14.15
NETWORKS
PROBLEM SET 3

JOHN WANG

Collaborators: Ryan Liu, Bonny Jain

1. Problem 1

1.1. **Problem 1.a. Problem:** Let A_1 denote the event that node 1 has at least $l \in \mathrm{Z}^+$ neighbhors. Do we observe a phase transition for this event? If so, finnd the threshold function and explain your reasoning.

Solution: Yes, we do observe a phase transition when $t(n) = \frac{l}{n-1}$. To show that there exists a phase transition, lut us examine the case when $\frac{p(n)}{t(n)} = \frac{p(n)(n-1)}{l} \to 0$. In this case we can use Markov's inequality. We define X as a random variable of the number of neighbors of node 1 and we note that $P(A_1) = P(X \geq l)$. Thus, we want to obtain the expected value of X. However, we know that a node has an expected $(n-1)p(n)$ neighbors, because each neighbor has a $p(n)$ probability of being connected by an edge, and there are $n-1$ neighbors. Thus, we see that $E[A_1] = (n-1)p(n)$. Markov's inequality allows us to obtain the following bound:

$$P(A_1) = P(X \geq l) \leq \frac{E[A_1]}{l} = \frac{p(n)(n-1)}{l} \tag{1}$$

However, since $\frac{p(n)}{t(n)} = \frac{p(n)(n-1)}{l} \to 0$, we see that $P(A_1) \to 0$ as well. Thus, we have shown that the first part of the phase transition, namely that under the threshold, A_1 does not occur. Now we shall show that above the threshold, the event occurs almost surely. We want to see what happens when $\frac{p(n)}{t(n)} = \frac{p(n)(n-1)}{l} \to \infty$. We shall use Chebyshev's inequality:

$$\begin{align} P(|X - E[X]| \geq |E[X] - l|) &\leq \frac{\mathrm{Var}(X)}{(E[x] - l)^2} \tag{2} \\ &= \frac{p(n-1)}{((n-1)p - l)^2} \tag{3} \\ &= \frac{p(n-1)}{(n-1)^2 pl^2 + l^2 - 2(n-1)pl} \tag{4} \\ &= \frac{1}{(n-1)l^2 + \frac{l^2}{(n-1)p} - 2l} \tag{5} \end{align}$$

Where we know that $\mathrm{Var}(X) = E[X] = (n-1)p$ because X can be approximated as a poisson random variable (see Newman p. 402) for large n. Moreover, we know that l is a constant so that as $\frac{p(n)(n-1)}{l} \to \infty$ we have $\frac{l^2}{(n-1)p} \to 0$. Therefore, we also see that $P(|X - E[X]| \geq |E[X] - l|) \leq \frac{1}{(n-1)l^2 - 2l}$ as $\frac{p(n-1)}{l} \to \infty$. Since l is a constant we see that $n \to \infty$ implies that this probability goes to zero. In other words, the probability that X deviates by more than $E[X] - l$ from its expected value goes to zero. This shows that $P(A_1) \to 1$ when $E[X] \geq l$. $\square$

1.2. **Problem 1.b. Problem:** Let B denote the event that a cycle with k edges (for a fixed k) emerges in the graph. Do we observe a phase transition of this event? If so, find the threshold function and explain your reasoning.

Solution: Yes a threshold function does exist for this event. Take $t(n) = \frac{1}{n}$ as the threshold function. We will first show that as $\frac{p}{t(n)} = pn \to 0$, then $P(B) \to 0$. We first want to find the expected number of cycles of length k in the graph. Let X denote the number of cycles of length k. We know that there are $\binom{n}{k}$ ways to select k nodes, and that for each of these subsets of nodes, there are $(k-1)!/2$ ways to create a cycle. This follows because we set the first node of the cycle, then there are $k-1$ ways of picking the second node, $k-2$ ways of picking the third node, etc. We divide by two because we could go either backwards or forwards in this cycle (clockwise or counterclockwise). Each of these cycles has probability p^k of emerging. Therefore, we have:

$$E[X] = \binom{n}{k}\frac{(k-1)!}{2}p^k \tag{6}$$

$$= \frac{n!}{k!(n-k)!}\frac{(k-1)!}{2}p^k \tag{7}$$

$$= \frac{n!}{2k(n-k)!}p^k \tag{8}$$

However, we know that $\frac{n!}{(n-k)!} = n(n-1)\dots(n-k+1) \leq n^k$. This means that we have $E[X] \leq \frac{(np)^k}{2k}$. Thus, as $np \to 0$, we see that $E[X] \to 0$, which implies that $P(X \geq 1) \leq \frac{E[X]}{1} = E[X] \to 0$. Since $P(X \geq 1) = P(B)$, we see that $P(B) \to 0$ as $\frac{p}{t(n)} \to 0$. Now to show the second half of the phase transition, we need to show that as $\frac{p}{t(n)} \to \infty$ we have $P(B) \to 1$.

To show this we note that $P(X \leq 0) = P(E[X] - X \geq E[X]) \leq \frac{\text{Var}(X)}{E[X]^2}$ by Chebyshev. We can also bound $E[X]$ from below by using the fact that $\frac{n!}{(n-k)!} = n(n-1)\dots(n-k+1) \geq (n-k)^k$. This shows that $E[X] \geq \frac{(n-k)^k p^k}{2k}$. Since k is a constant we know that $nk - kp \to \infty$ when $np \to \infty$. Thus, we see that $(n-k) \to \infty$, which implies that $E[X] \geq \infty$ when $np \to \infty$.

