Copied from Lecture Notes by Remco van der Hofstad "Random Graphs & Complex Networks" Vol. 1

Chapter 8
Preferential attachment models

Abstract

In preferential attachment models, vertices having a fixed number of edges are sequentially added to the network. Given the graph at time t, the edges incident to the vertex with label $t+1$ are attached to older vertices that are chosen according to a probability distribution that is an affine function of the degree of the older vertices. This way, vertices that already have a high degree are more likely to attract edges of later vertices, which explains why such models are called 'rich-get-richer' models. In this chapter, we introduce and investigate such models, focusing on the degree structure of preferential attachment models. We show how the degrees of the old vertices grow with time, and how the proportion of vertices with a fixed degree converges as time grows. Since the vertices with the largest degrees tend to be the earliest vertices in the network, preferential attachment models could also be called 'old-get-rich' models.

8.1 Motivation for the preferential attachment model

Refer to IMDb!

The generalized random graph model and the configuration model described in Chapters 6 and 7, respectively, are *static* models, i.e., the size of the graph is *fixed*, and we have not modeled the *growth* of the graph. There is a large body of work investigating *dynamic* models for complex networks, often in the context of the World-Wide Web, but also for citation networks or biological networks. In various forms, such models have been shown to lead to power-law degree sequences, and, thus, they offer a possible *explanation* for the occurrence of power-law degree sequences in real-world networks. The existence of power-law degree sequences in various real networks is quite striking, and models offering a convincing explanation can teach us about the mechanisms that give rise to their scale-free nature.

A possible explanation for the occurrence of power-law degree sequences is offered by the *preferential attachment paradigm*. In preferential attachment models, vertices are added sequentially with a number of edges connected to them. These edges are attached to a receiving vertex with a probability proportional to the degree of the receiving vertex at that time, thus favoring vertices with large degrees. For this model, it is shown that the number of vertices with degree k decays proportionally to k^{-3} [58], and this result is a special case of the more general result that we prove in this chapter.

The idea behind preferential attachment is simple. In a graph that evolves in time, the newly added vertices are connected to the already existing vertices. In an Erdős-Rényi random graph, which can also be formulated as an evolving graph where edges are added and removed, these edges would be connected to each individual with equal probability (see Exercise 8.1).

Now think of the newly added vertex as a new individual in a social population, which we model as a graph by letting the individuals be the vertices and the edges be the acquaintance relations. Is it then realistic that the edges connect to each already present individual with equal probability, or is the newcomer more likely to get to know socially active individuals, who already know many people? If the latter is true, then we should forget about equal probabilities for the receiving ends of the edges of the newcomer, and introduce a bias in his/her connections towards more social individuals. Phrased in a mathematical way, it should be more likely that the edges are connected to vertices that already have a high degree. A possible model for such a growing graph was proposed by Barabási and Albert [26], and has incited an enormous research effort since.

Strictly speaking, Barabási and Albert in [26] were not the first to propose such a model, and we start by referring to the old literature on the subject. Yule [259] was the first to propose a growing model where preferential attachment is present, in the context of the evolution of species. He derives the power-law distribution that we also find in this chapter. Simon [234] provides a more modern version of the preferential attachment model, as he puts it

> "*Because Yule's paper predates the modern theory of stochastic processes, his derivation was necessarily more involved than the one we shall employ here.*"

The stochastic model of Simon is formulated in the context of the occurrence of words in large pieces of text (as in [260]), and is based on two assumptions, namely (i) that the probability that the $(k+1)$st word is a word that has already appeared exactly i times is proportional to the number of occurrences of words that have occurred exactly i times, and (ii) that there is a constant probability that the $(k+1)$st word is a new word. Together, these two assumptions give rise to frequency distributions of words that obey a power law, with a power-law exponent that is a simple function of the probability of adding a new word. We shall see a similar effect occurring in this chapter. A second place where the model studied by Simon and Yule can be found is in work by Champernowne [72], in the context of income distributions in populations.

