ECE 232E Lecture 5 and Lecture 6 notes

Professor Vwani Roychowdhury

May 1, 2018

In this set of lecture notes, we will study the dynamics of random networks created using various types of preferential attachment protocol. In order to study the dynamics we will need some differential equation background which we provide at the beginning of the notes.

1 Differential equation prelude

Suppose the dynamics of $x(t)$ is described the differential equation 1

$$\frac{dx(t)}{dt} = \alpha \frac{x(t)}{t} \tag{1}$$

with initial condition $x(i) = m$. The solution to the differential equation 1 is given by

$$x(t) = ct^{\alpha}$$

where c is a constant. We can use the initial condition to find c,

$$ci^{\alpha} = m$$
$$\Rightarrow c = \frac{m}{i^{\alpha}}$$

Therefore, the complete solution to the differential equation 1 is given by

$$x(t) = m(\frac{t}{i})^{\alpha} \tag{2}$$

We will use the result given by equation 2 to study the dynamics of random networks created using various types of preferential attachment (PA) protocol.

2 Dynamics of random networks created using PA protocol

In this section, we will derive expressions for tracking the degree distribution over time for a random network created using PA protocol. To facilitate the derivation, let's introduce some notation:

- $k(i,t)$ is the average degree of the i^{th} node (the node added at the i^{th} time step) after time step t
- m is the number of edges added at each time step
- $k_{max}(t)$ is the maximum degree after time step t
- $P_k(t)$ is the probability that a randomly picked node has a degree k after time step t
- $\overline{\mathcal{N}_k(t)}$ is the average number of nodes with degree k after time step t

We want to derive an expression for $k(i,t)$. In order to do so, let's write the one step forward difference equation

$$k(i,t+1) = k(i,t) + mf_{t+1} \tag{3}$$

where f_{t+1} is the probability that the node added at the i^{th} time step forms an edge with the node added at the $(t+1)^{th}$ time step. The expression for f_{t+1} is given below

$$f_{t+1} = \frac{k(i,t)}{2mt} \tag{4}$$

Plugging in the above expression into equation 3 and rearranging we get,

$$\begin{aligned} k(i,t+1) - k(i,t) &= \frac{1}{2} \times \frac{k(i,t)}{t} \\ \Rightarrow \frac{k(i,t+1) - k(i,t)}{(t+1) - t} &= \frac{1}{2} \times \frac{k(i,t)}{t} \end{aligned}$$

Now for large enough value of t,

$$\lim_{t \to \infty} \frac{k(i,t+1) - k(i,t)}{(t+1) - t} = \frac{\partial k(i,t)}{\partial t}$$

Therefore, we have the following partial differential equation (PDE),

$$\frac{\partial k(i,t)}{\partial t} = \frac{1}{2} \times \frac{k(i,t)}{t} \tag{5}$$

Using the result of equation 2, we can write the solution to the PDE given by equation 5 as

$$k(i,t) = m\left(\frac{t}{i}\right)^{\frac{1}{2}} \tag{6}$$

The expression given by equation 6 tracks the average degree of a node added at the i^{th} time step over time. For large enough t, the plot of $k(i,t)$ against i is shown below

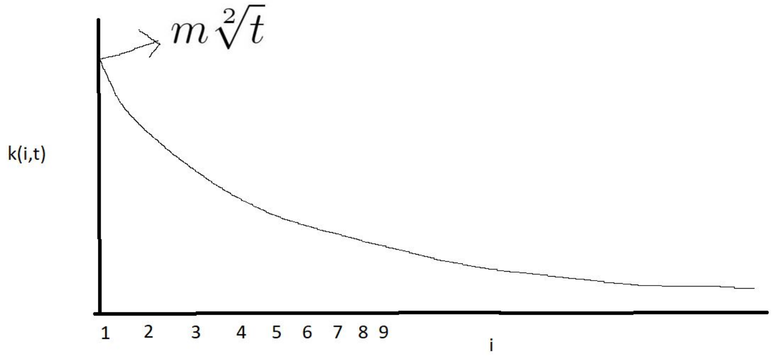

Figure 1: Plot of $k(i,t)$ against i

From the above plot, it can be observed that on average older nodes will have higher average degree and $k(i,t)$ increases monotonically with $\frac{1}{i}$. Also, we have

$$k_{max}(t) = m\sqrt[2]{t} \tag{7}$$

Having derived an expression for $k(i,t)$, we will now use that expression to derive an expression for $P_k(t)$. From the previous lecture, we have that

$$P_k(t) = \frac{\overline{\mathcal{N}_k(t)}}{t} \tag{8}$$

Now we can approximate $\overline{\mathcal{N}_k(t)}$ by finding the number of i for which $k \leq k(i,t) < k+1$. We can find the number of i using the following expression,

$$\overline{\mathcal{N}_k(t)} \approx \left| \frac{\partial k(i,t)}{\partial i}\bigg|_{i=i_k} \right|^{-1} \tag{9}$$

