Using filters

Clément Levallois

2017-02-10

Table of Contents

download a network file for this tutorial 1
open the network in Gephi 1
getting a sense of the attributes in the data laboratory 3
discovering the filter panel 5
combining 2 filters 7
Filter operators 12
    The MASK operator 12
(to be continued) 14
More tutorials on using filters in Gephi 14
the end 14

last modified: 2017-02-13

download a network file for this tutorial

download this zip file and unzip it on your computer.

You should find the file `miserables.gexf` in it.

Save it in a folder you will remember (or create a folder specially for this small project).

This file contains a network representing "who appears next to whom" in the 19th century novel *Les Misérables* by Victor Hugo [1: D. E. Knuth, The Stanford GraphBase: A Platform for Combinatorial Computing, Addison-Wesley, Reading, MA (1993)].

A link between characters A and B means they appeared on the same page or paragraph in the novel.

The file name ends with ".gexf", which just means this is a text file where the network information is stored (name of the characters, their relations, etc.), following some conventions.

open the network in Gephi

- open Gephi. On the Welcome screen that appears, click on `Open Graph File`
- find `miserables.gexf` on your computer and open it

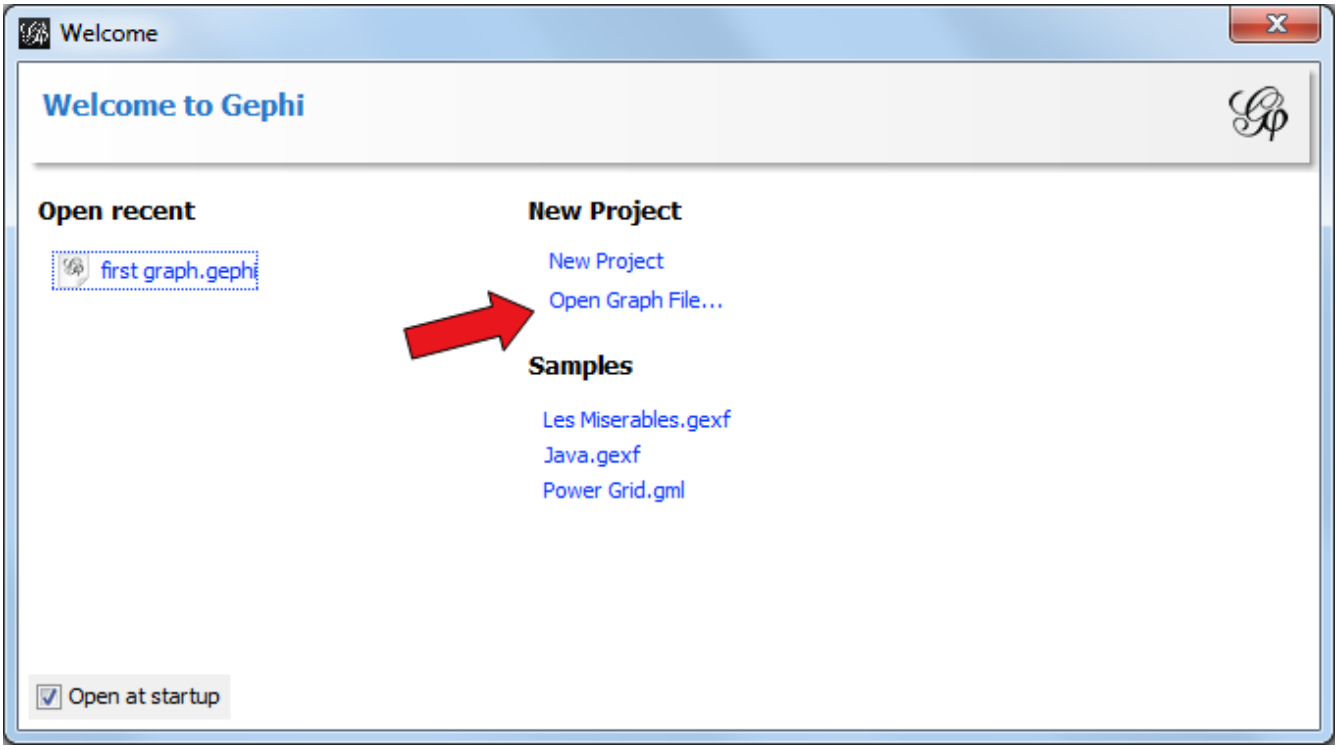

Figure 1. welcome screen

A report window will open, giving you basic info on the network you opened:

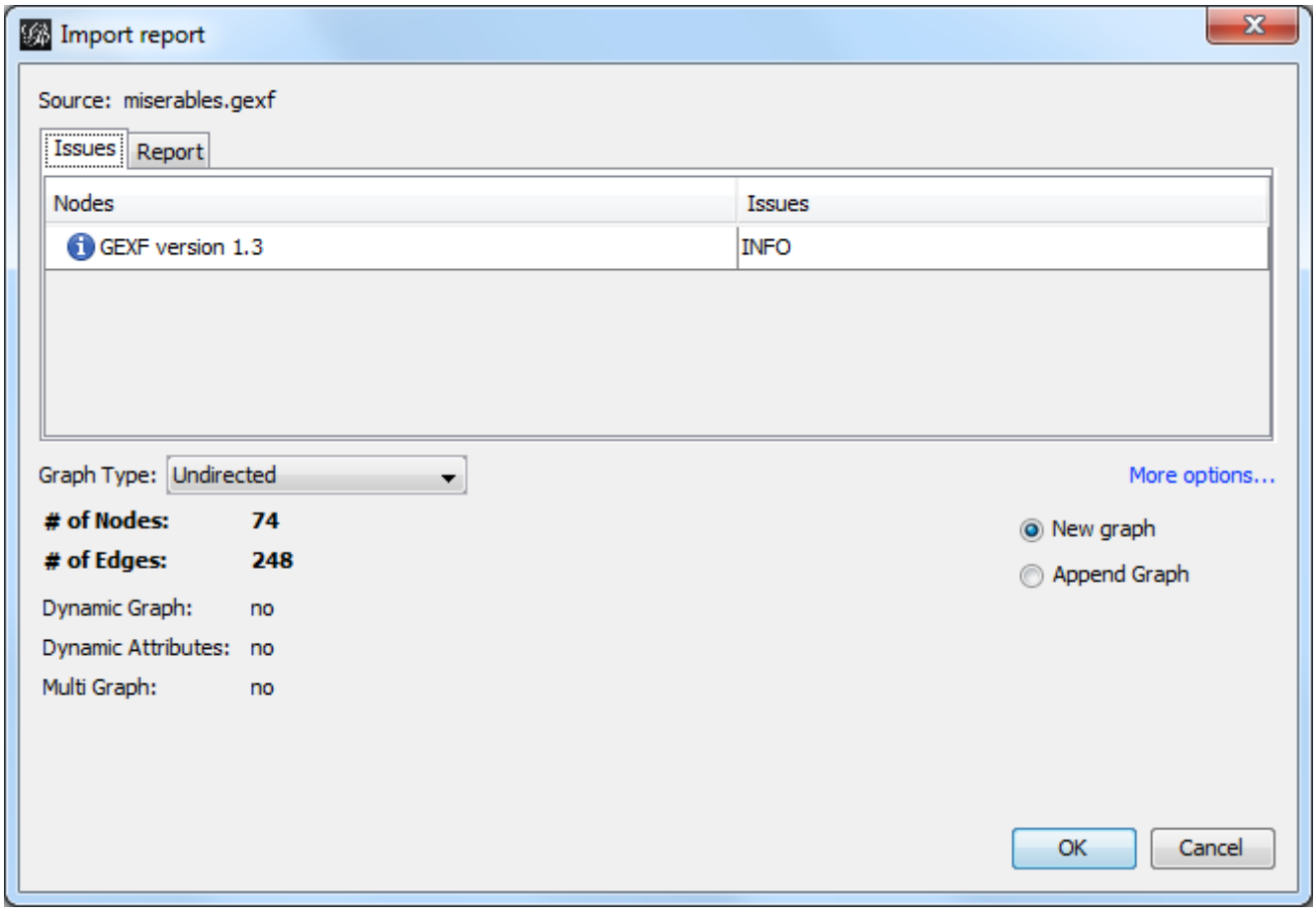

Figure 2. report window

This tells you that the network comprises 74 characters, connected by 248 links.

