CSE 610 HW #1

(due Sep 23)

Homeworks are to be done individually. Show your work and explain any assumption you make. Please use an editor, submissions with hand-writing will not be accepted. Email your work in pdf format before the class starts at 3.30. Have fun!

1. **(40 pts)** Consider the directed network in Figure 1;

   - Can you classify the nodes according to the bow-tie structure? Which nodes take part in SCC, IN, OUT, and tendrils?
   - How do you increase the size of SCC by a single edge modification in Figure 1? You can either insert an edge or remove an existing one.
   - What if you want to maximize the size of SCC by a single edge modification? Consider the modification which would increase the SCC size most.
   - How do you increase the size of IN by a single edge modification in Figure 1? You can either insert an edge or remove an existing one.
   - How do you increase the size of OUT by a single edge modification in Figure 1? You can either insert an edge or remove an existing one.

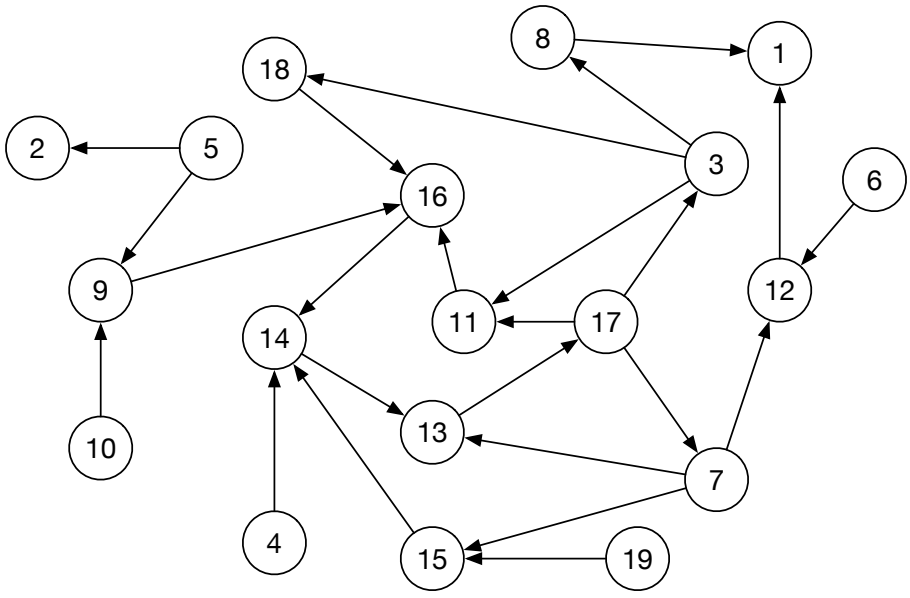

Figure 1: A directed network

2. **(30 pts)** Consider the attached bipartite network `bipartite.txt`. Construct the unweighted projection graph based on 1) left set, 2) right set. For each projection, give the following:

   (a) Number of nodes (with non-zero degree) and number of edges.

   (b) Degree distribution (i.e., histogram of degrees). Create at least 10 bins. Show the minimum and maximum values on both axes.

3. **(30 pts)** Consider the attached directed network `directed.txt`.

   - Find the top 10 vertices with the largest pagerank values – give the vertex ids and pagerank values. Set the damping factor to 0.75 (i.e., teleport probability is 0.25) and the number of iterations to 50.
   - Modify the pagerank algorithm such that if a vertex u with pagerank value r_u has out-degree d_u, each of u's out-neighbors gets $r_u/({d_u}^2 - d)$ votes (instead of r_u/d_u as in the original version). Keep the damping factor and the number of iterations same as above. Find the top 10 vertices with the largest pagerank values – give the vertex ids and pagerank values.