COMP4641: Social Information Networks Analysis and Engineering 2020 Spring Semester Assignment 1

Due time: 23:59pm, Apr 01 (Wed), 2020.

Problem 1 (40%)

(a) $\bar{k} = (n-1)p$

(b)

1) True. $m = \bar{k}$ implies that every nodes in both models are expected to connect same number of neighbours (lets say X). For static geographic model, all the nodes connect the m nearest neighbours. But nodes in Erdos-Renyi model connect neighbours randomly. Therefore static geographic model is expected to have a higher clustering coefficient, that is stronger locality.

2) False. Static geographic model connect the nearest neighbours, this mean that the shortest path inside the component is short, but it is long outside the component. However, Erdos-Renyi model will connect the whole graph randomly. That mean the distances are normally distributed, node can go to another node with a average shorter distance. In conclusion, Erdos-Renyi model is expected to have a shorter average shortest path.

(c) By definition, the more the number of nodes, the lower the probability for a new node to connect other nodes. So the decreasing probability make the links unevenly distributed.

Erdos-Renyi model: $P(deg = 1) = \dfrac{4(N-1)}{2^N}$ with N nodes and $p = 0.5$.

Random growth model: $P(deg = 1) = \displaystyle\sum_{n=1}^{N} \frac{1}{n}$ with N total nodes.

Therefore random growth model has more nodes with degree 1.

(d) Model 1: preferential attachment model; Model 2: random growth model.

For preferential attachment model, a new node is more likely to connect a node with higher degree $\dfrac{k_i}{\sum_j k_j}$. For random growth model, a new node to all other nodes with equal probability $\dfrac{1}{\sum_j 1}$. Since rich get richer, preferential attachment model will present some large node, while random growth model will present all the nodes with similar size.

(e) $$\int dK^{t_0}(t) = \int \frac{m}{t} dt$$

$$K^{t_0}(t) = mln(t) + C$$

at $t = t_0$:

$$K^{t_0}(t_0) = mln(t) + C = m$$

$$C = m - mln(t_0)$$

$$\int dK^{t_0}(t) = \int \frac{m}{t} dt$$

$$\begin{aligned} K^{t_0}(t) &= mln(t) + C \\ &= mln(t) + m - mln(t_0) \\ &= m(1 + ln(t) - ln(t_0)) \\ &= m(1 + ln(\frac{t}{t_0})) \end{aligned}$$

(f) The probability for new node connecting node $i = \frac{k_i}{\sum_j k_j}$, that means the new nodes are more likely to connect to some "large node"(high degree). At the same time, total degree of the graph is increasing, so the probability of connecting node with small degree is dropping. Just as the formula describe, when t is getting larger, the probability of smaller than $K^{t_0}(t)$ is also getting larger.

$P(k)$ should perform similar to Figure 1(a).

Problem 2 (30%)

(a) normalised degree centrality for node 2: $\frac{k_i}{n-1} = \frac{3}{6} = \frac{1}{2}$

normalised betweenness centrality node 4: $\frac{9}{(6)(5)/2} = \frac{3}{5}$

normalised closeness centrality node 4: $\left[\frac{1+1+1+1+2+2}{6}\right]^{-1} = \frac{6}{8} = \frac{3}{4}$

clustering coefficient for node 4: $\frac{2*2}{4*3} = \frac{1}{3}$

P	1	2	3	4	5	6	7
1		1/2		1/2			
2	1/3		1/3	1/3			
3		1/2		1/2			
4	1/4	1/4	1/4		1/4		
5				1/3		1/3	1/3
6					1/2		1/2
7					1/2	1/2	

(b)

PP	1	2	3	4	5	6	7
1		1/8	1/6 + 1/8	1/6	1/8		
2	1/12		1/12	1/6	1/12		
3	1/6 + 1/8	1/8		1/6	1/8		
4	1/12	1/8 + 1/8	1/12			1/12	1/12
5	1/12	1/12	1/12			1/6	1/6
6				1/6	1/4		1/6
7				1/6	1/4	1/6	

$$c_{54} = (\frac{1}{3} + 0)^2 = \frac{1}{9}$$

$$c_{56} = (\frac{1}{3} + \frac{1}{6})^2 = \frac{1}{4}$$

$$c_{57} = (\frac{1}{3} + \frac{1}{6})^2 = \frac{1}{4}$$

$$\begin{aligned} c_5 &= \frac{1}{9} + \frac{1+1}{4} \\ &= \frac{16}{144} + \frac{72}{144} \\ &= \frac{88}{144} \\ &= \frac{11}{18} \end{aligned}$$

(c)

```
[-0.22222222  0.66666667  0.55555556 -0.22222222 -0.33333333 -0.22222222  -0.22222222]
[ 0.66666667 -0.5         0.33333333  0.66666667 -0.5        -0.33333333  -0.33333333]
[-0.22222222  0.66666667  0.55555556 -0.22222222 -0.33333333 -0.22222222  -0.22222222]
[ 0.55555556  0.33333333 -0.88888889  0.55555556  0.33333333 -0.44444444  -0.44444444]
[-0.33333333 -0.5         0.33333333 -0.33333333 -0.5         0.66666667   0.66666667]
[-0.22222222 -0.33333333 -0.44444444 -0.22222222  0.66666667 -0.22222222   0.77777778]
[-0.22222222 -0.33333333 -0.44444444 -0.22222222  0.66666667  0.77777778  -0.22222222]
```