Now we use the fact that X can be approximated as a poission distribution to note the fact that $\text{Var}(X) = E[X]$. This means that $P(X \leq 0) \leq \frac{1}{E[X]} \to 0$. This implies that $P(X \geq 0) \to \infty$, which means that $P(B) \to \infty$, just as we wanted. □

2. Problem 2

2.1. Problem 2.a.

Problem: Show that the mean degree of a vertex in this network is $2c$.

Solution: We know that the expected number of connected trios is the total possible number of triples of nodes times the probability each triple becomes a triangle. This is $\binom{n}{3}\frac{c}{\binom{n-1}{2}} = \frac{n!}{3!(n-3)!}\frac{c}{(n-1)!}(n-1-2)!2! = \frac{2!nc}{3!} = \frac{nc}{3}$. Since we know that the total number of edges is just 3 times the number of triangles we have nc total expected edges. Moreover, the expected total degree is 2 times the total number of expected edges (since each edge has two endpoints). This means we expect $2nc$ total degree in the graph, and since there are n nodes, each node has an expected degree of $\frac{2nc}{n} = 2c$. □

2.2. Problem 2.b.

Problem: Show that the degree distribution is

$$p_k = \left\{ \begin{array}{ll} e^{-c}c^{k/2}/(k/2)! & \text{if } k \text{ is even} \\ 0 & \text{if } k \text{ is odd} \end{array} \right\} \tag{9}$$

Solution: We can assume that the degree distribution is a poisson random variable because of the fact that each triangle is selected randomly with a particular probability. This is a binomial distribution, which converges to a poission distribution as n grows large. Therefore, we only need to find the expected degree of a node. The probability that a node has degree k is really equal to the probability that it is inside of $k/2$ triangles (because each triangle provide two edges). The expected value is given by the number of triangles that a node can connect to times the probability of occuring. This is just $\binom{n-1}{2}\frac{c}{\binom{n-1}{2}}$.

Therefore, we see that $\lambda = c$ in this poisson distribution and that there are $m = k/2$ different triangle possibilities. We therefore have the probability distribution of a possion random variable $\frac{\lambda^m e^{-\lambda}}{m!}$ which we can substitute $\lambda = c$ and $m = k/2$ to obtain $\frac{c^{k/2}e^{-c}}{(k/2)!}$. This is the degree distribution for when k is even, because k cannot be odd. This is because whenever a new triangle is added, 2 more edges are added to each node, and thus, one cannot have an odd degree. Therefore, we have shown that the degree distribution is given by:

$$p_k = \left\{ \begin{array}{ll} e^{-c}c^{k/2}/(k/2)! & \text{if } k \text{ is even} \\ 0 & \text{if } k \text{ is odd} \end{array} \right\} \tag{10}$$

□

2.3. Problem 2.c.

Problem: Show that the clustering coefficient is $C = \frac{1}{2c+1}$.

Solution: □

2.4. **Problem 2.d. Problem:** Show that when there is a giant component in the network, its expected size S, as a fraction of the network size, satisfies $S = 1 - e^{-cS(2-S)}$.

Solution: Suppose that u_i is the probability that node i does not belong to the giant component. For i to not belong in the giant component we must either have 1) that i is not connected to some triangle j or 2) that i is connected to the triangle j but j itself is not part of the giant component. The first occurs with probability $1-p$, while the second occurs with probability pu_j^2, since both nodes of the connecting triangle cannot belong to the giant component. This means the probaiblity that i is not connected to the giant component through traingle j is $1 - p + pu_j^2$.

Since each node is independent and triangles are randomly selected, we can say $u_i = u_j = u$. Since there are $\binom{n-1}{2}$ triangles for which we must do this analysis for, we see the following:

$$u = (1 - p + pu^2)^{\binom{n-1}{2}} \tag{11}$$

$$= \left(1 - \frac{c}{\binom{n-1}{2}} + \frac{c}{\binom{n-1}{2}}u^2\right)^{\binom{n-1}{2}} \tag{12}$$

Taking logs of both sides we see that we can simplify the expression:

$$\ln u = \binom{n-1}{2} \ln\left(1 - \frac{c}{\binom{n-1}{2}} + \frac{c}{\binom{n-1}{2}}u^2\right) \tag{13}$$

$$= \binom{n-1}{2} \frac{cu^2 - u}{\binom{n-1}{2}} \tag{14}$$

$$= cu^2 - u \tag{15}$$

Where we have used the fact that $ln(1 + x) = x$ for small x. Exponentiating both sides of the resulting expression gives us that $u = e^{cu^2 - u}$. However, we know that $S = 1 - u$ because u is the probability of any particular node not belong to the giant component, and S is the probability of belonging to the giant component. Therefore, we see that $1 - S = e^{-c+c(1-S)^2}$ which simplifies to $S = 1 - e^{-cS(2-s)}$, which proves our result. $\square$

2.5. **Problem 2.e. Problem:** What is the value of the clustering coefficient when the giant component fills half of the network?

Solution: We choose $S = 1/2$ and we solve for c. We have:

$$\frac{1}{2} = 1 - e^{\frac{1}{2}c\left(\frac{1}{2}-2\right)} \tag{16}$$

$$\frac{1}{2} = 1 - e^{-\frac{3}{4}c} \tag{17}$$

$$e^{-\frac{3}{4}c} = \frac{1}{2} \tag{18}$$

$$-\frac{3}{4} = -\ln(2) \tag{19}$$

$$c = \frac{4}{3}\ln(2) \tag{20}$$

Now, we substitute into our equation for the clustering coefficient of $C = \frac{1}{2c+1}$ so we obtain $C = \frac{1}{8\ln(2)/3+1}$. $\square$

3. Problem 3

Problem:
Solution: $\square$ **Problem:**
Solution: $\square$ **Problem:**
Solution: $\square$ **Problem:**
Solution: $\square$ **Problem:**
Solution: $\square$