Social Network Analysis

5/04/2018

Name ____________________________
Student ID ____________________________

Note: Whenever an exercise requires the application of a known formula both the formula and its solution must be reported and discussed.

Exercise 1: Graph Modelling [4 points]
Given the matrix $\mathcal{G}$

$$\mathcal{G} = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 \end{bmatrix}$$

- Draw $\mathcal{G}$;
- Synthetically characterize the graph $\mathcal{G}$ describes (directedness, number of nodes/edges, density, components, max/min/avg degrees...).
- Is the graph planar[1]?

Solution

- The resulting graph is shown in Fig. 1. We assume outgoing edges on the rows, incoming on the columns.
- The graph is composed by 6 nodes and 12 edges: it is directed, has density $\frac{12}{30} = 0.4$, it is composed by a single weakly connected component. $min(d_{in}) = min(d_{out}) = 1$, $max(d_{in}) = 3$, $max(d_{out}) = 4$, $avg(d) = 2$.
- Yes the graph is planar.

[1] In graph theory, a planar graph is a graph that can be embedded in the plane, i.e. , it can be drawn on the plane in such a way that its edges intersect only at their endpoints.

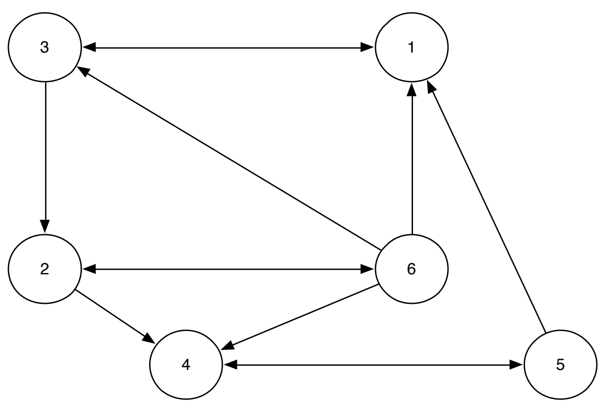

Figure 1:

Exercise 2: Synthetic graphs [7 points]
Let $\mathcal{G}$ be a Barabasi-Albert graph with $N = 2300$ and $m = 7$:

- How many edges are in $\mathcal{G}$?
- What is the expected degree of the largest hub?
- What is the expected clustering coefficient of $\mathcal{G}$?

Let $\mathcal{E}$ be an Erdos-Renyi graph having the same size and order (i.e., number of nodes and edges) of $\mathcal{G}$:

- What is the value of p that allows to generate $\mathcal{E}$?
- What will be the average degree of $\mathcal{E}$ nodes? and the graph density?
- Describe the regime of $\mathcal{E}$.

Solution

- $|E| = mN = 7 * 2300 = 16100$
- $k_{max} = k_{min} N^{\frac{1}{\gamma-1}} = 7 * 2300^{\frac{1}{2}} \simeq 336$ ($K_{min} = m$, in BA $\gamma = 3$)
- $CC = \frac{(ln|N|)^2}{|N|} \simeq 0,026$
- $p = \frac{2|E|}{|N|(|N|-1)} \simeq 0.006$
- $avg(d) = \frac{2|E|}{|N|} = 14$
- $D(\mathcal{E}) = p$
- $p > \frac{ln(|N|)}{|N|}$, the regime is connected.

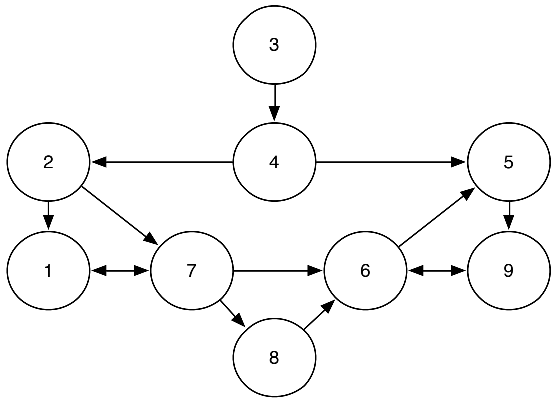

Figure 2:

Exercise 3: Paths [7 points]
Given the *directed* graph $\mathcal{G}$ shown in Figure 2:

- Is the graph *strongly* connected? If so, compute its diameter, otherwise compute the diameter of its biggest strongly connected component.
- List all the shortest paths among the pairs [1,5], [3,8] and [9,2];
- Compute the edge betweenness of (3,4), (2,1) and (6,5);
- Is it possible to identify an Hamiltonian cycle on $\mathcal{G}$? If not specify the minimum set of edges to needed to build it.

