SEN124A: Exercise Python and Gephi

For this exercise you are expected to use the networkX python package and the Gephi network analysis tool. To start you need two files:

- `ex-length.xlsx` - an excel file with a network where the weight is representing the length of the different edges.
- `ex-traveltime.xlsx` - an excel file with the same network, where the edge weights are representing travel time.

Note that networkX documentation can be found here: `https://networkx.github.io/documentation/stable/index.html`. Use the search bar to find useful function for network analysis.

Part 1: network analysis in Python using networkX

Make sure that the two files above as well as the file `Exercise SEN124A.ipynb` are in the same folder. Start Jupyter notebook, browse to the folder and open the notebook.

1. Start by loading both files and create for each a directed graph.
2. Visualize both graphs and very that they are both directed graphs (indicated by arrows) and that they differ form each other.
3. For both graphs calculate the shortest path between nodes 'A' and 'I'. Note that the paths should be different.
4. For both graphs calculate the pagerank values for the nodes.
5. For both graphs determine the weight edge betweenness centrality values.

Part 2: network visualization in Gephi

After finishing the first part, open the file `Exercise SEN124A.gephi` in Gephi. This is the same network as in the excel file `ex-length.xlsx`.

1. Verify that the file represents indeed the same network.
2. Add the pagerank values from the python analysis as node attributes.

3. Add the normalized edge betweenness centrality from the python analysis as edge attributes
4. Visualize the network, whereby:
   - node size represents the relative pagerank value of a node
   - edge thickness represents the normalized edge betweenness centrality
   - the nodes are labeled by name
   - the edges don't overlap

   The final network should look something like shown in Figure 1.

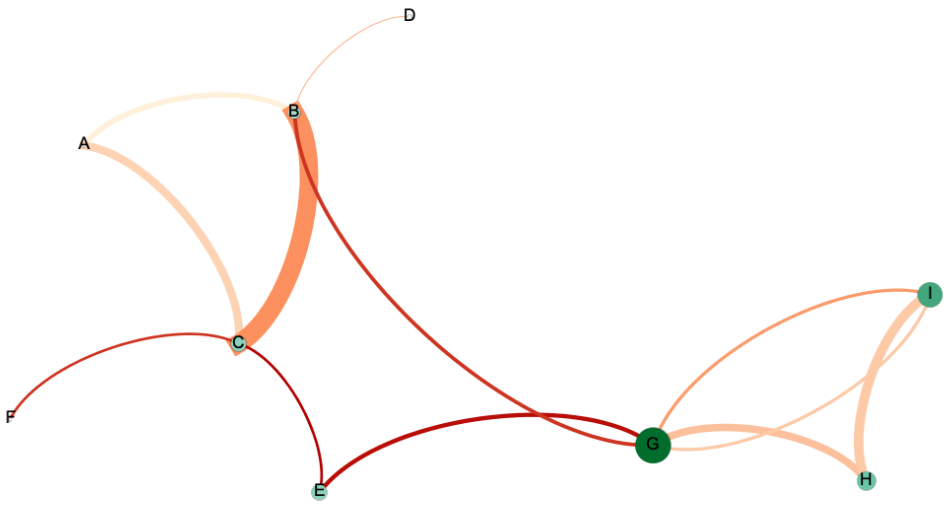

Figure 1: Sample solution of the exercise.

Part 3: Bonus

Visualize the network with travel time as edge attribute in the same way as the network in Part 2. Compare the two networks. Can you say something about the efficiency of both networks?