A mini-projet on Complex Network Analysis

Duration

The duration of this project is typically 5 to 6 hours.

Objective

The goal of this mini-project is to develop your skills on complex network (graphs) analysis. You will choose a dataset (see list of dataset repository on the slide 41 of the course) and develop an exploratory analysis to understand and discuss some important properties of the graph.

Keys points

You are free to focus your analysis on the factors you consider relevant for the network dataset you will have chosen.

You will at least study the following points:

1. Basic connectivity analysis
    1. Data description: origin of the data, and for each graph, number of nodes and edges.
    2. Is the graph connected (or how many connected components ?)
    3. Compute, plot and comment the distribution of nodes' degrees
    4. Are the (these) graph(s) scale free ?
    5. According to various **centrality measures** (betweenness, etc), which are the most important nodes ?
2. Find the communities.
    - Use the Louvain algorithm to find the communities and visualize the results.
    - Describe quantitatively and qualitatively the communities.

Note: if the graph is large, communities can be used to produce a large scale plot (like the example we saw in the course).

Tools

Choose the right tool. Among the possibilities, consider NetworkX and Gephi, or a combination of both. It's important to produce nice figures.