

# Complex Networks - TD3

Remy Cazabet, Lorenza Pacini

2019

## 1 Network analysis

### 1.1 In short

1. Choose a real network
2. Use what you learned in the lectures to analyse it
3. Write a report illustrating your experiments and insights

### 1.2 Format

We expect at least a jupyter notebook, containing both code, results and discussions. If there is a need to develop some ideas or include external resources, a notebook might not be enough. Feel free to join a text document, and any another useful resource.

### 1.3 Choosing a network

You are totally free in choosing your network. You can probably do one of the following:

#### 1.3.1 Network repository

You can download an existing dataset already in network format from a network repository.

You can find a list of network repositories, for instance, here: <https://kateto.net/2016/05/network-datasets/>. Be careful that if your data does not contain node attributes (e.g., age, location, country...), you won't be able to interpret much. If the network is on a topic you have no idea about, it will also be hard to make sense out of it.

### 1.3.2 Create you own network

If you're interested in a topic in particular, you might be able to easily create your own network. Basically, what you need is to end up with an edge list (that can be easily imported in networkx and any other tool). For instance, if you're interested in movies, you could download the MovieLens dataset <https://grouplens.org/datasets/movielens/latest/>, that collects rankings given by users to movies. From it, you can build a bipartite, weighted network (user-movie), and/or derive a movie similarity network (counting the number of similar votes by the same people ? Or something more clever).

## 1.4 Visualization

A nice graph drawing is worth 1000 words. You can learn to use a software like Gephi (<https://gephi.org>) to make nice drawings. It is much easier than doing the same thing with networkx. Everything is done using the graphical interface, so you should master it quickly. The recommended way to use it is to import csv tables in which you already have computed all the nodes/edges properties that you like, and then to color/size nodes according to those properties. Note that some basic things like some centralities can also be computed directly with Gephi.

## 1.5 What should I do ?

Everything is fine. Try to tell a story. Show us what you have learned.

Ok, here is for instance a naive example: <https://networkofthrones.wordpress.com/the-novels/a-game-of-thrones/>, but you should be able to do much better. If you have a real question you would like to answer on real data, it's the occasion.

If you're particularly interested in a method rather than a dataset, you can also investigate more precisely this method (or set of methods) on several datasets. For instance, you might want to compare several methods for community detection on several datasets, or try to design your own network generator and show how good its properties are... The only limit is your imagination ! (And time, unfortunately...)