Test Exam (Re-exam 2017/2018)

You have two hours to complete this exam.

Write your name and student number on each sheet of paper that you hand in. Write the total number of sheets on the first sheet.

Give concise answers in clear Dutch or English. Write clearly. When not asked explicitly, you may directly use results discussed in class; mention this explicitly when you do.

If you are asked for a formula or calculation, make sure you explain your intermediate steps for possible partial credit.

The exam has five exercises. Make sure to allocate your time well; some exercises may be harder than others.

1. Erdős-Rényi ($2 = 1 + 0.5 + 0.5$ points)

Consider the Erdős-Rényi random graph model.

(a) Give the two main reasons why this model does not accurately capture most real networks. Explain your answers.
`The degree distribution is too concentrated around the mean, i.e. not power-law. The local clustering coefficient does not decrease with the degree of a vertex. +0.5 for each.`

(b) What is (roughly) the expected size of the largest component at the critical point of the model?
$N^{2/3}$. `+0.5 for the answer.`

(c) Consider a graph generated from this model such that $N = 5000$ and $p = 10^{-5}$. In which regime is the graph? *The correct answer yields 0.5 points; a wrong answer yields -0.25 points; no answer yields 0 points.*

- A Supercritical regime.
- B Anomalous regime.
- C Subcritical regime.
- D Ultra-Small World regime.

`C`

2. Clustering coefficient ($2.25 = 1 + 0.5 + 0.75$ points)

(a) Define and explain in detail, the global clustering coefficient and the average local clustering coefficient of a graph.
$C_i = \frac{2L_i}{k_i(k_i-1)}$`, which measures the density of the neighbors of the`

vertex. It is the ratio of the actual number of edges between the neighbors of a vertex and the maximum possible number of edges between the neighbors of a vertex. Then $\langle C \rangle$ is the average of the values C_i. $C_\Delta = \frac{3 \cdot \#\text{triangles}}{\#\text{triples}}$, which measures how dense the graph is on a global scale. +0.25 for each definition, +0.25 for each explanation.

(b) Construct, for any N, an example of a graph of N vertices where these coefficients differ.
For example, a double star with $n+2$ vertices, such that vertex 1 and 2 are adjacent and adjacent to all other vertices. Then $C_\Delta = \frac{3n}{2\binom{n}{2}+2n} \approx 1/n$. However, C_i is 1 for vertices $i \geq 3$ and $\approx 2/n$ for $i = 1, 2$. Hence, $\langle C \rangle \approx 1$. +0.25 for a single example, +0.25 for an example that extends to any N.

(c) Consider a graph where each vertex has degree 2 and the average local clustering coefficient is 1. What does this graph look like? Explain your answer.
It is a disjoint union of triangles. Since each vertex has degree 2, the graph is a disjoint union of cycles. Since the local clustering coefficient is 1, all edges between the neighbors are present, so each cycle must be a triangle. +0.5 for the answer, +0.25 for the explanation.

3. Configuration model ($2.25 = 0.5 + 0.75 + 0.75$ points)

Consider a graph generated by the configuration model.

(a) State the Molloy-Reed criterion for the existence of a giant component in a graph.
$\frac{\langle k^2 \rangle}{\langle k \rangle} > 2$. +0.25 for the fraction, +0.25 for the > 2.

(b) Suppose that the degree distribution p_k satisfies $p_0 = 0$ and $p_k = 0$ for $k > 3$; that is, only p_1, p_2, p_3 are potentially bigger than 0. Argue that the graph has a giant component if $p_1 < 3p_3$.
$\langle k^2 \rangle = p_1 + 4p_2 + 9p_3$. $\langle k \rangle = p_1 + 2p_2 + 3p_3$. $\frac{\langle k^2 \rangle}{\langle k \rangle} > 2$ holds if and only if $p_1 < 3p_3$. +0.25 for applying Molloy-Reed, +0.25 for $\langle k^2 \rangle$ and +0.25 for $\langle k \rangle$.

(c) Suppose the degree sequence used to generate the configuration model graph specifies degree k_i for vertex i. Derive a simple expression for the expected number of common neighbors of two vertices i and j.

$\sum_{v\neq i,j} \frac{k_i k_v}{2m}\frac{k_j k_v}{2m} = \frac{k_i k_j}{4m^2}\sum_{v\neq i,j} k_v^2$. +0.25 for the probability of an edge, +0.25 for the sum, +0.25 for the simplification.

4. Network growth ($1.75 = 0.5 + 0.5 + 0.75$ **points**)

Preferential attachment is a model for network growth. A vertex joining at time i has degree $k_i(t)$ at time t. Recall that a new vertex joining at time t connects to m existing vertices; it connects to the vertex that joined at time $i < t$ with probability proportional to $k_i(t)$.

(a) Give the formula for $k_i(t)$, the degree of vertex i at time t, as a function of t.
$m(\frac{t}{i})^\beta$. Here β is the dynamical exponent, which indicates the growth of the degrees and is generally equal to $1/2$. +0.25 for the formula, +0.25 for describing β.

(b) Suppose the new vertex connects to an existing vertex with probability proportional to $(k_i(t))^3$. Describe the expected structure of the graph.
This is superlinear preferential attachment. Almost all vertices will connect to a few super-hubs. Hence, the graph becomes a hub-and-spoke graph. +0.25 for hub-and-spoke, +0.25 for the explanation.

(c) Name three other models of network growth discussed in class.
Bianconi-Barabasi, double preferential attachment, random attachment, accelerated growth. +0.25 for each correct model.

5. Robustness ($1.75 = 0.5 + 0.5 + 0.75$ **points**)

(a) Consider a very large scale-free graph with power-law exponent $2 < \gamma < 3$. An attacker disables 80% of the vertices of the graph at random. Does this fragment the graph into many small components? Explain your answer.
No. The critical threshold is $1-\frac{1}{\frac{\langle k^2\rangle}{\langle k\rangle}-1}$. Since the second moment diverges for these values of γ, the critical threshold tends to 1 and a giant component remains. +0.25 for the answer, +0.25 for the explanation using moments.

(b) Consider a very large scale-free graph with a large power-law exponent γ. An attacker disables the 80% *highest-degree* vertices of the graph. Does this fragment the graph into many small components? Explain your answer.

`No. For large` γ`, the graph behaves like a random graph and random failures and targeted attacks become indistinguishable. Hence, the critical threshold is again close to` 1`. +0.25 for the answer, +0.25 for the argument.`

(c) Describe the structure of graphs that have the highest tolerance against both random and targeted attacks (i.e. the removal of high-degree vertices).

`Bimodal graphs, where a small` r `fraction of vertices have the same large degree and a large` $1-r$ `fraction have the same small degree. The small degree vertices make sure the graph remains connected under targeted attacks, while a random attack is unlikely to remove the high-degree vertex. +0.5 for the answer bimodal, +0.25 for the explanation.`