In [26], Barabási and Albert describe the preferential attachment graph informally as follows:

> "*To incorporate the growing character of the network, starting with a small number* (m_0) *of vertices, at every time step we add a new vertex with* $m(\leq m_0)$ *edges that link the new vertex to* m *different vertices already present in the system. To incorporate preferential attachment, we assume that the probability* Π *that a new vertex will be connected to a vertex* i *depends on the connectivity* k_i *of that vertex, so that* $\Pi(k_i) = k_i/\sum_j k_j$.

After t time steps, the model leads to a random network with $t+m_0$ vertices and mt edges."

This description of the model is informal, but it must have been given precise meaning in [26] (since, in particular, Barabási and Albert present simulations of the model predicting a power-law degree sequence with exponent close to $\tau = 3$). The model description does not explain how the first edge is connected (note that at time $t = 1$, there are no edges, so the first edge can not be attached according to the degrees of the existing vertices), and does not give the dependencies between the m edges added at time t. We are left wondering whether these edges are independent, whether we allow for self-loops, whether we should update the degrees after each attachment of a single edge, etc. In fact, each of these choices has, by now, been considered in the literature, and the results, in particular the occurrence of power laws and the power-law exponent, do not depend sensitively on the respective choices. See Section 8.9 for an extensive overview of the literature on preferential attachment models.

The first to investigate the model rigorously, were Bollobás, Riordan, Spencer and Tusnady [58]. They complain heavily about the lack of a formal definition in [26], arguing that

"The description of the random graph process quoted above (i.e, in [26], edt.) is rather imprecise. First, as the degrees are initially zero, it is not clear how the process is started. More seriously, the expected number of edges linking a new vertex v to earlier vertices is $\sum_i \Pi(k_i) = 1$, rather than m. Also, when choosing in one go a set S of m earlier vertices as the neighbors of v, the distribution of S is not *specified by giving the marginal probability that each vertex lies in S."*

One could say that these differences in formulations form the heart of much confusion between mathematicians and theoretical physicists. To resolve these problems, choices had to be made, and these choices were, according to [58], made first in [57], by specifying the initial graph to consist of a vertex with m self-loops, and that the degrees will be updated in the process of attaching the m edges. This model will be described in full detail in Section 8.2 below.

This chapter is organized as follows. In Section 8.2, we introduce the model. In Section 8.3, we investigate how the degrees of fixed vertices evolve as the graph grows. In Section 8.4, we investigate the degree sequences in preferential attachment models. The main result is Theorem 8.2, which states that the preferential attachment model has a power-law degree sequence. The proof of Theorem 8.2 consists of two key steps, which are formulated and proved in Sections 8.5 and 8.6, respectively. In Section 8.7, we investigate the maximal degree in a preferential attachment model. In Section 8.8 we discuss some further results on preferential attachment models proved in the literature, and in Section 8.9 we discuss many related preferential attachment models. We close this chapter with notes and discussion in Section 8.10.

8.2 Introduction of the model

We start by introducing the model. The model we investigate produces a *graph sequence* which we denote by $(\mathrm{PA}_t(m,\delta))_{t\geq 1}$, which for every t yields a graph of t vertices and mt edges for some $m = 1, 2, \ldots$ We start by defining the model for $m = 1$. In this case, $\mathrm{PA}_1(1,\delta)$ consists of a single vertex with a single self-loop. We denote the vertices of $\mathrm{PA}_t(1,\delta)$ by $\{v_1^{(1)}, \ldots, v_t^{(1)}\}$. We denote the degree of vertex $v_i^{(1)}$ in $\mathrm{PA}_t(1,\delta)$ by $D_i(t)$, where a self-loop increases the degree by 2.

