ECE 232E Lecture 3 and Lecture 4 notes

Professor Vwani Roychowdhury

April 18, 2018

In this set of lecture notes, we analyze power law network and generative models for random network.

1 Power law network

The defining feature of a power law network is given by equation 1

$$\lim_{k\to\infty} \mathbb{P}(\text{a randomly chosen node has degree } k) \propto \frac{1}{k^{\gamma}}, \gamma > 0 \tag{1}$$

If we denote,

$$P_k = \mathbb{P}(\text{a randomly chosen node has degree } k)$$

Then the degree distribution of a power law network, plotted in a log-log scale, is shown in figure 1

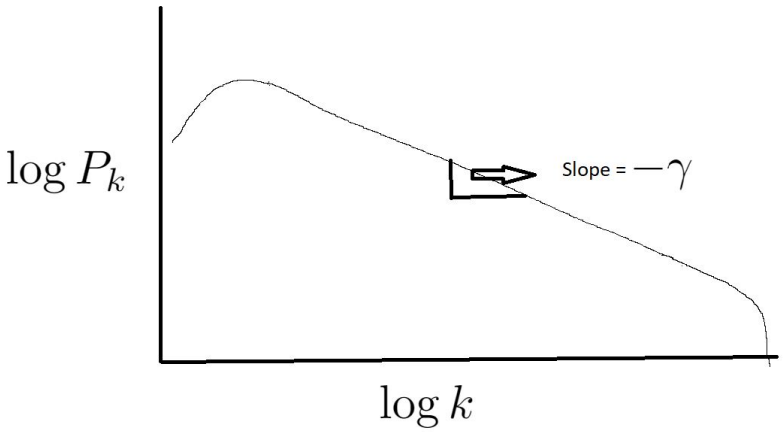

Figure 1: Degree distribution of a power law network plotted on a log-log scale

1.1 Structural properties of a power law network

Having defined a power law network, we will study some of the structural properties of a power law network. We will analyze the following structural properties:

- Degree distribution
- Average degree and variance

1.1.1 Degree distribution

Before stating the expression for the degree distribution, let's introduce some notation:

k_{max} : Maximum degree
$c(k_{max}, \gamma)$: Normalization constant as a function of maximum degree and exponent γ

Using the above notation, the degree distribution of a power law network is given by equation 2

$$P_k = \frac{c(k_{max}, \gamma)}{k^\gamma}, \quad k = 1, 2, 3, \cdots, k_{max} \tag{2}$$

We can compute the value of the normalization constant ($c(k_{max}, \gamma)$) using the law of total probability

$$\sum_{k=1}^{k_{max}} P_k = 1$$

$$\sum_{k=1}^{k_{max}} \frac{c(k_{max}, \gamma)}{k^\gamma} = 1$$

$$c(k_{max}, \gamma) = \frac{1}{\sum_{k=1}^{k_{max}} \frac{1}{k^\gamma}} \tag{3}$$

The normalization constant, given by equation 3, converges to different values depending on the values of k_{max} and γ.

i) For $k_{max} >> 1$ and $\gamma < 1$, we have

$$c(k_{max}, \gamma) \to 0$$

Therefore, the degree distribution is ill-defined for $\gamma < 1$.

ii) For $k_{max} >> 1$ and $\gamma = 1$, we have

$$c(k_{max}, 1) = \frac{1}{\sum_{k=1}^{k_{max}} \frac{1}{k}}$$

We can derive an approximation to the above equation using some bounds on the summation. First let's derive a lower bound on the summation

$$\sum_{k=1}^{k_{max}} \frac{1}{k} >= \int_1^{k_{max}} \frac{1}{k} dk = ln(k_{max}) \tag{4}$$

Now let's derive an upper bound on the summation,

$$\begin{aligned}\sum_{k=1}^{k_{max}} \frac{1}{k} =& 1 + \sum_{k=2}^{k_{max}} \frac{1}{k} \\ \sum_{k=2}^{k_{max}} \frac{1}{k} =& \sum_{k=1}^{k_{max}} \frac{1}{k} - 1 \leq ln(k_{max}) \\ \sum_{k=1}^{k_{max}} \frac{1}{k} \leq& ln(k_{max}) + 1 \end{aligned} \tag{5}$$