Links are undirected, meaning that if A is connected to B, then it is the same as B connected to A.

The report also tells us the graph is not dynamic: it means there is no evolution or chronology, it won’t "move in time".

Click on `OK` to see the graph in Gephi.

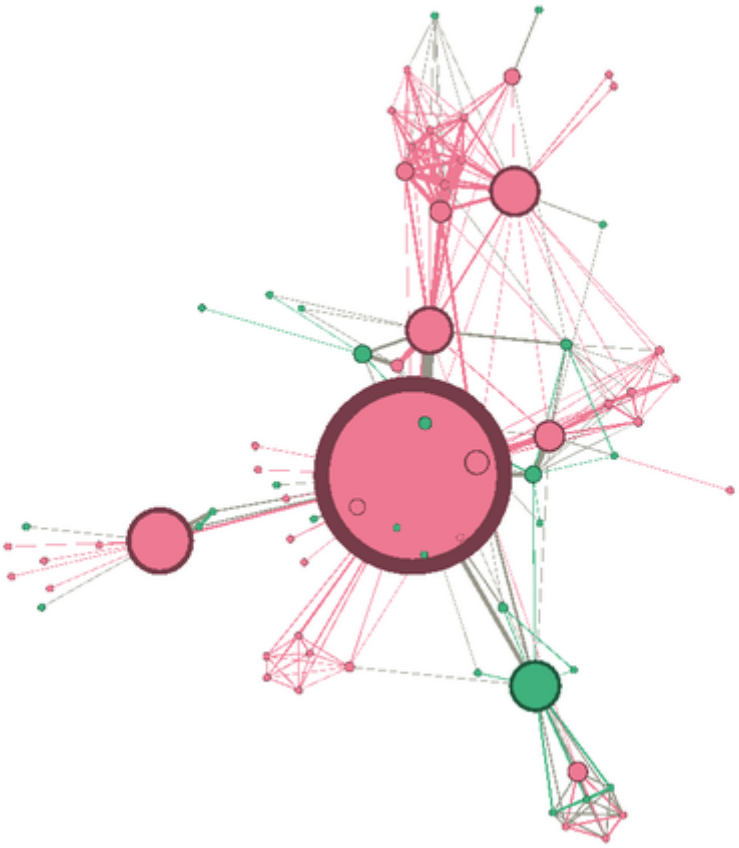

Figure 3. The network we will use in this tutorial

getting a sense of the attributes in the data laboratory

We can switch to the data laboratory to see the underlying data:

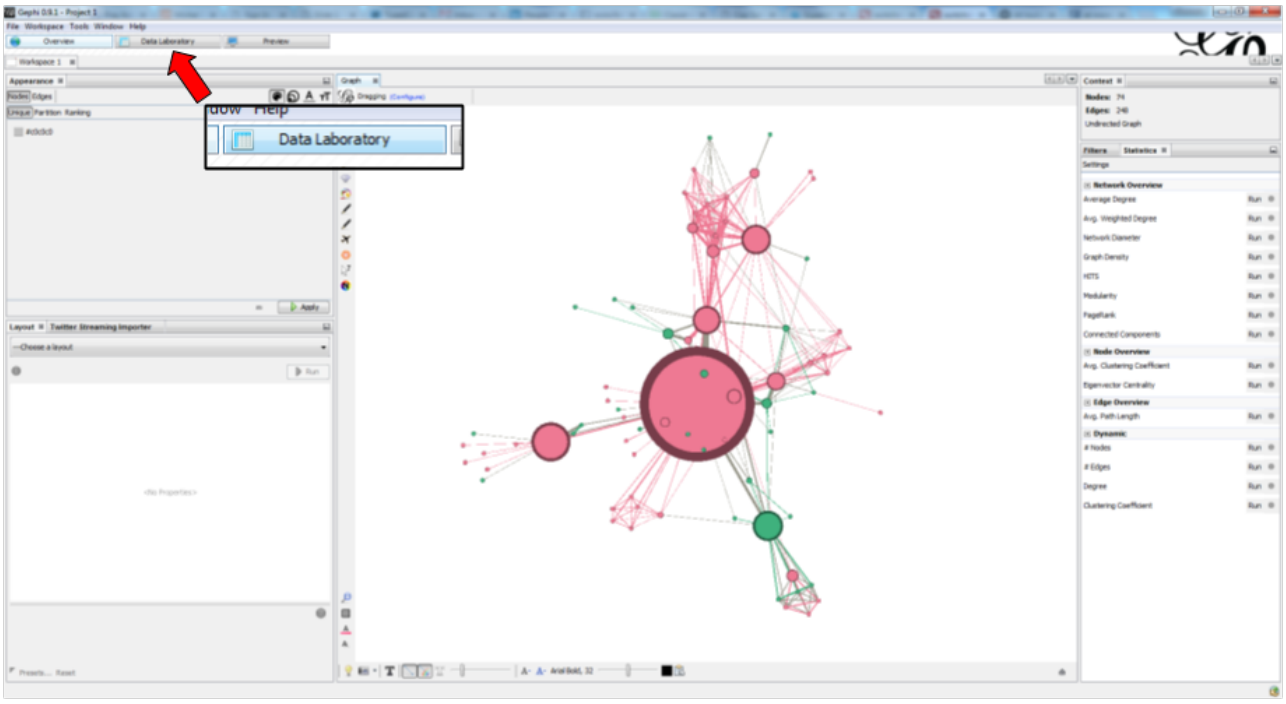

Figure 4. Switching the view to the data laboratory

We see that the nodes of the network have many attributes. In particular, each have a Gender and a measure of how central they are:

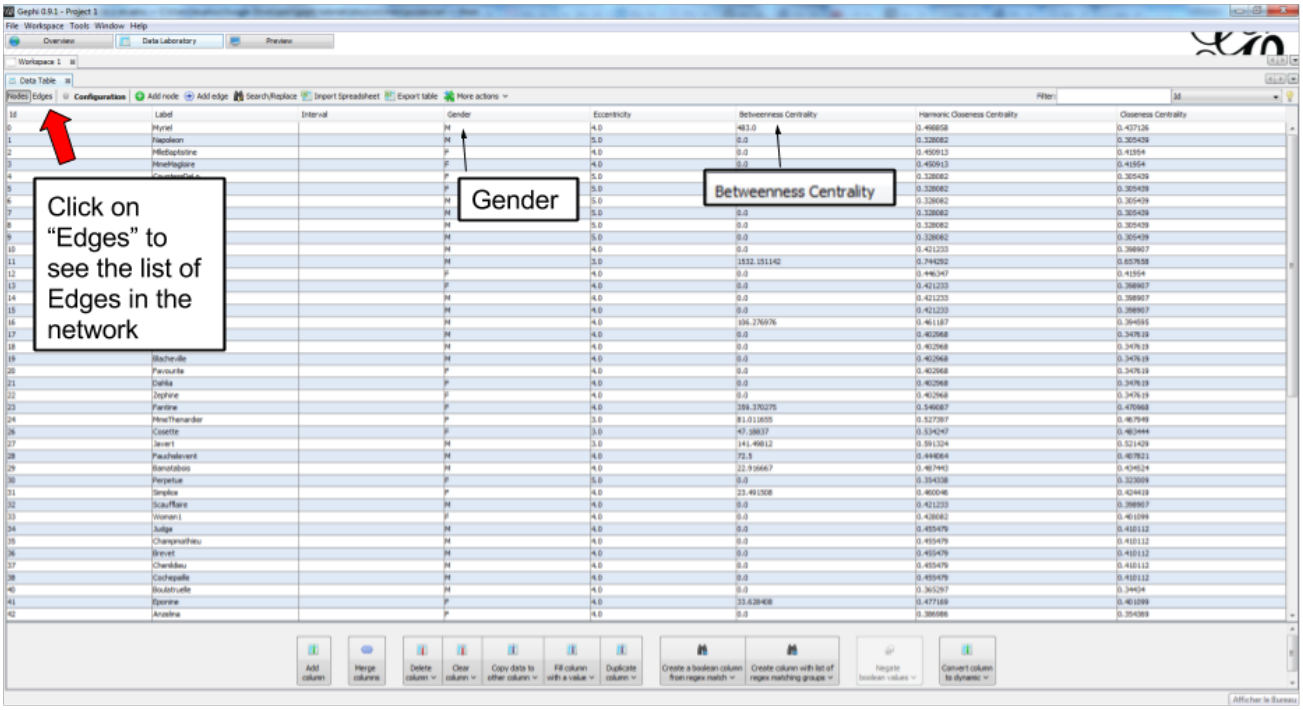

Figure 5. Nodes attributes.

This is the list of edges (relations) in the network. Notice that they have a "weight" (a "strength").

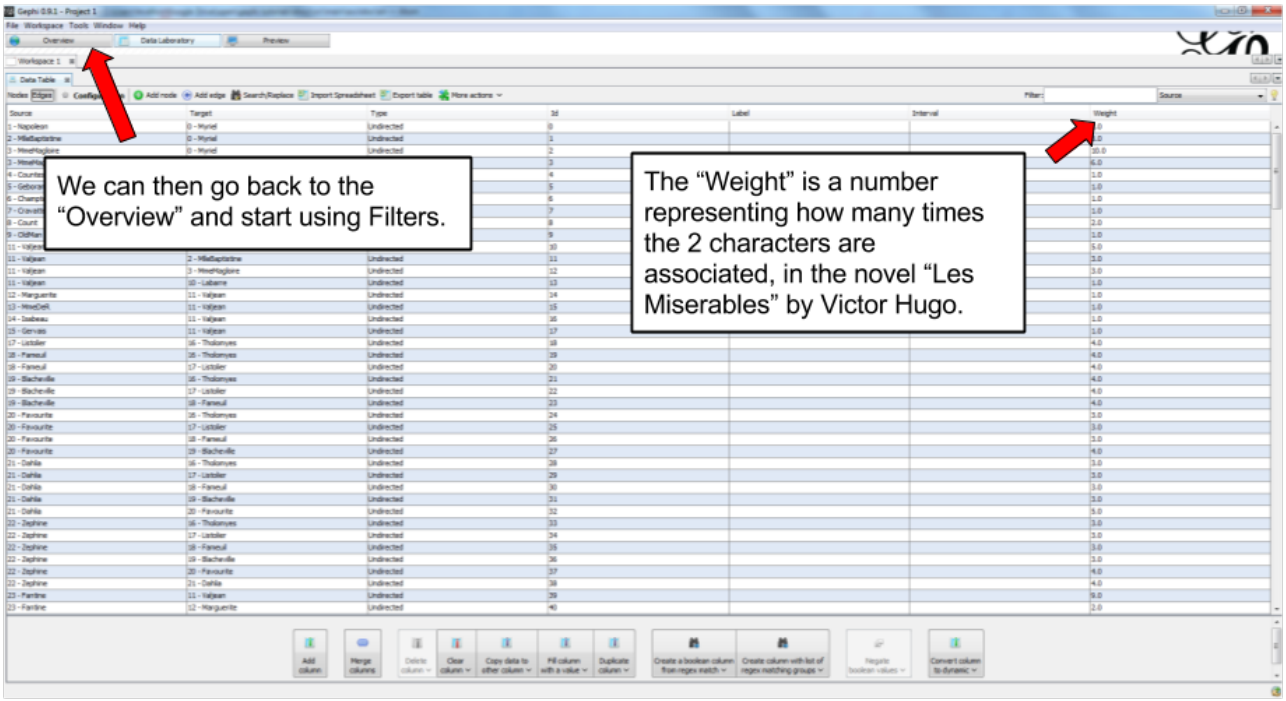

Figure 6. Edges attributes

discovering the filter panel

In the overview, make sure the Filter panel is displayed:

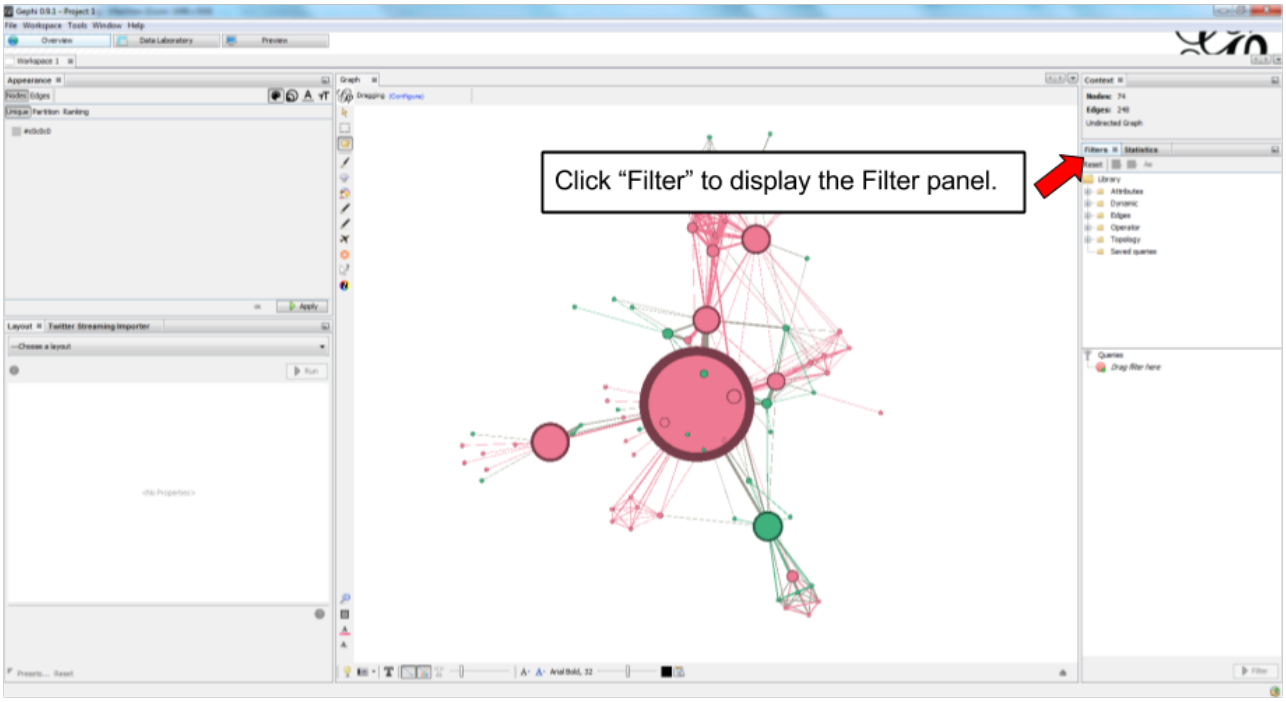

Figure 7. Making the Filter panel visible.

