Network Science Exam I

November 10th, 2023, Fall Semester 2023/2024 2 hours

The exam can be answered in Portuguese or in English in a separated exam sheet. All 8 questions are equally valued.

I. Let G be a large undirected network and assume that the Louvain algorithm is used to determine the community structure of G. What graph representation should be considered for better performance? Justify your answer and explain how that representation affects the algorithm running time and memory requirements.

Solution hints: Since the Louvain algorithm iterates at each step over all neighbors of each vertex, to achieve $O(n+m)$ time per pass, either adjacency lists or sorted edge lists should be used.

II. Define and explain the difference between closeness centrality and betweenness centrality, detailing the concepts captured by each one.

Solution hints: Closeness is a measure of how close an individual is to all other nodes of the network. Betweenness is a measure of the influence a node has over the spread of information through the network.

III. Growth and preferential attachment are considered two fundamental principles underlying the emergence of power-law degree distributions in real systems. But most empirically observed networks also portray a significant saturation for low degrees, prominent exponential cutoffs for high degrees, and often a high clustering coefficient. Propose an algorithm capable of generating a network portraying a power-law degree distribution and at least one of the previous properties observed in real networks. Briefly explain why your algorithm works.

Solution hints: For example, one can obtain significant saturation for low degrees with the Barabási-Albert model with initial attractiveness, prominent exponential cutoffs for high degrees with the Barabási-Albert model with ageing, and an high clustering coefficient with the Dorogovtsev-Mendes-Samukhin (DMA) minimal model.

IV. Scale-free networks are known to be extremely robust against random failures. How does network robustness changes with degree assortativity? Justify by relating to what is known for real networks.

Solution hints: Degree assortativity refers to the tendency for vertices to attach to others that have similar degree. There is no strong correlation between changes in assortativity and changes in network robustness against random failures, although empirical analysis of real networks, for which degree assortativity effects have been observed, shows that robustness against random failures tends to increase.

V. Provide a model that can help in characterizing the cascading effects on complex networks.

Solution hints: A simple model for cascading events: i) Consider a network of size N; ii) All nodes have a state initialized to "0" (healthy); At each time step, each node will turn "1" if at least a fraction φ of its neighbors is also "1" (i.e., have also failed). Run many simulations for the same φ and the average degree, $\langle k \rangle$. Vary φ and $\langle k \rangle$.

VI. Network topology drastically impacts the dynamics of spreading processes, offering distinct predictions for the outbreak of diseases on random (Erdős-Rényi, ER) graphs and on scale-free networks. Resort to the SIS model and the second moment of the degree distribution to illustrate such differences.

Solution hints: Taking the SIS model as a reference, a random ER network has a finite epidemic threshold λ_c, implying that a pathogen with a small spreading rate ($\lambda < \lambda_c$) is likely to die out. If, however, the spreading rate of the pathogen exceeds λ_c, the pathogen becomes endemic, and a finite fraction of the population is infected at any time. For a scale-free network, we have $\lambda_c = 0$. Hence, even viruses with a very small spreading rate λ can persist in the population. In general, λ_c scales with the inverse of the second moment, $< k^2 >$, of the degree distribution. Since the second moment of the degree diverges for scale-free networks with an exponent $\gamma \leq 3$, the outbreak threshold will vanish, i.e., $\lambda_c = 0$.

VII. Clarify the distinction between simple and complex contagion processes. To help clarify the difference, please provide an example for each.

Solution hints: Contagion processes on networks, including disease spreading, can be modelled as simple contagion, i.e., a contagion process where it is sufficient to come into contact with an infected individual to be infected (e.g., flu). Many social phenomena, however, are better modeled as complex contagion processes, in which multiple interactions are required for a contagion event to occur. The diffusion of memes, products, and behavior is often explained by the concept of complex contagion. It highlights the fact that most individuals do not adopt a new meme, product or behavioral pattern as a result of a single contact. Instead, adoption requires multiple contacts with several individuals who have already adopted the behavior. For instance, the more friends a person has who own a mobile phone, the higher the likelihood that she will also purchase one.

VIII. Please indicate whether each of the following statements is TRUE or FALSE. (Note: For each wrong answer we discount a correct one.)

a) The local clustering coefficient C_i for a vertex i is given by the ~~proportion~~ number of links between the vertices within its neighborhood divided by the number of links that could possibly exist between them.

b) The Erdős–Rényi network model is a famous proof of principle whose main purpose is to capture the basic mechanisms responsible for the emergence of highly clustered structures in real-world networks.

c) The stag-hunt game (or coordination game) is a dilemma involving two players, leading to a stable co-existence between cooperation and defection.

d) Degree assortativity is a preference for a network's nodes to attach to others that have similar degree, k. It can be assessed through the shape of the k_{nn} distribution, i.e., the average degree of neighbors of a node with degree k. If this function decreases with increasing k, the network is assortative; if the function increases with increasing k, the network is said to be disassortative.

e) The three degrees of influence theory states that each person is separated from everyone else by only three degrees.

f) A k-core is the largest subgraph whose nodes have at most k interconnections.

Solution: True, False, False, False, False, False.