Using the lower and upper bounds, given by equations 4 and 5 respectively

$$\begin{aligned} ln(k_{max}) \leq \sum_{k=1}^{k_{max}} \frac{1}{k} \leq ln(k_{max}) + 1 \\ 0 \leq \sum_{k=1}^{k_{max}} \frac{1}{k} - ln(k_{max}) \leq 1 \end{aligned}$$

Now, using the result given by equation 6 (can be derived using numerical analysis)

$$\lim_{k_{max} \to \infty} \left(\sum_{k=1}^{k_{max}} \frac{1}{k} - ln(k_{max}) \right) = 0.5772 \tag{6}$$

we have the following approximation for the normalization constant

$$c(k_{max}, 1) \approx \frac{1}{ln(k_{max}) + 0.5772} \tag{7}$$

Therefore, the degree distribution is well-defined for $\gamma = 1$.

iii) For $k_{max} >> 1$ and $\gamma = 2$, we have

$$c(k_{max}, 2) = \frac{1}{\sum_{k=1}^{k_{max}} \frac{1}{k^2}}$$

Using the result given by equation 8,

$$\lim_{k_{max} \to \infty} \sum_{k=1}^{k_{max}} \left(\frac{1}{k^2} \right) = \frac{\pi^2}{6} \tag{8}$$

we get

$$c(k_{max}, 2) = \frac{6}{\pi^2} \tag{9}$$

In general, for $k_{max} >> 1$ and $\gamma > 1$, we have $c(k_{max}, 2) > 0$. Therefore, we also get a well-defined degree distribution for $\gamma > 1$.

1.1.2 Average degree and variance

Before deriving the average degree of a power law network, let's introduce some notation:

$$\begin{aligned} m &: \text{Number of nodes picked} \\ \mathcal{N}_k &: \text{Number of nodes picked with degree} = k \\ P_k &: \text{Probability that a randomly picked node has degree } k \\ \mathbb{E}_m(deg) &: \text{Empirical average degree} \\ \mathbb{E}(deg) &: \text{True average degree} \end{aligned}$$

Then, with the above notation

$$\mathbb{E}_m(deg) = \frac{1}{m} \times \sum_{k=1}^{k_{max}} k \times \mathcal{N}_k$$

$$\mathbb{E}_m(deg) = \sum_{k=1}^{k_{max}} k \times \frac{1}{m}\mathcal{N}_k$$

Now, for large enough picks

$$\lim_{m \to \infty} \frac{1}{m}\mathcal{N}_k = P_k$$

Using the above relation, the expression for true average degree is given by equation 10

$$\mathbb{E}(deg) = \sum_{k=1}^{k_{max}} k \times P_k \tag{10}$$

Plugging in the expression for P_k for a power law network (equation 2) we get

$$\mathbb{E}(deg) = \sum_{k=1}^{k_{max}} \frac{c(k_{max}, \gamma)}{k^{\gamma - 1}} \tag{11}$$

Therefore, the average degree of a power law network is given by equation 11. From equation 11, it can be observed that the average degree is a function of the maximum degree (k_{max}) and the exponent (γ).

i) For $k_{max} >> 1$ and $\gamma < 1$, we have $c(k_{max}, \gamma) = 0$ and the distribution is ill-defined.

ii) For $k_{max} >> 1$ and $1 \leq \gamma \leq 2$, we have

$$\lim_{k_{max} \to \infty} \mathbb{E}(deg) \to \infty \tag{12}$$

iii) For $k_{max} >> 1$ and $\gamma > 2$, we have

$$\lim_{k_{max} \to \infty} \mathbb{E}(deg) < \infty \tag{13}$$

The expression for the second moment of the degree distribution of a power law network is given by equation 14

$$\mathbb{E}(deg^2) = c(k_{max}, \gamma) \sum_{k=1}^{k_{max}} \frac{1}{k^{\gamma-2}} \tag{14}$$