How the Filter panel works:

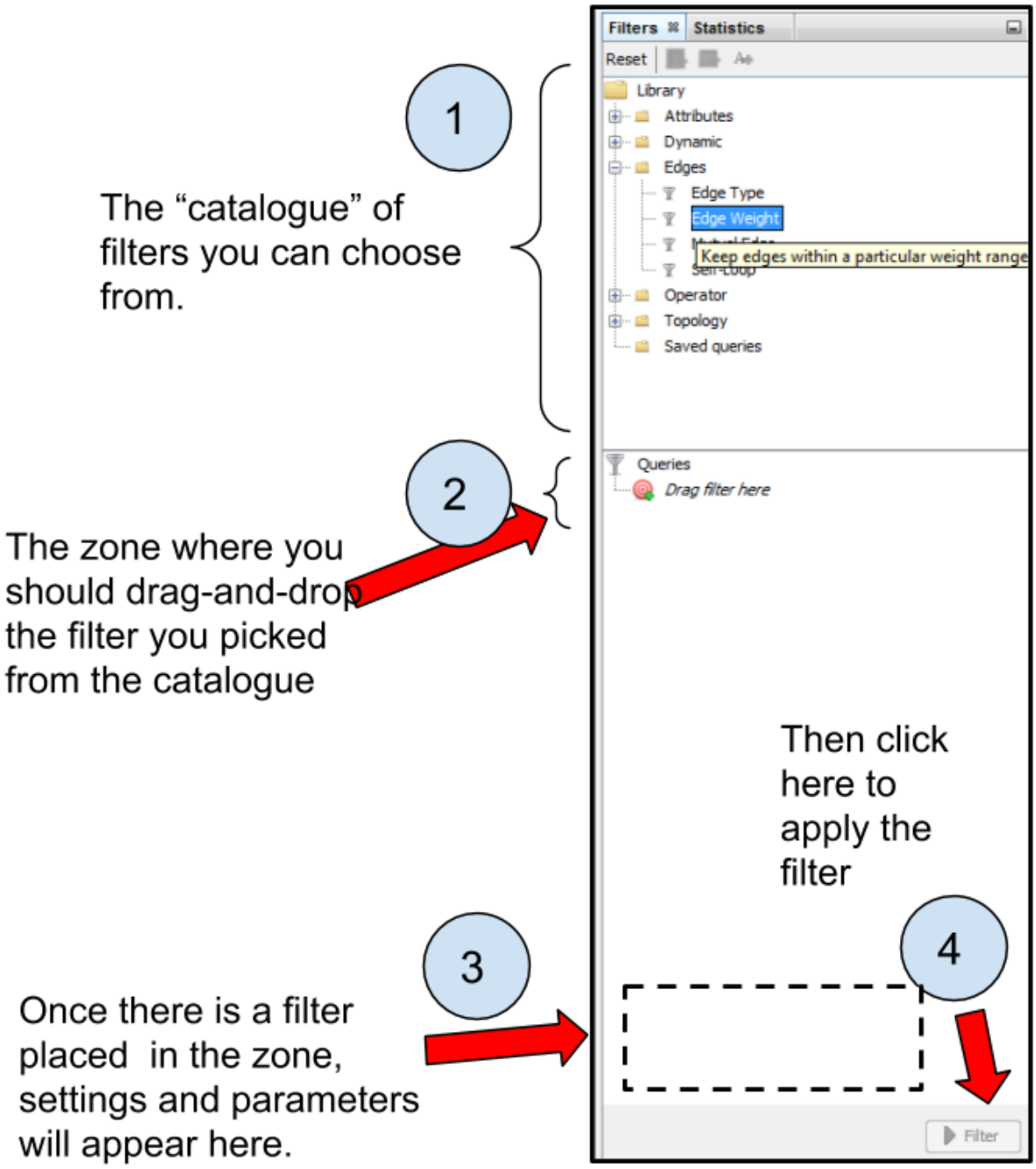

Figure 8. Workflow of filters.

An example: filtering out the edges which have a weight value lower than 2:

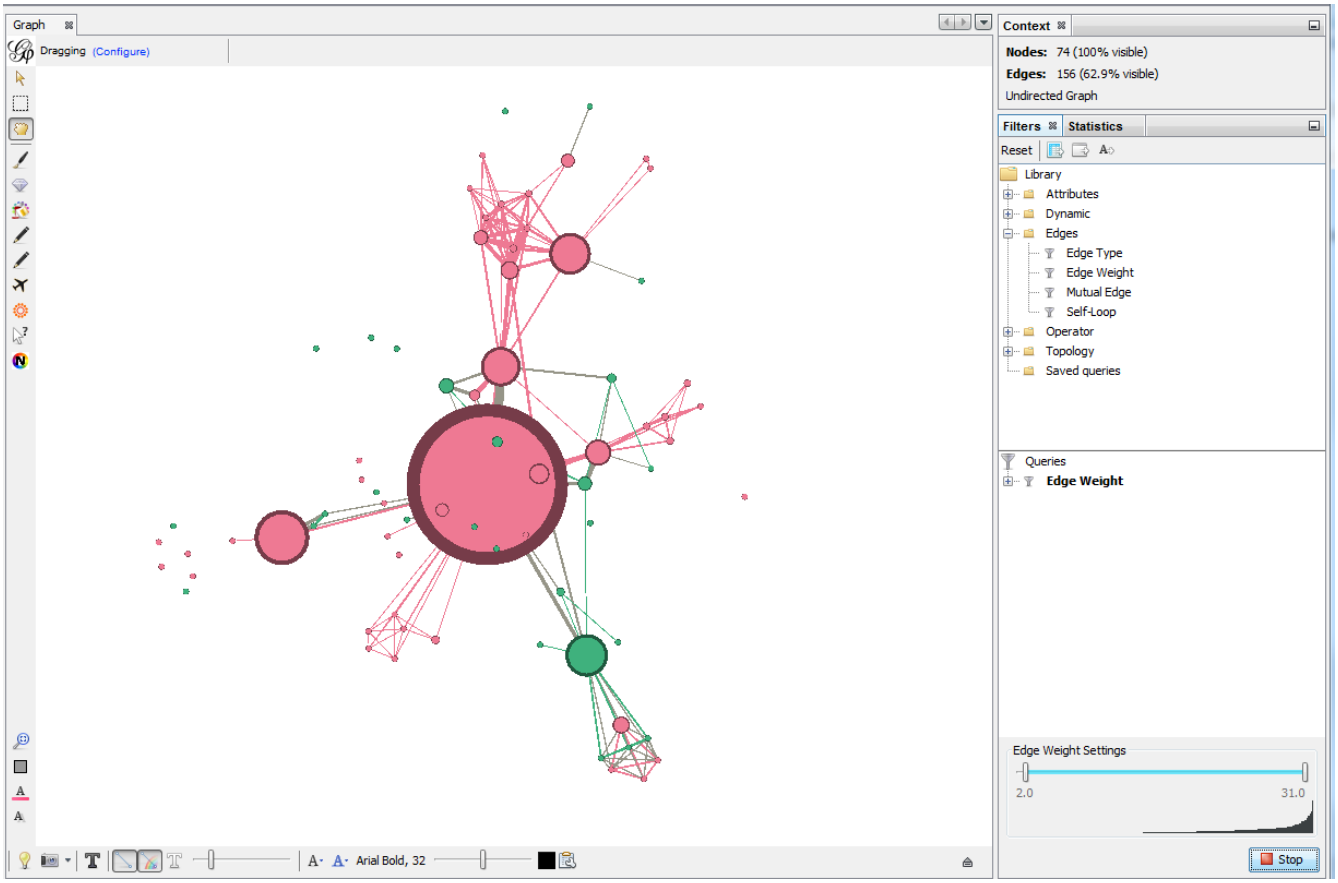

Figure 9. Filtering out edges with weight lower than 2.

view online animation - link: https://tinyurl.com/gephi-tuto-2

When you are finished using a filter in the zone, right click on it and select "remove".

combining 2 filters

One filter is applied AFTER this other:

The first filter to be applied is NESTED (placed inside) the second one as a "subfilter"

Which filter should be placed inside which? Let's look at different examples:

1st Example:

Keeping on screen only the female characters which have a tie (an edge, a relaion) of at least strength 2:

→ place the filter "edge weight" inside the filter "Gender":

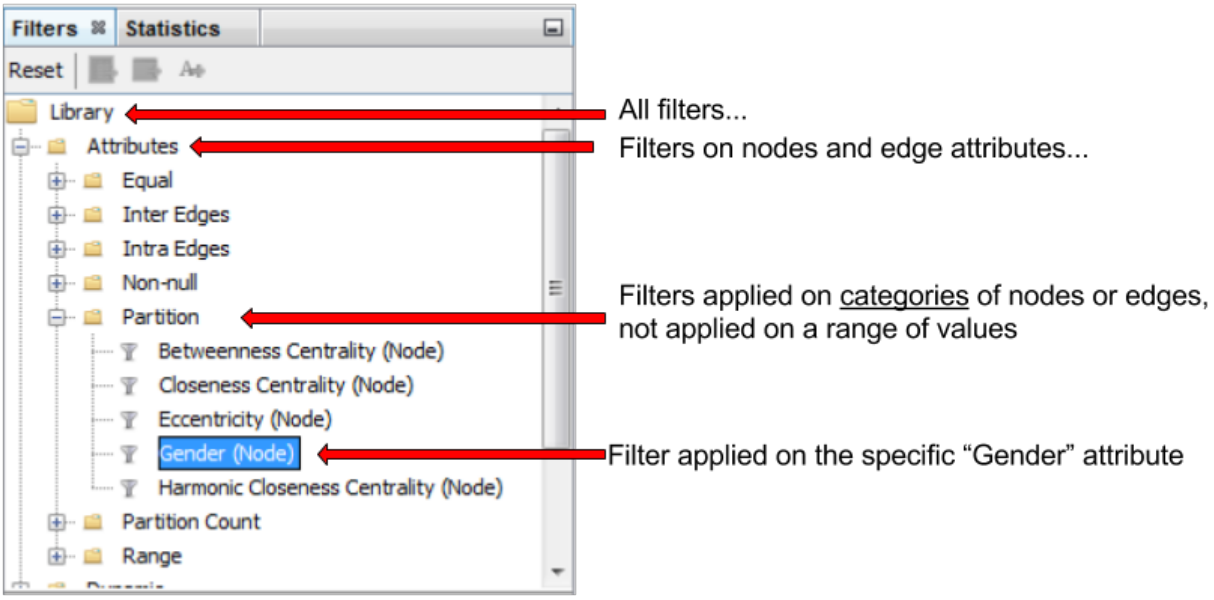

Figure 10. Where to find the filter on the Gender attribute

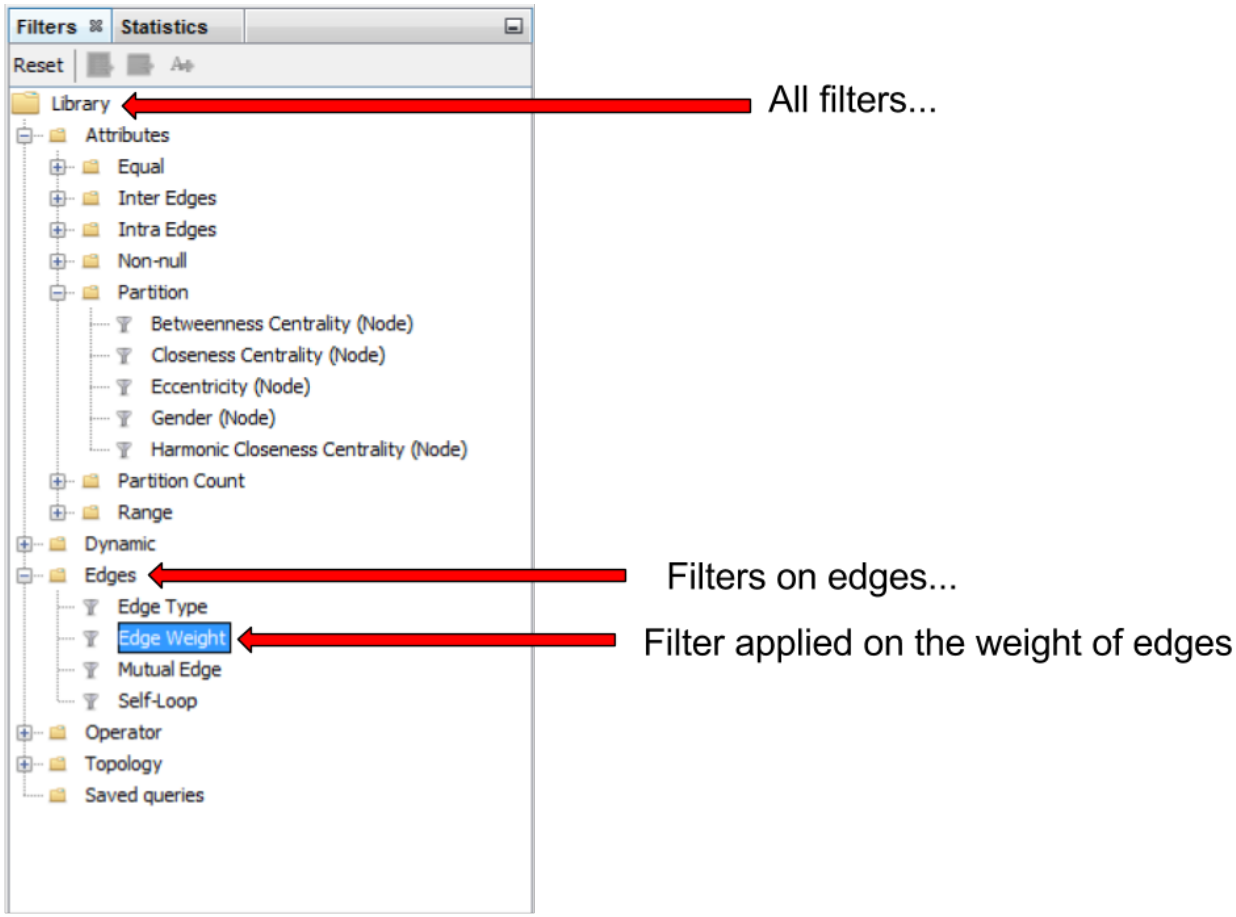

Figure 11. Where to find the filter on edge weight

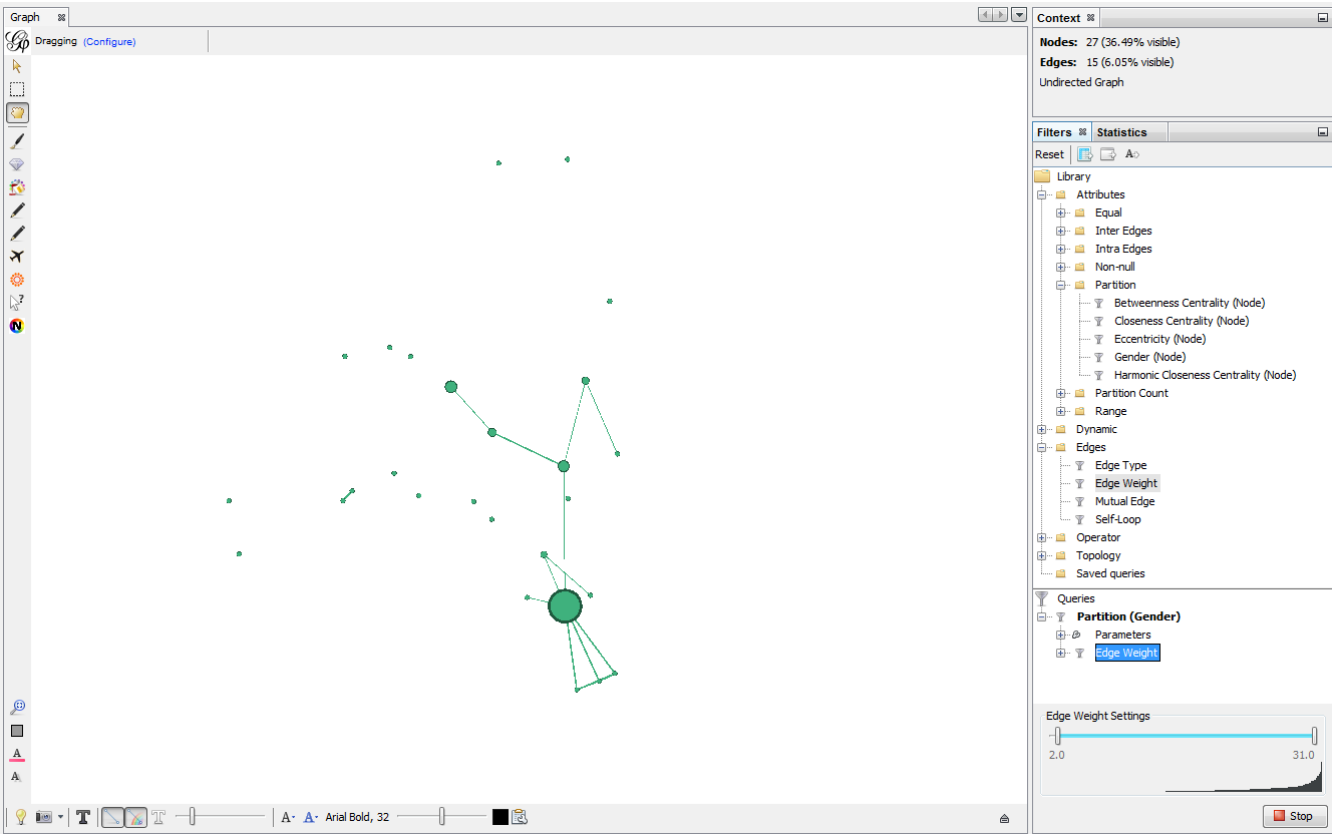

Figure 12. Keeping only female characters with tie of at least 2.