Then, conditionally on $\mathrm{PA}_t(1,\delta)$, the growth rule to obtain $\mathrm{PA}_{t+1}(1,\delta)$ is as follows. We add a single vertex $v_{t+1}^{(1)}$ having a single edge. This edge is connected to a second end point, which is equal to $v_{t+1}^{(1)}$ with probability $(1+\delta)/(t(2+\delta)+(1+\delta))$, and to a vertex $v_i^{(1)} \in \mathrm{PA}_t(1,\delta)$ with probability $(D_i(t)+\delta)/(t(2+\delta)+(1+\delta))$, where $\delta \geq -1$ is a parameter of the model. Thus,

$$\mathbb{P}\big(v_{t+1}^{(1)} \to v_i^{(1)} \big| \mathrm{PA}_t(1,\delta)\big) = \begin{cases} \dfrac{1+\delta}{t(2+\delta)+(1+\delta)} & \text{for } i = t+1, \\ \dfrac{D_i(t)+\delta}{t(2+\delta)+(1+\delta)} & \text{for } i \in [t]. \end{cases} \tag{8.2.1}$$

The above preferential attachment mechanism is called *affine*, since the attachment probabilities in (8.2.1) depend in an affine way on the degrees of the random graph $\mathrm{PA}_t(1,\delta)$. Exercises 8.2 and 8.3 show that (8.2.1) is a probability distribution.

The model with $m > 1$ is defined in terms of the model for $m = 1$ as follows. Fix $\delta \geq -m$. We start with $\mathrm{PA}_{mt}(1,\delta/m)$, and denote the vertices in $\mathrm{PA}_{mt}(1,\delta/m)$ by $v_1^{(1)}, \ldots, v_{mt}^{(1)}$. Then we identify or collpase $v_1^{(1)}, \ldots, v_m^{(1)}$ in $\mathrm{PA}_{mt}(1,\delta/m)$ to the vertex $v_1^{(m)}$ in $\mathrm{PA}_t(m,\delta)$, and $v_{m+1}^{(1)}, \ldots, v_{2m}^{(1)}$ in $\mathrm{PA}_{mt}(1,\delta/m)$ to be $v_2^{(m)}$ in $\mathrm{PA}_t(m,\delta)$, and, more generally, $v_{(j-1)m+1}^{(1)}, \ldots, v_{jm}^{(1)}$ in $\mathrm{PA}_{mt}(1,\delta/m)$ to be $v_j^{(m)}$ in $\mathrm{PA}_t(m,\delta)$. This defines the model for general $m \geq 1$. We note that the resulting graph $\mathrm{PA}_t(m,\delta)$ is a multigraph with precisely t vertices and mt edges, so that the total degree is equal to $2mt$ (see Exercise 8.4).

To explain the description of $\mathrm{PA}_t(m,\delta)$ in terms of $\mathrm{PA}_{mt}(1,\delta/m)$ by collapsing vertices, we note that an edge in $\mathrm{PA}_{mt}(1,\delta/m)$ is attached to vertex $v_k^{(1)}$ with probability proportional to the weight of vertex $v_k^{(1)}$. Here the weight of $v_k^{(1)}$ is equal to the degree of vertex $v_k^{(1)}$ plus δ/m. Now, the vertices $v_{(j-1)m+1}^{(1)}, \ldots, v_{jm}^{(1)}$ in $\mathrm{PA}_{mt}(1,\delta/m)$ are collapsed to form vertex $v_j^{(m)}$ in $\mathrm{PA}_t(m,\delta)$. Thus, an edge in $\mathrm{PA}_t(m,\delta)$ is attached to vertex $v_j^{(m)}$ with probability proportional to the total weight of the vertices $v_{(j-1)m+1}^{(1)}, \ldots, v_{jm}^{(1)}$. Since the sum of the degrees of the vertices $v_{(j-1)m+1}^{(1)}, \ldots, v_{jm}^{(1)}$ is equal to the degree of vertex $v_j^{(m)}$, this probability is proportional to the degree of vertex $v_j^{(m)}$ in $\mathrm{PA}_t(m,\delta)$ plus δ. We note that in the above construction and for $m \geq 2$, the degrees are updated after each edge is attached. This is what we refer to as *intermediate updating of the degrees.*

The important feature of the model is that edges are more likely to be connected to vertices with large degrees, thus making the degrees even larger. This effect is called *preferential attachment.* Preferential attachment may explain *why* there are quite large degrees. Therefore, the preferential attachment model is sometimes called the *Rich-get-Richer model.* The preferential attachment mechanism is quite reasonable in many real networks. For example, one is more likely to get to know a person who already knows many people, making preferential attachment not unlikely in social networks. However, the precise form of preferential attachment in (8.2.1) is only one possible example. Similarly, a paper is more likely to be cited by other papers when it has already received many citations. Having this said, it is not obvious *why* the preferential attachment rule should be affine. This turns out to be related to the degree-structure of the resulting random graphs $\mathrm{PA}_t(m,\delta)$ that we investigate in detail in this chapter.