Then using equations 11 and 14, the variance of the degree distribution of a power law network is given by equation 15

$$Var(deg) = \left(c(k_{max}, \gamma) \sum_{k=1}^{k_{max}} \frac{1}{k^{\gamma-2}} \right) - \left(\sum_{k=1}^{k_{max}} \frac{c(k_{max}, \gamma)}{k^{\gamma-1}} \right)^2 \tag{15}$$

From equation 15, it can be observed that the variance is a function of the maximum degree (k_{max}) and the exponent (γ).

i) For $k_{max} >> 1$ and $\gamma < 1$, we have $c(k_{max}, \gamma) = 0$ and the distribution is ill-defined.

ii) For $k_{max} >> 1$ and $1 \leq \gamma \leq 3$, we have

$$\lim_{k_{max} \to \infty} Var(deg) \to \infty \tag{16}$$

iii) For $k_{max} >> 1$ and $\gamma > 3$, we have

$$\lim_{k_{max} \to \infty} Var(deg) < \infty \tag{17}$$

Combining all the results above, we have the following table for $\lim_{k_{max} \to \infty}$

	$\gamma \leq 1$	$1 < \gamma \leq 2$	$2 < \gamma \leq 3$	$\gamma > 3$
$c(k_{max}, \gamma)$	0	> 0	> 0	> 0
$\mathbb{E}(deg)$	ill-defined	∞	$< \infty$	$< \infty$
$Var(deg)$	ill-defined	∞	∞	$< \infty$

If,

$$\lim_{k_{max} \to \infty} \left(\frac{\mathbb{E}(deg)}{\mathbb{E}(deg^2)} \right) = 0$$

then the distribution is called heavy-tailed. This type of distribution implies that there are significant number of nodes with high degree. From the above table, we can conclude that if $2 < \gamma \leq 3$ then we have a heavy-tailed degree distribution. This range of γ gives rise to a sparse network (average degree is bounded) with a significant number of high-degree nodes (unbounded variance).

1.2 Generating a power law network

In this section, we derive the algorithm for generating a power law network. The algorithm can be divided into two steps:

1. Generating the nodes of the network along with their degrees
2. Connecting the nodes of the network by edges

There are two methods to accomplish step 1, which we will describe below:

1.2.1 Method 1

In Method 1, we randomly sample n numbers from the power law degree distribution

$$P_k = \frac{c(k_{max}, \gamma)}{k^\gamma}, \quad \gamma > 1$$

to get the nodes of the network along with their degrees. The value of the i^{th} sample is the degree of the i^{th} node.

In this method, let $\mathcal{N}_k$ denote the number of nodes with degree k. Then $\mathcal{N}_k$ has a binomial distribution given by equation 18

$$P(\mathcal{N}_k = j) = \binom{n}{j} \left(\frac{c(k_{max}, \gamma)}{k^\gamma} \right)^j \left(1 - \frac{c(k_{max}, \gamma)}{k^\gamma} \right)^{n-j} \tag{18}$$