Solution

- *"In the mathematical theory of directed graphs, a graph is said to be strongly connected or diconnected if every vertex is reachable from every other vertex. The strongly connected components or diconnected components of an arbitrary directed graph form a partition into subgraphs that are themselves strongly connected."* The graph is not strongly connected. The biggest strongly connected component is (6,5,9) since there exist at least path connecting all the possible pairs in such set. The diameter of the biggest strongly connected component is 2.
- $[1,5] = \{1-7-6-5\}$
  $[3,8] = \{3-4-2-7-8\}$
  There not exist paths connect node 9 to node 2.
- $BE(3,4) = 8$
  $BE(2,1) = 3$
  $BE(6,5) = 6$
- No, it is not. The minimum set of edges needed to build an Hamiltonian cycle is $\{(9,3)\}$.

Exercise 4: Indicators [7 points]
Given the graph $\mathcal{G}$ shown in Figure 2 compute:

- Degree Centrality (in/out/global) of all nodes;
- Closeness Centrality of 3, 8, 2;
- Betweenness Centrality of 8, 9;
- Local Clustering Coefficient of 2, 6, 9 (assuming the graph undirected).

Solution

- $d_1^{out} = d_3^{out} = d_5^{out} = d_8^{out} = d_9^{out} = 1$
  $d_2^{out} = d_4^{out} = d_6^{out} = 2$
  $d_7^{out} = 3$
  $d_3^{in} = 0$
  $d_2^{in} = d_4^{in} = d_8^{in} = 1$
  $d_1^{in} = d_5^{in} = d_7^{in} = d_9^{in} = 2$
  $d_6^{in} = 3$
  $d_3 = 1$
  $d_1 = d_2 = d_4 = d_5 = d_9 = 3$
  $d_6 = d_7 = 5$
  $d_8 = 2$

- Closeness centrality
  In directed graphs the Closeness centrality index of a node u needs to be computed taking into account only the nodes reachable from u. This is to avoid generating an erroneous ranking. Although ill-defined, for this exercise only, we will accept also spurious solutions that compute the centrality over the complete graph.

  $Cl(3) = \frac{1}{8}(1*1+2*2+3*3+2*4) = 2,75$
  $Cl(8) = \frac{1}{3}(1*1+2*2) \simeq 1,7$ (Substitute $\frac{1}{3}$ with $\frac{1}{8}$ for the compete graph)
  $Cl(2) = \frac{1}{6}(2*1+2*2+2*3) = 2$ (Substitute $\frac{1}{6}$ with $\frac{1}{8}$ for the compete graph)

- $BE(8) = 0$
  $BE(9) = 2$

- $CC(2) = \frac{1}{3}$
  $CC(6) = \frac{2}{6} = \frac{1}{3}$
  $CC(9) = \frac{1}{1} = 1$

Exercise 5: Graph Construction [7 points]
Given 10 nodes and, at most, 20 edges build a graph such that all the following conditions hold:

- The graph is composed by two separated components;
- There exists a path of length 4 between nodes 1 and 2;
- Node 2 has a clustering coefficient of $\frac{2}{3}$;
- The shortest path among 3 and 1 is equal to 2;
- Node 6 has the lowest Degree Centrality;
- Node 7 has the highest Closeness Centrality;
- Edge (8,10) has the lowest betweenness centrality.

Solution
Several solutions are possible, among them the one reported in Fig 3. N.B.: Closeness and Betweenness – as all measures involving paths – are defined, and computable, only component wise.

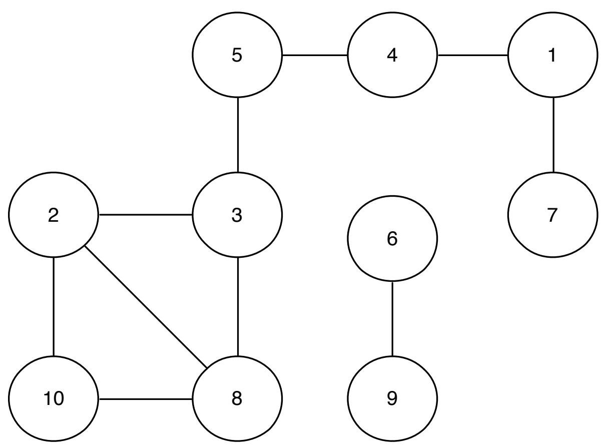

Figure 3: