



---

# COMP5313

# Lab 5

---

## Network visualization

The goal of this lab is to understand how to use [Gephi](#) to visualize large-scale networks and to outline properties like communities (i.e., regions), betweenness and degrees.

Make sure your machine runs Linux or reboot it.

### Exercise 1: Get and launch Gephi

Get Gephi, unzip it and launch it: Download <https://launchpad.net/gephi/0.8/0.8.2beta/+download/gephi-0.8.2-beta.tar.gz>:

```
1 wget https://launchpad.net/gephi/0.8/0.8.2beta/+download/gephi-0.8.2-beta.tar.gz .
2 tar -xvzf gephi-0.8.2-beta.tar.gz
3 ./bin/gephi
```

Duration: 5 min

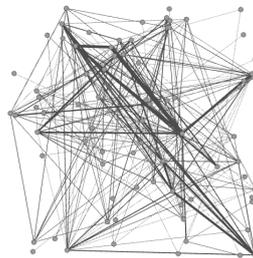


Figure 1: The initial graph corresponding to the dataset (node position is random at first)

### Exercise 2: Loading a network dataset

Collect the coappearance network of characters in the novel *Les Misérables* [1]. Its vertices are the characters of the novel and an edge indicates that the two characters appear together in the same chapter of the novel, at least once, as extracted by Knuth [2]. Download <http://poseidon.it.usyd.edu.au/~gramoli/largescalenetworks/labs/gephi/lesmiserables.zip>:

```
1 wget http://poseidon.it.usyd.edu.au/~gramoli/largescalenetworks/labs/gephi/lesmiserables.zip
2 unzip lesmiserables.zip
```

There is a list of nodes located in `lesmiserables_csv/node_list.csv` and a list of edges located in `lesmiserables_csv/edge_list.csv` as well as a GML file: `lesmiserables.gml`

representing the network of co-appearance of the characters of the book “Les Miserables” from Victor Hugo.

Create a new project in Gephi by going into the tab **File > New project**. Import the dataset by “opening” the GML file and specifying the graph as ‘undirected’.

**Remark:** You can alternatively choose to import the node list csv file, `lesmiserables_csv/node_list.csv` with a separator ‘Tab’ as table ‘Node table’ and click the ‘Next’ button. In import settings, choose main character: ‘Integer’ and click finish. You also need to import the edge list but it is typically harder to specify that the graph is undirected.

*Duration: 10 min*

### Exercise 3: Scattering and coloring nodes by degree

Click on the ‘Choose a layout’ on the left column of the window and choose “Force Atlas” with a repulsion strength of 5000 and press ‘Run’. (Note that there are other layout method you could use like Fruchterman Reingold.)

Now, rank node by ‘Degree’ in the top left corner of the window and ‘Apply’. Adapt the colours with ‘Spline’ and the parameter icon at the bottom.

On the left side panel called ‘Ranking’, choose the second icon from the left  also called size/weight. You can rank by degree by choosing “Degree” between 4 and 15 and the following spline:

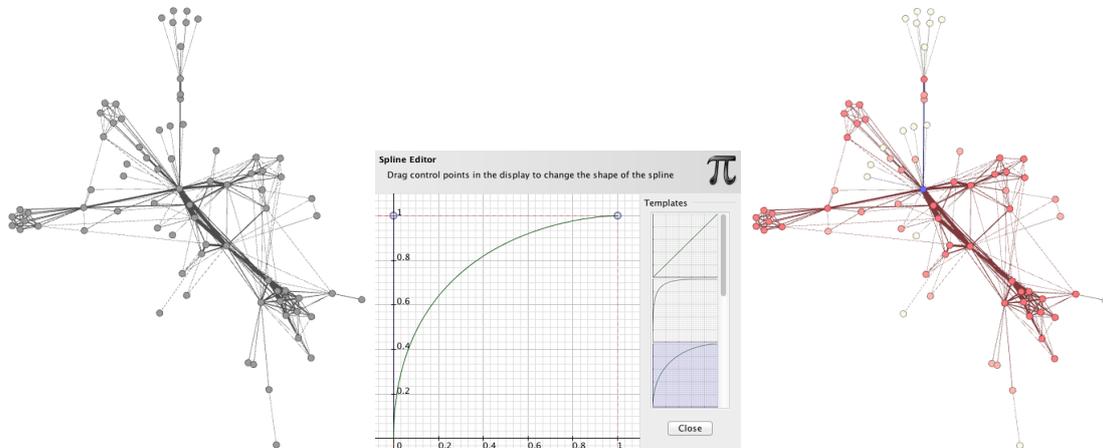


Figure 2: The graph with scattered nodes, the spline chosen and the graph obtained after degree coloring

*Duration: 10 min*

### Exercise 4: Computing more properties

You can compute and use more metrics using the ‘Statistics’ panel on the right-hand side of the window. For example, click the ‘Run’ button on the line of ‘Average degree’ to observe the average distribution.