Now, using the expression for $k(i,t)$, we have

$$\left| \frac{\partial k(i,t)}{\partial i} \right| = mt^{\beta}\beta i^{-(\beta+1)}; \quad \beta = 0.5 \tag{10}$$

Also, the expression for i_k is given by

$$i_k = (\frac{k}{m})^{\frac{-1}{\beta}} t \tag{11}$$

Combining equations 10 and 11 we get,

$$\left| \frac{\partial k(i,t)}{\partial i}\bigg|_{i=i_k} \right| = m^{\frac{-1}{\beta}} \beta t^{-1} (\frac{k}{m})^{1+\frac{1}{\beta}} \tag{12}$$

Therefore,

$$\overline{\mathcal{N}_k(t)} \approx m^{\frac{1}{\beta}} \beta^{-1} t (\frac{k}{m})^{-1-\frac{1}{\beta}} \tag{13}$$

Plugging in the above approximation into equation 8 and simplifying, we get

$$P_k(t) = c \times \frac{1}{k^{1+\frac{1}{\beta}}} \tag{14}$$

where c is approximation constant. For random networks created using PA we have $\beta = 0.5$ and therefore we have

$$P_k \propto \frac{1}{k^3} \quad \text{PA networks} \tag{15}$$

The above result derived using the dynamics of the network matches with the result derived in the previous lecture. Now for PA networks the value of the exponent is 3 and hence the degree distribution is not very fat tailed. In order to get a more fat tailed distribution we need to make the value of the exponent closer to 2. To be concrete, we want to come up with a protocol that further amplifies the degrees of high-degree nodes. This amplification can be achieved using the Doubly Preferential Attachment protocol which we analyze in the next section.

3 Doubly Preferential Attachment

The protocol for creating a network using doubly preferential attachment is given below:

At i^{th} time step,

1. A node joins the network
2. The node makes m preferential edges with the existing nodes in the network
3. One node is picked preferentially and it makes b preferential edges

Having described the protocol, we can now analyze the dynamics of the networks created using doubly preferential attachment protocol. Using the same notation from before, we will first derive an expression for $k(i, t)$. In order to do so, let's write the one step forward difference equation

$$k(i, t+1) = k(i, t) + m f_{t+1} + b r_{t+1} + b g_{t+1}(1 - r_{t+1}) \tag{16}$$

where f_{t+1} is the probability that the node added at the i^{th} time step forms an edge with the node added at the $(t+1)^{th}$ time step (step 2 of the protocol), r_{t+1} is the probability that the node added at the i^{th} time step is picked to make b preferential edges (step 3 of the protocol), and g_{t+1} is the probability that the node added at the i^{th} time step forms an edge with the node that has been picked preferentially in step 3 of the protocol. The expressions for f_{t+1}, r_{t+1}, and g_{t+1} are given below:

$$f_{t+1} = r_{t+1} = g_{t+1} = \frac{k(i, t)}{2(m + b)t} \tag{17}$$

Plugging in the expressions into equation 16 and assuming $\frac{k(i,t)}{2(m+b)t} <<< 1$, we get

$$\begin{aligned} k(i,t+1) - k(i,t) &= \frac{m+2b}{2(m+b)} \frac{k(i,t)}{t} \\ \Rightarrow \frac{k(i,t+1) - k(i,t)}{(t+1) - t} &= \frac{m+2b}{2(m+b)} \frac{k(i,t)}{t} \end{aligned}$$

Now for large enough value of t,

$$\lim_{t \to \infty} \frac{k(i,t+1) - k(i,t)}{(t+1) - t} = \frac{\partial k(i,t)}{\partial t}$$

Therefore, we have the following partial differential equation (PDE)

$$\frac{\partial k(i,t)}{\partial t} = \frac{m+2b}{2(m+b)} \frac{k(i,t)}{t} \tag{18}$$

If we denote $\alpha = \frac{m+2b}{2(m+b)}$, and then using the result of equation 2 we can write the solution to the PDE given by equation 18 as

$$k(i,t) = m \left(\frac{t}{i} \right)^{\alpha} \tag{19}$$

Then, the expression for the exponent γ is given by

$$\gamma = 1 + \frac{1}{\alpha} \tag{20}$$

By selecting the value of b, we can control the value of the exponent γ, which in turn effects the heavy tail property of the distribution. To be specific, the closer the value of the exponent γ is to 2 the more heavy tailed the distribution is.