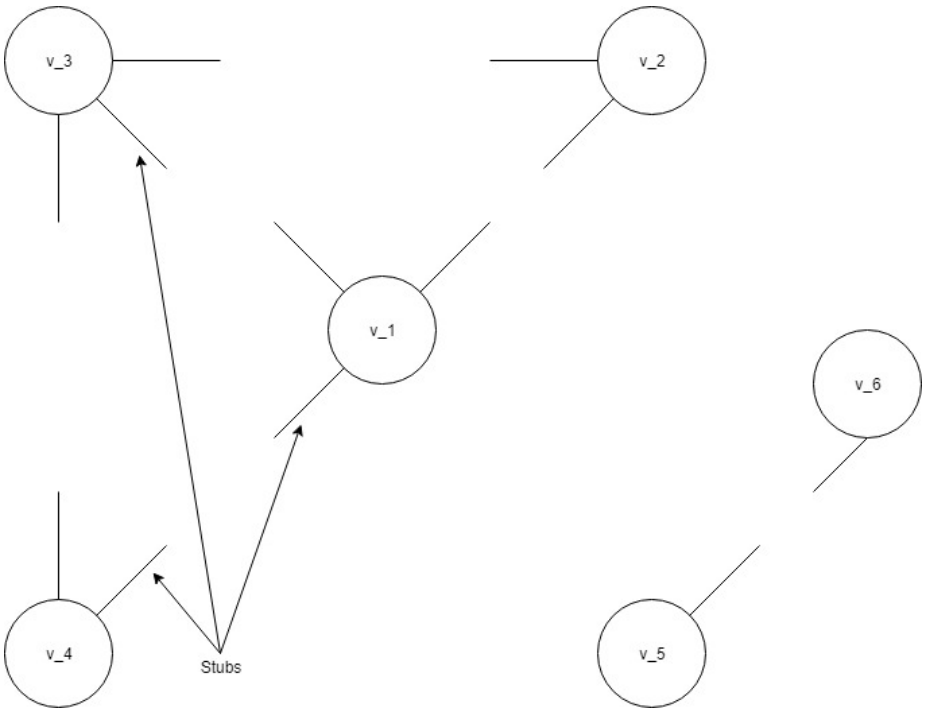

Figure 2: Generating the 6 nodes of the network along with the stubs

1.2.2 Method 2

In Method 2, we generate the n nodes of the network by fixing the number of nodes of a given degree. To be specific, we fix

$$\mathcal{N}_k = \mathbb{E}[\mathcal{N}_k] = \left\lceil n \times \frac{c(k_{max}, \gamma)}{k^\gamma} \right\rceil$$

and generate $\mathcal{N}_1, \mathcal{N}_2, \cdots, \mathcal{N}_{k_{max}}$ such that

$$\mathcal{N}_1 + \mathcal{N}_2 + \cdots + \mathcal{N}_{k_{max}} = n$$

We can fix the value of k_{max} using equation 19

$$2 \times |E| = c(k_{max}, \gamma) \times \mathcal{N}_k \times \sum_{k=1}^{k_{max}} \frac{1}{k^{\gamma-1}} \tag{19}$$

1.2.3 Step 2: stub matching

After we have generated the nodes of the network along with their degrees (using one of the above methods), then we randomly match the stubs to connect the nodes of the network by edges.

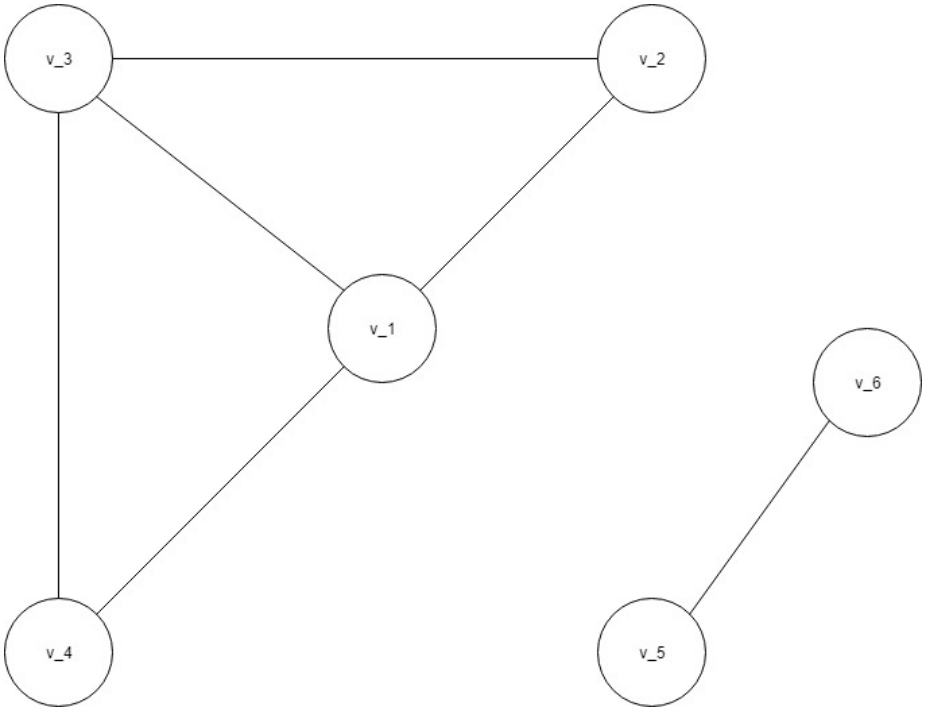

Figure 3: Network after random stub matching

1.2.4 Connectivity of networks created using this algorithm

Networks created using this algorithm will most likely have many connected components. However, if we create the nodes of the network along with their

degrees using Method 2 then it gives rise to a lot more connected and stable network.

$$\left(\frac{|GCC|}{n}\right)_{\text{Method 2}} > \left(\frac{|GCC|}{n}\right)_{\text{Method 1}}$$

2 Generative models for random networks

In this section, we will first state the algorithm for generating a random network using the preferential attachment model and then study the structural properties of the random network generated using the preferential attachment model.

2.1 Preferential attachment model

The preferential attachment model builds an evolving random network using the following algorithm

At the i^{th} time step:

1. A node, v_i, joins the network and brings in m edges
2. For each edge, v_i chooses a node v_j proportional to the degree of v_j and makes m such edges

From step 2 of the algorithm, it can be observed that a node v_i picks it's neighbors preferentially and this model will amplify the number of neighbors of high-degree nodes. Also, by construction, networks created using preferential attachment model are connected.

2.2 Structural properties

Having defined the network generation process using preferential attachment model, we can now analyze some of the structural properties of the network. We will analyze the following structural properties:

- Average degree
- Steady state degree distribution

2.2.1 Average degree

In the preferential attachment model, at each time step we are adding m edges. Therefore,

$$\text{Total number of edges after time step } t = |E|_t = mt$$

Using equation 20,

$$2 \times |E|_t = \sum_{i=1}^{t} deg(v_i) \tag{20}$$

we get,

$$\text{Average degree} = \frac{2mt}{t} = 2m \tag{21}$$

2.2.2 Steady state degree distribution

To facilitate the derivation of the steady state degree distribution, let's introduce some notation:

$\mathcal{N}_k(t)$: Number of nodes of degree k after time step t

$P_k(t)$: Probability of a randomly picked node after time step t having degree k

P_k : Probability of a randomly picked node having degree k

Then using the above notation, at steady state

$$\lim_{t\to\infty} P_k(t) = \lim_{t\to\infty} \frac{\mathcal{N}_k(t)}{t} = P_k \tag{22}$$

Equation 22 states that at steady state, degree distribution is independent of t. Since $\mathcal{N}_k(t)$ is a random variable,so we can rewrite equation 22 to get a steady state condition in the expected sense

$$P_k = \frac{\overline{\mathcal{N}_k(t)}}{t} = \frac{\overline{\mathcal{N}_k(t+1)}}{t+1} \tag{23}$$

where $\overline{\mathcal{N}_k(t)} = \mathbb{E}[\mathcal{N}_k(t)]$. Equation 23 is a representation of the steady state degree distribution in the expected sense. We can use mean-field theory to solve for the steady state degree distribution, P_k. In order to do so, first let's derive a relationship between $\overline{\mathcal{N}_k(t+1)}, \overline{\mathcal{N}_k(t)}$, and $\overline{\mathcal{N}_{k-1}(t)}$

$$\overline{\mathcal{N}_k(t+1)} = \overline{\mathcal{N}_k(t)} - f_k\overline{\mathcal{N}_k(t)} + f_{k-1}\overline{\mathcal{N}_{k-1}(t)} \tag{24}$$