Run 'Network diameter' and tick the box 'UnDirected', you can now see the betweenness centrality of the nodes and use it for representing nodes visually.

Duration: 5 min

## Exercise 5: Adjusting node sizes by betweenness centrality

The *betweenness centrality* of a node is similar to the edge betweenness in that it is the number of shortest paths of the graph that traverse this node. On the left-hand side panel called 'Ranking', choose the second icon from the left  also called size/weight.

With the scroll-bar you can now select 'Betweenness centrality' with a min size of 5 and max of 50. You should see the number of shortest paths traversing each node you have computed earlier by looking at the value associated with each node. Rerun "Force Atlas" after ticking the box 'adjusting by size'.

Show and adjust node labels with label and degree using the bottom bar



- show the label 
- show the font proportional to the node size , and
- adjust the font range.

Duration: 10 min

## Exercise 6: Community detection

Go to the Statistics panel on the right side and run 'Modularity' with default parameter to detect communities. Go to the Partition tab of the left panel and click the refresh button, . Scroll down to 'Modularity class', right-click on the proposed colors to randomise them.

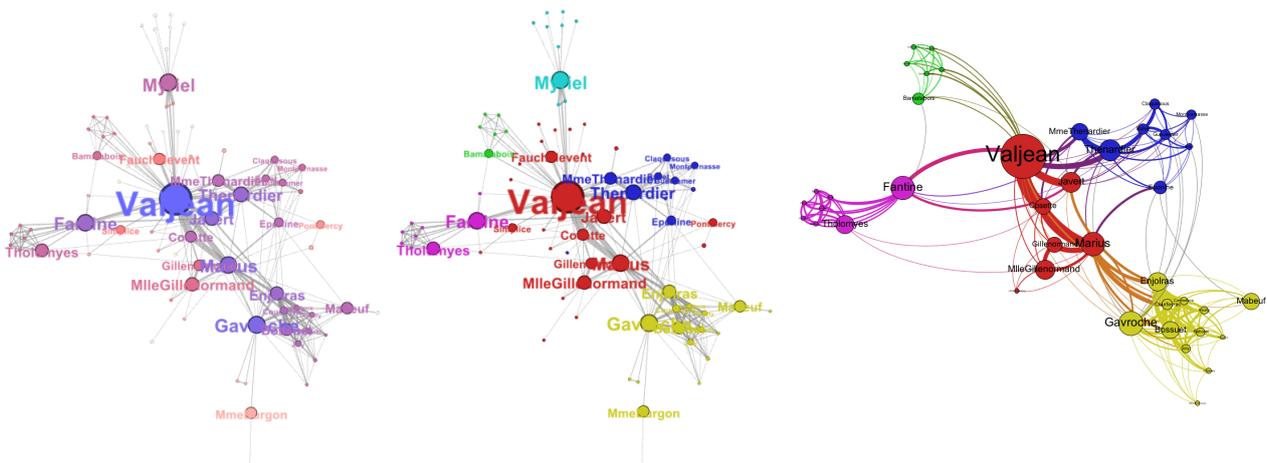


Figure 3: The graph with node resized (left), filtered communities (centre) and after preview (right)

To hide some nodes or edges, use the 'Filters' tab on the right-hand side panel. As a filter, use Topology, K-core and select 4 as the K-core setting. Finally, choose the 'Pre-

view' mode at the top of the main window, tick the 'Show Labels' box and choose **File > Export SVG/PDF/PNG file...**

*Duration: 10 min*

## References

- [1] Victor Hugo. Les Misérables. 1862.
- [2] Knuth, D.E. The Stanford GraphBase: A Platform for Combinatorial Computing. Reading (MA): Addison-Wesley. 1993.