- If $b = 0$, then $\gamma = 3$
- If $b = m$, then $\gamma = 2.33$
- If $b = 2m$, then $\gamma = 2.2$

4 Preferential Attachment with Deletion

The protocol for creating a network using preferential attachment with deletion is given below:

At i^{th} time step,

1. A node joins the network and makes m preferential edges with the existing nodes in the network
2. An existing node is picked at random and is deleted with probability c

Since nodes are deleted randomly in this protocol, the average number of deleted edges at time t is a function of the average degree of a node at time t; this implies that we need to solve for the expected number of edges as a function of time step t. To facilitate the derivation, let's introduce some notation:

- $\mathcal{N}(t)$ is the expected number of nodes in the network after time step t

$$\mathcal{N}(t) = (1-c)t$$

- $E(t)$ is the expected number of edges in the network after time step t

With the above notation, the one step forward difference equation is given by

$$E(t+1) = E(t) + m - z_{t+1} \tag{21}$$

where z_{t+1} is the number of edges deleted at time step $t+1$. The expression for z_{t+1} is

$$\begin{aligned} z_{t+1} &= c\frac{2E(t)}{\mathcal{N}(t)} \\ z_{t+1} &= \frac{2c}{(1-c)t}E(t) \end{aligned} \tag{22}$$

Plugging in the expression for z_{t+1} into equation 21 and rearranging it, we get

$$\begin{aligned} E(t+1) - E(t) &= m - \frac{2c}{(1-c)t}E(t) \\ \Rightarrow \frac{E(t+1) - E(t)}{(t+1) - t} &= m - \frac{2c}{(1-c)t}E(t) \end{aligned}$$

Now for large enough value of t,

$$\lim_{t\to\infty} \frac{E(t+1) - E(t)}{(t+1) - t} = \frac{dE(t)}{dt}$$

Therefore, we have the differential equation,

$$\frac{dE(t)}{dt} = m - \frac{2c}{(1-c)}\frac{E(t)}{t} \tag{23}$$

The solution to the differential equation given by 23 is

$$E(t) = m\left(\frac{1-c}{1+c}\right)t \tag{24}$$

You can verify that expression given by equation 24 is a solution to the differential equation.

We will now derive an expression for $D(i,t)$, the probability that the i^{th} node is still alive after time step t. The one step forward difference equation is given by

$$D(i,t+1) = D(i,t)a_{t+1} \tag{25}$$

where a_{t+1} is the probability that the i^{th} node is not deleted at time step $t+1$. The expression for a_{t+1} is given by

$$a_{t+1} = 1 - \frac{c}{(1-c)t} \tag{26}$$

Plugging in the above expression into equation 25 and rearranging it, we get

$$\begin{aligned} D(i,t+1) - D(i,t) &= -\frac{c}{1-c}\frac{D(i,t)}{t} \\ \Rightarrow \frac{D(i,t+1) - D(i,t)}{(t+1)-t} &= -\frac{c}{1-c}\frac{D(i,t)}{t} \end{aligned}$$

Now for large enough value of t,

$$\lim_{t\to\infty} \frac{D(i,t+1) - D(i,t)}{(t+1)-t} = \frac{\partial D(i,t)}{\partial t}$$

Therefore, we have the partial differential equation (PDE),

$$\frac{\partial D(i,t)}{\partial t} = -\frac{c}{1-c}\frac{D(i,t)}{t} \tag{27}$$

Assuming $D(i,i) = 1$, the solution to the PDE is given by

$$D(i,t) = \left(\frac{t}{i}\right)^{-\frac{c}{1-c}} \tag{28}$$

We will now derive an expression for $k(i,t)$, the average degree of the i^{th} node after time step t (assuming it is still alive). The one step forward difference equation is given by

$$k(i,t+1) = k(i,t) + m\frac{k(i,t)}{2E(t)} - n_{t+1} \tag{29}$$

where n_{t+1} is the probability that a neighbor of the node is deleted at the $(t+1)^{th}$ time step. The expression for n_{t+1} is given by