view online animation - link: https://tinyurl.com/gephi-tuto-1

In this case, it was equivalent to:

- nest the "Gender" filter inside the "Edge weight" filter or
- nest the "Edge weight" filter inside the "Gender Filter"

→ The result was the same. Now we will see a case where the placement of filters matter.

Here, we want to visualize:

- only the nodes which have **less than** 10 relations <1>
- and among these, only those which form the "main island" of the network (we want to hide small detached groups of nodes) <2>

① in technical terms, nodes with a `degree` of less than 10.

② in technical terms, we are looking for the `giant component`

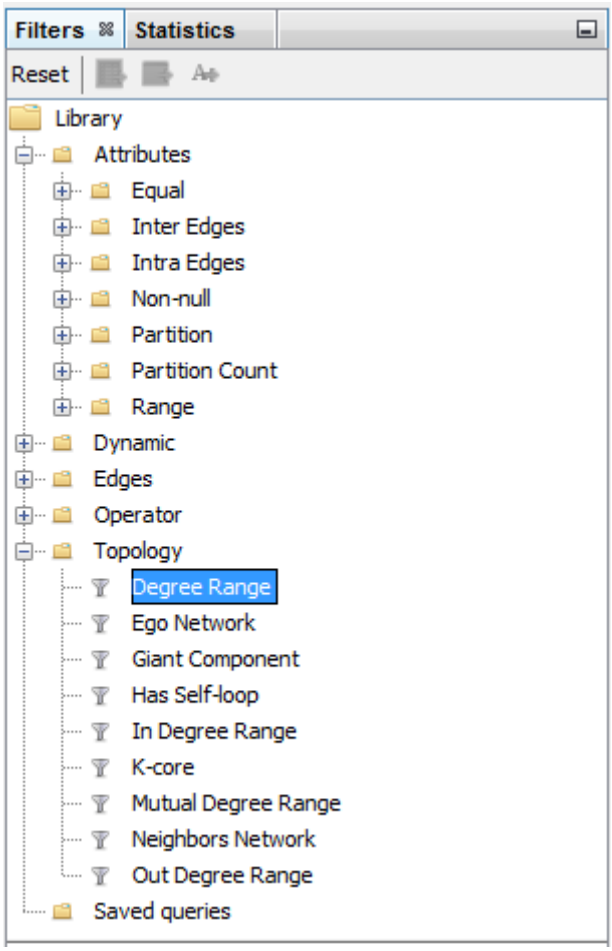

Figure 13. Where to find the filter on degree

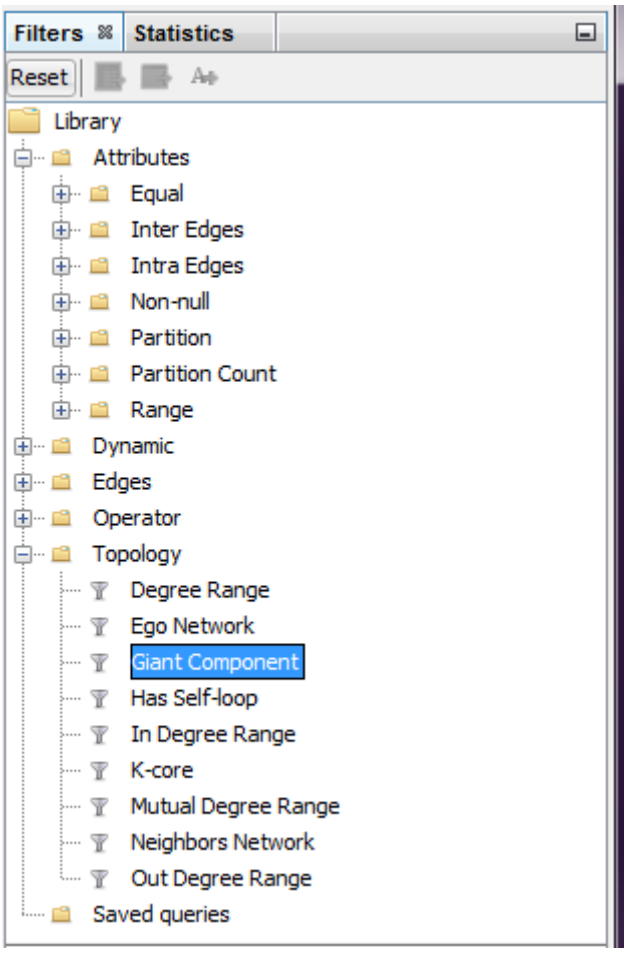

Figure 14. Where to find the filter on giant component

We will see that the placement on the filters in the zone will make a difference.

