CS5834, Fall 2023: Intro to Urban Computing
Homework Assignment 1
US domestic air traffic network analysis

Assigned: Sep 5, 2023   **Total points**: 100 points
Due: Sep 19, 2023, 11:59pm ET

In this homework, you will study the domestic air traffic network of the US. You should use the range of network science methods covered in class to visualize and analyze the network using the *networkx* library in Python and articulate various observations from the analysis. Particular tasks can be found below.

Your submission should contain the following:

1. A PDF file of executed code from Task 1.

2. A report answering questions from Task 2.

Please place the two files in a single folder and compress the folder. Please name the compressed file in this format: yourfullname_HW1.fileformat. The "fileformat" for Task 1 will be PDF and that for Task 2 will be DOC or PDF.

Link to code: Available with this assignment.

The assignment provides code that is very similar to what was shown during the lecture (the Citibike examples). The code is tailormade to be run on the US air traffic dataset.

Task 1 (10 points):

a. Run the *.ipynb* notebook from the given link, make necessary changes to the current path as mentioned during lecture 1 (please refer to the slides) and run all the cells.

b. After all the cells are run, please go to "Files" and select print. Here please select "Save as PDF" in the destination section in the print window and save the document in your system. Submit this PDF file.

Task 2:

Considering the context of domestic air traffic in US, study and explain various properties of the network.

a. What is a node and what is an edge in the context of air traffic? How many nodes and edges are present in the network? **(10 points)**

b. What can you can infer from the visualization of the network? Speak about the structure of the network and describe it qualitatively. **(10 points)**

c. What can you tell from the degree distribution and degree rank plots? **(10 points)**

d. How connected and dense is the network? What does this imply? (**10 points**)

e. What is the value of the average clustering coefficient? What does this imply? **(10 points)**

f. Which are the most important nodes based on the Pagerank score and what do you learn from this?**(10 points)**

g. Which are the most important hub and authority nodes and what do you infer from this? **(10 points)**

h. What are the values of the various centrality measures? What can you infer from these scores? **(10 points)**

i. What can you infer from the strongly and weekly connected component plots? **(10 points)**