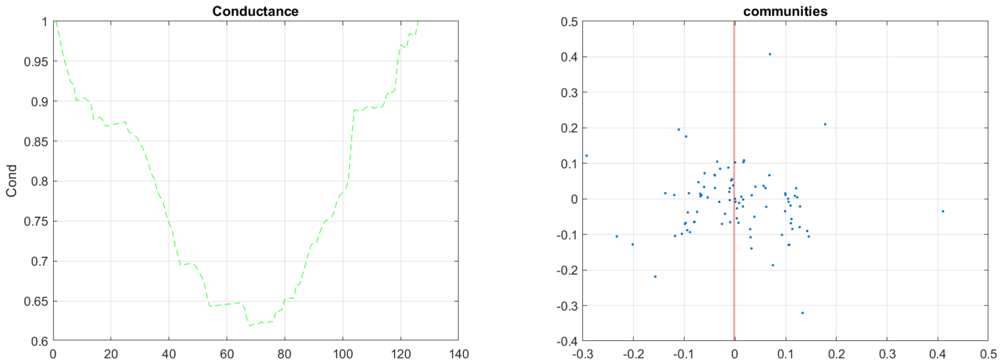

Figure 5: Communities separation and representation

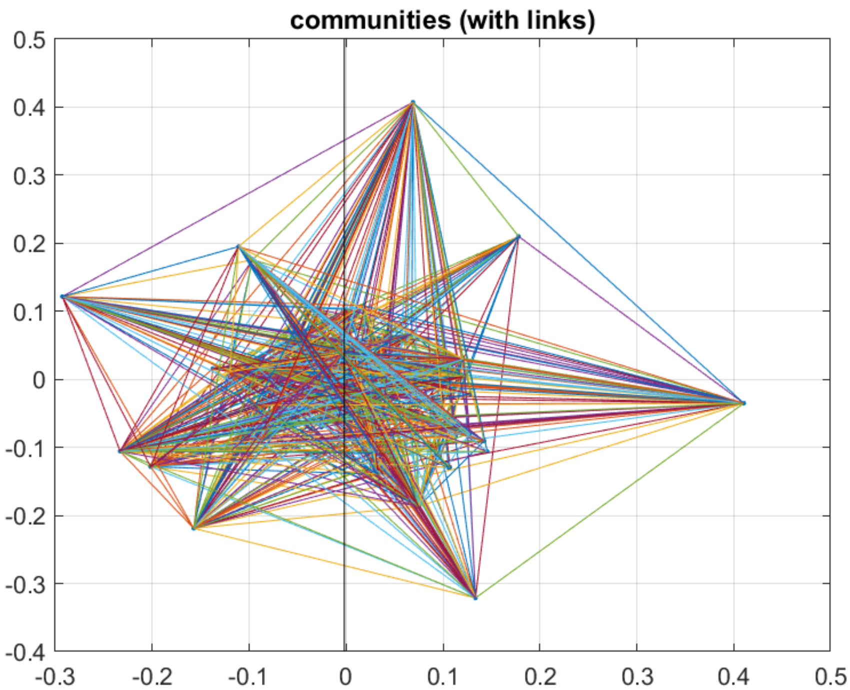

Figure 6: Network communities

Average clustering coefficient of the full component	Average clustering coefficient of the major connected component
0.7160	0.0253

Table 2: Network clustering coefficients

the previous section and Figure 7. To be specific, the four graphs are equal, since the original network is undirected and the output degrees are specular to the input degrees. The effects of it, can be also observed in Figure 6 where we can foresee how nodes tend to connect to other elements having a different degree, resulting in a sort of "uniform mesh effect". In this case the structural cutoff comes around the degree magnitude 6, which of course is smaller than the natural cutoff identifying the largest hub (19.14).

Network robustness

Finally, attention must be paid to the network robustness, which can be summarized by the values of the inhomogeneity ratio and the breaking point estimate. The former is equal to $25.8477 > 2$ and so we confirm the presence of a giant component; the latter is equal to 0.2264 which is rather low and signals a weak network. This conclusion is also confirmed by the low average degree of the nodes.

References

[1] R. A. Rossi and N. K. Ahmed, "The network data repository with interactive graph analytics and visualization," in *Proceedings of the Twenty-Ninth AAAI Conference on Artificial Intelligence*, 2015.

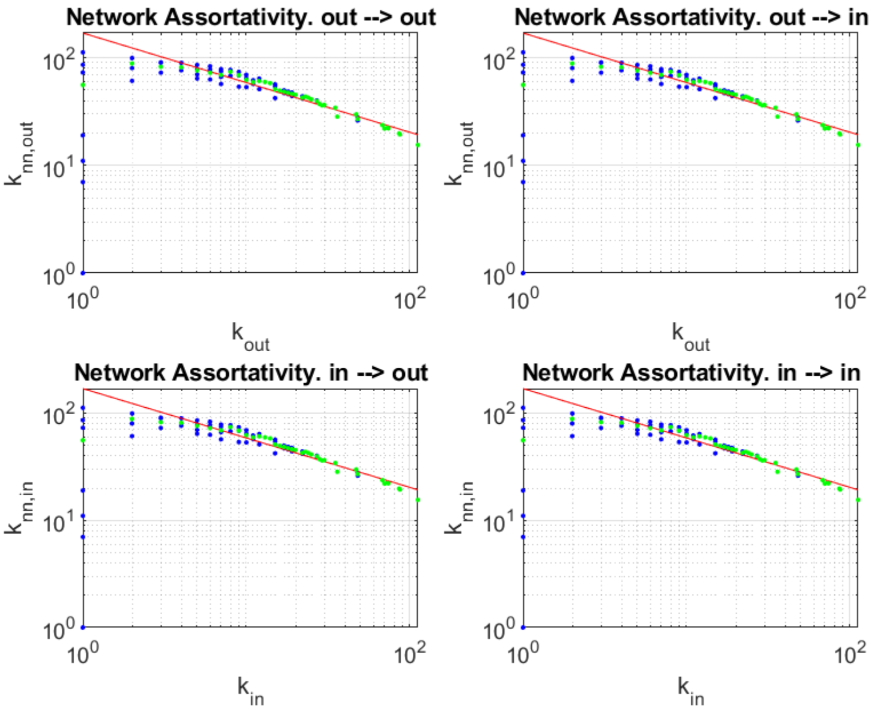

Figure 7: Neighbouring degrees and assortativity. Assortativity interpolation = red line, average neighbouring degree = green dots, average degree of the neighbours = blue dots.