$$n_{t+1} = c\frac{k(i,t)}{\mathcal{N}(t)} \tag{30}$$

Plugging in the above expression into equation 29 and rearranging it, we get

$$\begin{aligned} k(i,t+1) - k(i,t) &= m\frac{k(i,t)}{2E(t)} - c\frac{k(i,t)}{\mathcal{N}(t)} \\ \Rightarrow \frac{k(i,t+1) - k(i,t)}{(t+1)-t} &= m\frac{k(i,t)}{2E(t)} - c\frac{k(i,t)}{\mathcal{N}(t)} \end{aligned}$$

Now for large enough value of t,

$$\lim_{t \to \infty} \frac{k(i,t+1) - k(i,t)}{(t+1) - t} = \frac{\partial k(i,t)}{\partial t}$$

Plugging in the expression for $E(t)$ and simplifying, we get the partial differential equation (PDE),

$$\frac{\partial k(i,t)}{\partial t} = \frac{1}{2} \times \frac{k(i,t)}{t} \tag{31}$$

The solution to the PDE is given by

$$k(i,t) = m\left(\frac{t}{i}\right)^{\frac{1}{2}} \tag{32}$$

Now, the expression for $P_k(t)$ is given by

$$P_k(t) = \frac{(\#\text{ of } i\text{'s such that average degree is equal to } k)(\text{Probability that the nodes are still alive})}{\mathcal{N}(t)} \tag{33}$$

Using the above expression, and performing a similar calculation to that in section 2, it can be shown that

$$P_k \propto \frac{1}{k^\gamma} \tag{34}$$

where $\gamma = 1 + (\frac{1}{\beta}\frac{1}{1-c})$, and where we have assumed that in general,

$$k(i,t) = m\left(\frac{t}{i}\right)^{\frac{1}{\beta}} . \tag{35}$$

Since $\beta = 1/2$ the value of the exponent, $\gamma = 1 + (\frac{1}{2}\frac{1}{1-c})$, approaches ∞ as $c \longrightarrow 1$, destroying any heavy tail in the degree distribution:

- For $c = 0.5$, we get $\gamma = 5$
- For $c = 0.25$, we get $\gamma = 3.67$

Since $0 < c < 1$, $\gamma > 3$ and so the networks generated using PA with deletion does not have a very heavy tail.

5 Deletion-Compensation protocol

The protocol for creating a network using deletion-compensation protocol is given below:

At i^{th} time step,

1. A new node joins the network and makes m preferential edges with the existing nodes in the network
2. A node picked at random, leaves the network with probability c (deletion step)
3. If a node loses it's neighbor then with probability n it makes a new preferential edge with one of the existing nodes (compensation step)

Using the above protocol, the expressions for the number of nodes network after time step t is given by

$$\mathcal{N}(t) = (1-c)t \tag{36}$$

In order to get the expression for the number of edges in the network after time step t, we can write the one step forward difference equation

$$E(t+1) = E(t) + m - c(1-n)\frac{2E(t)}{\mathcal{N}(t)} \tag{37}$$

Now we can translate the above equation into a partial differential equation using a similar approach as before and solve the partial differential equation to get

$$E(t) = \frac{m(1-c)}{1+c-nc}t \tag{38}$$

Similarly, the one step forward difference equation for $k(i,t)$ is given by

$$k(i,t+1) = k(i,t) + \frac{mk(i,t)}{2E(t)} - \frac{ck(i,t)}{\mathcal{N}(t)} + \frac{nck(i,t)}{\mathcal{N}(t)} + \frac{2E(t)}{\mathcal{N}(t)}\frac{nck(i,t)}{2E(t)} \tag{39}$$

Translating the above equation into a partial differential equation using a similar approach as before and solve the PDE to get

$$k(i,t) = m\left(\frac{t}{i}\right)^{\beta} \tag{40}$$

where $\beta = \frac{1-c+2nc}{2(1-c)}$. Using a similar approach as before, it can be shown that

$$P_k \propto \frac{1}{k^{\gamma}} \tag{41}$$

where,

$$\gamma = 1 + \frac{1}{(1-c)\beta}$$
$$\Rightarrow \gamma = 1 + \frac{2}{1-c+2nc} \tag{42}$$

From expression 41, it can be observed that as c approaches 1, exponent approaches $1 + \frac{1}{n}$.