In the above expression, f_k is the probability with which a node with degree k gets a new connection at time step $t+1$ and f_{k-1} is the probability with which a node with degree $(k-1)$ gets a new connection at time step $t+1$. The expressions for f_k and f_{k+1} are given below

$$f_k = m \times \frac{k}{2mt}$$
$$f_{k-1} = m \times \frac{k-1}{2mt}$$

Plugging in the above expressions in equation 24, we get

$$\overline{\mathcal{N}_k(t+1)} = \overline{\mathcal{N}_k(t)} - \frac{k}{2}\frac{\overline{\mathcal{N}_k(t)}}{t} + \frac{k-1}{2}\frac{\overline{\mathcal{N}_{k-1}(t)}}{t}$$

Applying the steady state condition (given by equation 24) to the above expression, we get

$$\begin{aligned}(k+2) \times P_k &= (k-1) \times P_{k-1} \\ \frac{P_k}{P_{k-1}} &= \frac{k-1}{k+2}\end{aligned} \tag{25}$$

We can get the steady state degree distribution P_k by solving the recurrence equation 25. We claim (you can verify on your own) that the the steady state degree distribution P_k is given by the following expression

$$P_k = \frac{c}{k(k+1)(k+2)} \tag{26}$$

where c is the normalization constant. For large values of k we have,

$$\lim_{k\to\infty} \propto \frac{1}{k^3} \tag{27}$$

From equation 27, we can conclude that a random network generated using preferential attachment model has a power law degree distribution.

2.3 Picking a node preferentially

In preferential attachment model, an incoming node v_i picks it's neighbors preferentially. For this preferential attachment, the incoming node v_i needs to know the degrees of all the other nodes in the network. Therefore, the model requires global knowledge. However, there is also another alternative mechanism by which the incoming node choses it's neighbors preferentially. This alternative mechanism only requires local knowledge and is thus superior.

2.3.1 Preferential attachment using random walk

In this mechanism, you do a random walk on the existing network for a sufficiently large number of steps and then select the node you end at. The incoming node forms an edge with the end node of the random walk. Let's show that this mechanism selects the nodes preferentially, that is nodes with higher degree has a higher probability of getting selected. To facilitate the derivation, let's introduce some notation:

P_{ij} : Probability of going to node v_j given that you are at node v_i
P : Node transition matrix with entries P_{ij}
k_i : Degree of node v_i
$\pi_i(m)$: Probability of being at node i after m steps of random walk
A : Node-Node incidence matrix

Then with the above notation, we have the following relation between the node transition and node incidence matrices

$$P = diag(\frac{1}{k_1}, \frac{1}{k_2}, \cdots, \frac{1}{k_{|V|}})A \tag{28}$$

We say that the node occupancy probabilities have reached a steady state if,

$$\pi_i(m) = \pi_i(m+1), \quad i = 1, 2, 3, \cdots, |V| \tag{29}$$

where,

$$\pi_i(m+1) = \sum_{j=1}^{|V|} P_{ji}\pi_j(m) \tag{30}$$

The matrix form of equation 30 is given by equation 31

$$\pi(m+1) = P^T\pi(m) \tag{31}$$

where,

$$\pi(m+1) = [\pi_1(m+1), \pi_2(m+1), \cdots, \pi_{|V|}(m+1)]^T$$

is the node occupancy probability vector after m steps of the random walk. We can use this node occupancy probability vector to rewrite the steady state equation as

$$\pi = P^T \pi \tag{32}$$

The steady state node occupancy probabilities are given by the solution to equation 32. From equation 32, we can observe that the steady state occupancy probability vector,π, is the right eigenvector of P^T. Substituting the expression for P^T in equation 32, we get

$$\begin{aligned} \pi &= A^T diag(\frac{1}{k_1}, \frac{1}{k_2}, \cdots, \frac{1}{k_{|V|}})\pi \\ \pi &= Adiag(\frac{1}{k_1}, \frac{1}{k_2}, \cdots, \frac{1}{k_{|V|}})\pi \end{aligned} \tag{33}$$