First, let us place the filter on giant component **inside** the filter on degree:

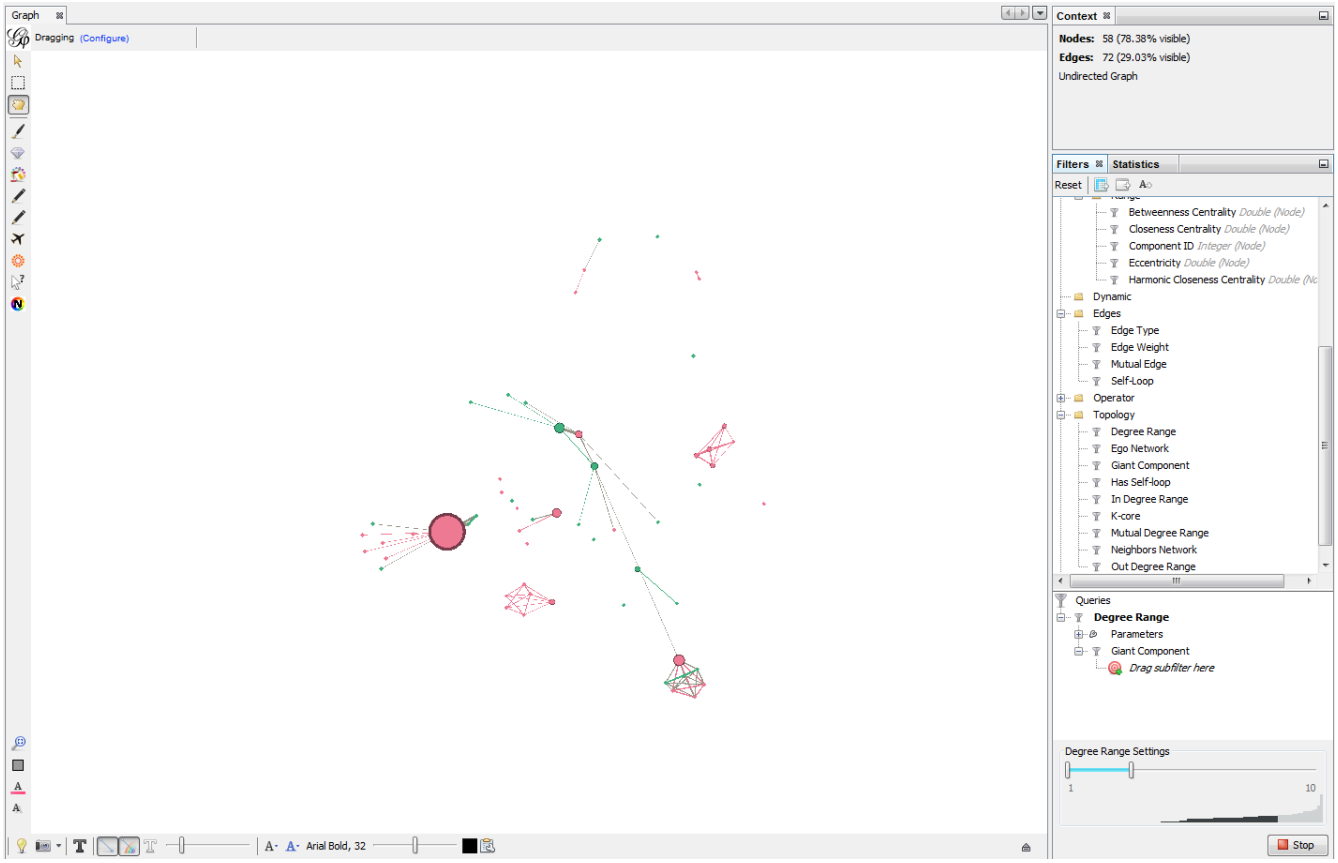

Figure 15. Filters in one configuration

In this first case,

- only the giant component of the network was made visible.

→ Since the network was just one big connected "island" to start with, it did not change a thing.

- then, all characters with more than 10 relations where hidden

→ this hides nodes which were connecting with many others, so that we end up with many groups, disconnected from each others.

Now instead, placing the filter degree **inside** the filter on giant component:

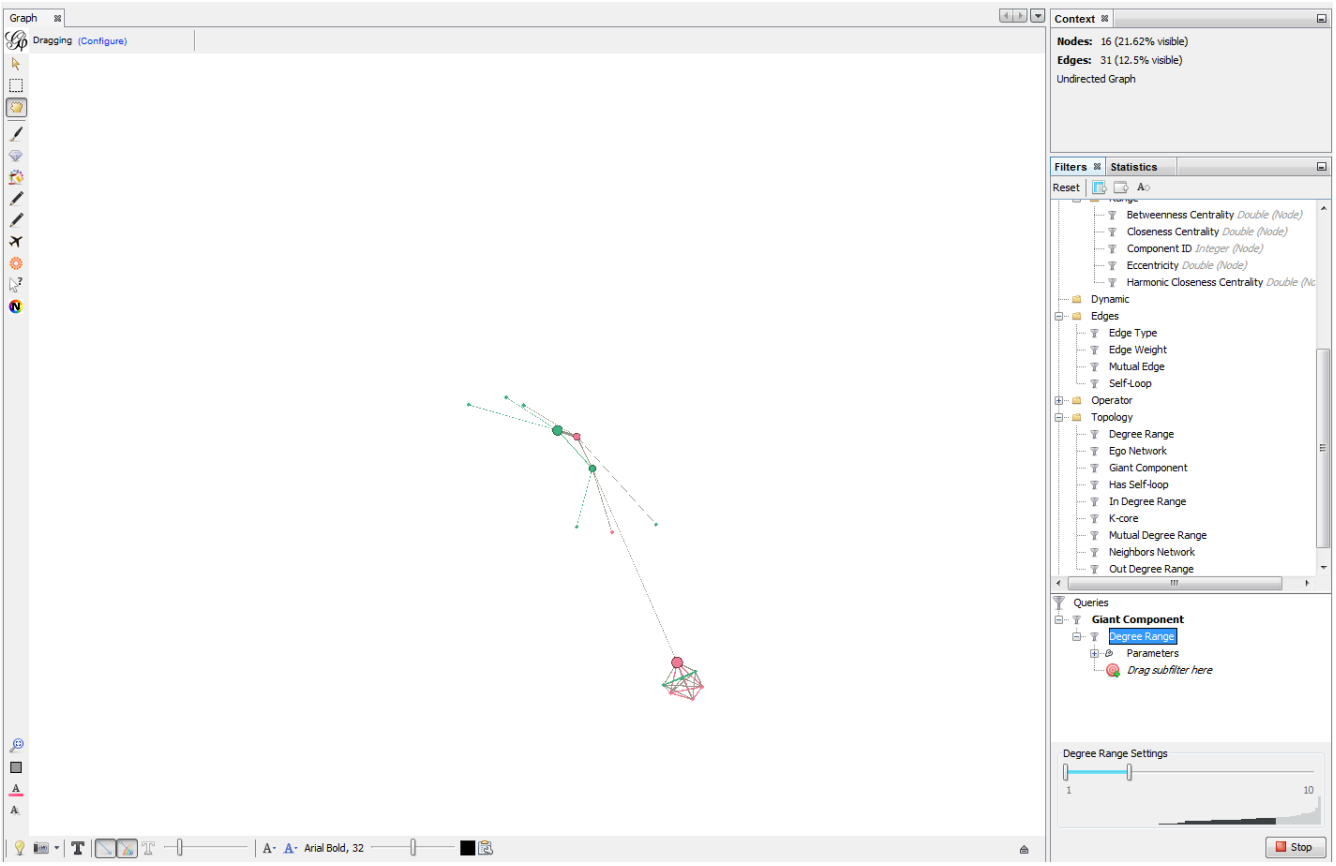

Figure 16. Same filters in another configuration

In this second case,

- starting from the complete network, all characters with more than 10 relations where deleted.

→ this created a network made of many disconnected groups of nodes

- then the giant component filter is applied,

→ which had for effect to hide small groups, to keep in view only the biggest group of connected nodes.

In summary: be careful how you apply several filters at once, this might have an effect on the logic of filtering.

Filter operators

The MASK operator

Imagine you are interested in the female characters of the novel "Les Miserables".

- you are interested in these characters and the relations among them
- you are interested in the relations between female characters and male characters
- you are **not** interested in the relations between male characters
- How to make appear only female characters, their relations (to female and male characters) and

only those?

The MASK operator enables you to: - apply a filter as you would usually do, - and then, to make visible again **only the relations that have a connection with the visible nodes**

So this shows female characters, relations between them, and their relations to male characters. Male-male relations are not visible:

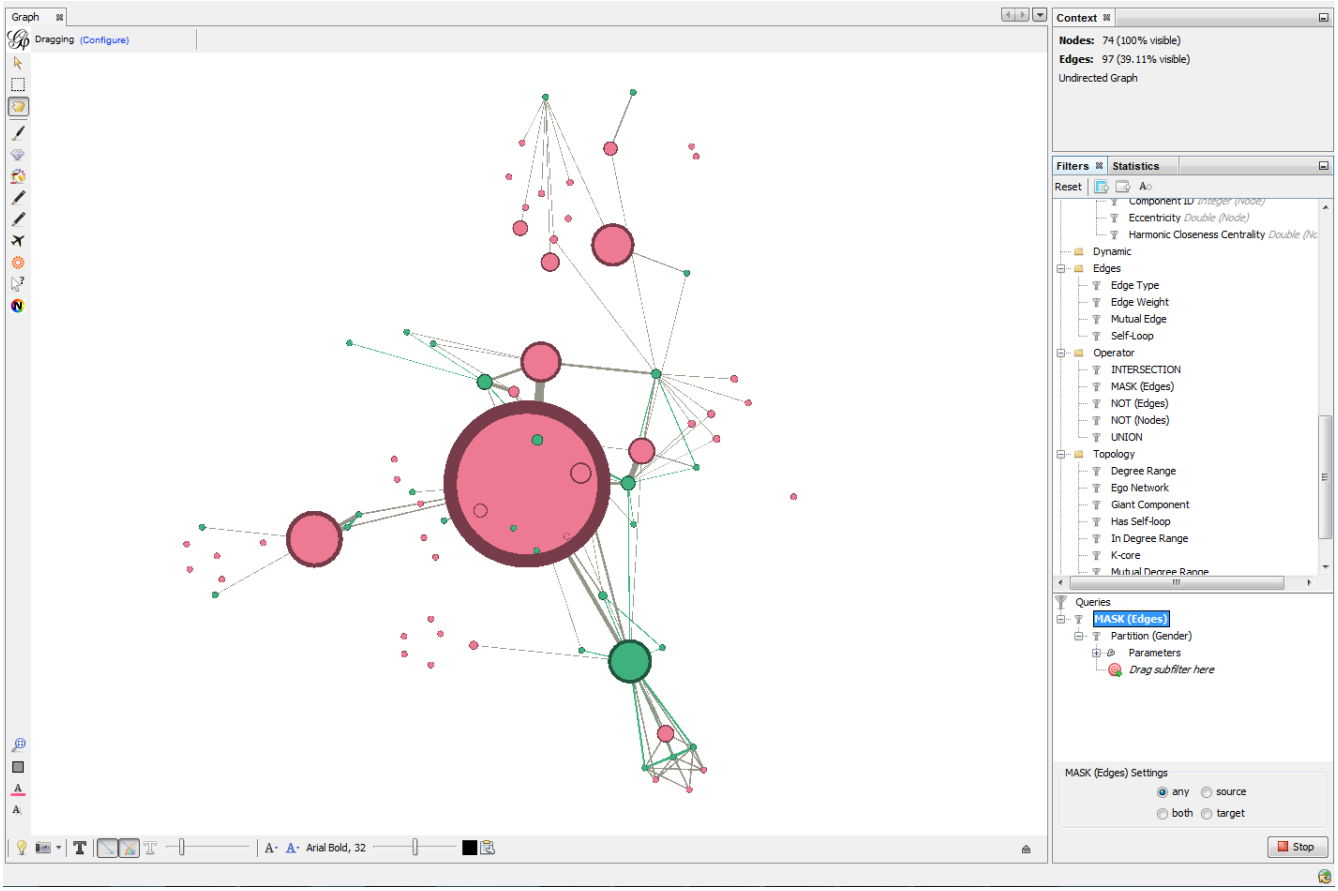

Figure 17. Using the MASK operator

It is also possible to hide / show only some of the directed relations between the visible graph and the filtered out graph:

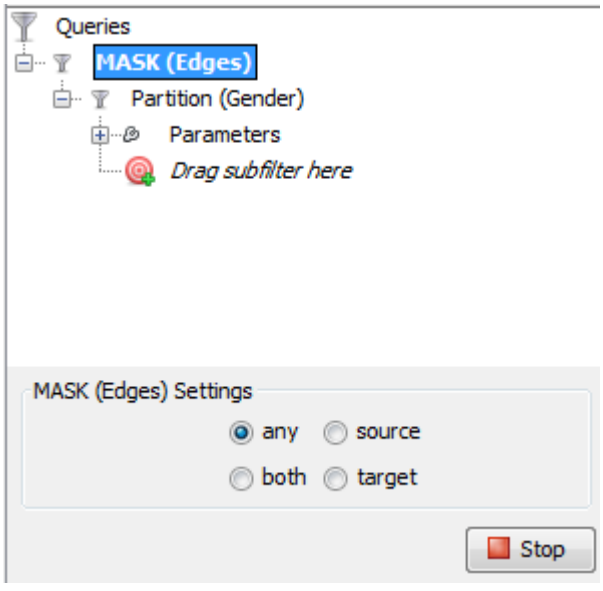

Figure 18. Parameters of the MASK operator

(to be continued)

More tutorials on using filters in Gephi

- Video on using filters by Jen Golbeck

the end

Visit the Gephi group on Facebook to get help,

or visit the website for more tutorials