We claim that the solution to equation 33 is given by,

$$\begin{aligned} \pi &= \frac{1}{\sum_{i=1}^{|V|} k_i} A\mathbf{1} \\ &= \frac{1}{\mathbf{1}^T A \mathbf{1}} A\mathbf{1} \end{aligned} \tag{34}$$

where, $\mathbf{1} = [1, 1, \cdots, 1]^T \in \mathbb{R}^{|V|}$. We verify that the π given by equation 34 is a solution to the steady state equation 32 in the following steps:

$$\begin{aligned} P^T\pi &= Adiag(\frac{1}{k_1}, \frac{1}{k_2}, \cdots, \frac{1}{k_{|V|}})\left(\frac{1}{\mathbf{1}^T A \mathbf{1}} A\mathbf{1}\right) \\ &= \frac{1}{\mathbf{1}^T A \mathbf{1}} Adiag(\frac{1}{k_1}, \frac{1}{k_2}, \cdots, \frac{1}{k_{|V|}})[k_1, k_2, \cdots, k_{|V|}]^T \\ &= \frac{1}{\mathbf{1}^T A \mathbf{1}} A\mathbf{1} \\ &= \pi \quad \text{(Verified)} \end{aligned}$$

Therefore, doing a random walk for a sufficiently large number of steps will give the steady state occupancy vector π. The steady state occupancy vector π can then be used to pick the nodes in a preferential manner. In the next section, we derive the number of steps required for the random walk to converge to a steady state occupancy vector.

2.3.2 Number of steps required for convergence

The largest eigenvalue of P^T is 1 and let's denote the second largest eigenvalue of P^T by λ_2. Then the rate of convergence of the occupancy vector ($\pi_(m)$) to the steady state occupancy vector (π) depends on the gap $(1 - \lambda_2)$. We have the following result characterizing the rate of convergence

If $(1 - \lambda_2) = \delta$, where $\delta > 0$, then

$$||\pi(m) - \pi||_2 \to 0 \text{ as } (e^{-\delta})^m; \ (e^{-\delta}) < 1 \tag{35}$$

Then using the result given by equation 35 and letting $m = ln(t)$, we have

$$||\pi(m) - \pi||_2 \propto (e^{-ln(t)})^{\delta} = t^{-\lambda}$$

where t is the size of the network. Therefore, $ln(t)$ steps are sufficient for exponential convergence.

3 Random network visualizations

In the earlier sections we analyzed the various structural properties of power law networks (PL) and networks generated using preferential attachment (PA). In this section we visualize those networks and show that although PL and PA networks have the same power law degree distribution but their structure is very different.

Figure 4: Power law network with average degree of 2

Figure 5: Power law network with average degree of 4

From the above figures we can see that power law networks are disconnected and have many connected components. However, for a higher average degree the size of the GCC is larger and there are less isolated nodes.

Figure 6: Preferential attachment network with $m = 1$

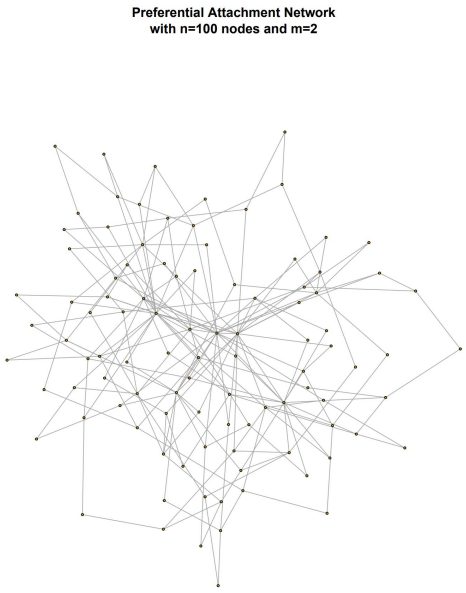

Figure 7: Preferential attachment network with $m = 2$

From the above figures we can see that preferential attachment networks are connected, but their structure is different depending on the value of m. For $m = 1$ we have a tree like structure with a very high modularity.