The above model is a slight variation of models that have appeared in the literature. The model with $\delta = 0$ is the *Barabási-Albert model,* which has received substantial attention in the literature and which was first formally defined in [57]. We have added the extra parameter δ to make the model more general.

The definition of $(\mathrm{PA}_t(m,\delta))_{t\geq 1}$ in terms of $(\mathrm{PA}_t(1,\delta/m))_{t\geq 1}$ is quite convenient. However, we can also equivalently define the model for $m \geq 2$ directly. We start with $\mathrm{PA}_1(m,\delta)$ consisting of a single vertex with m self-loops. To construct $\mathrm{PA}_{t+1}(m,\delta)$ from $\mathrm{PA}_t(m,\delta)$, we add a single vertex with m edges attached to it. These edges are attached *sequentially* with intermediate updating of the degrees as follows. The eth edge is connected to vertex $v_i^{(m)}$, for $i \in [t]$ with probability proportional to $D_i(e-1,t)+\delta$, where, for $e = 1,\ldots,m$, $D_i(e,t)$ is the degree of vertex i after the eth edge is attached, and to vertex $v_{t+1}^{(m)}$ with probability proportional to $D_{t+1}(e-1,t)+1+e\delta/m$. Here we make the convention that $D_{t+1}(0,t) = 0$. This alternative definition makes it perfectly clear how the choices by Barábasi and Albert missing in [26] are made. Indeed, the degrees are updated during the process of attaching the edges, and the initial graph at time 1 consists of a single vertex with m self-loops. Naturally, the edges could also be attached sequentially by a different rule, for example by attaching the edges independently according to the distribution for the first edge. Also, one has the choice to allow for self-loops or not. See Figure 8.1 for a realization of $(\mathrm{PA}_t(m,\delta))_{t\geq 1}$ for $m = 2$ and $\delta = 0$, and Figure 8.2 for a realization of $(\mathrm{PA}_t(m,\delta))_{t\geq 1}$ for $m = 2$ and $\delta = -1$. Exercises 8.5 investigates related growth rules.

In the literature, slight variations on the above model have been considered. We will discuss two of those. In the first, and for $m = 1$ and $\delta \geq -1$, self-loops do not occur. We denote this variation by $\big(\mathrm{PA}_t^{(b)}(m,\delta)\big)_{t\geq 2}$ and sometimes refer to this model by *model (b)*. To define $\mathrm{PA}_t^{(b)}(1,\delta)$, we let $\mathrm{PA}_2^{(b)}(1,\delta)$ consist of two vertices $v_1^{(1)}$ and $v_2^{(1)}$ with two edges between them, and we replace the growth rule in (8.2.1) by the rule that, for all $i \in [t]$,

$$\mathbb{P}\big(v_{t+1}^{(1)} \to v_i^{(1)} \big| \mathrm{PA}_t^{(b)}(1,\delta)\big) = \frac{D_i(t)+\delta}{t(2+\delta)}. \tag{8.2.2}$$

The advantage of this model is that it leads to a *connected graph.* We again define the model with $m \geq 2$ and $\delta \geq -m$ in terms of $\big(\mathrm{PA}_t^{(b)}(1,\delta/m)\big)_{t\geq 2}$ as below (8.2.1). We also note that the differences between $\big(\mathrm{PA}_t(m,\delta)\big)_{t\geq 1}$ and $\big(\mathrm{PA}_t^{(b)}(m,\delta)\big)_{t\geq 2}$ are minor,

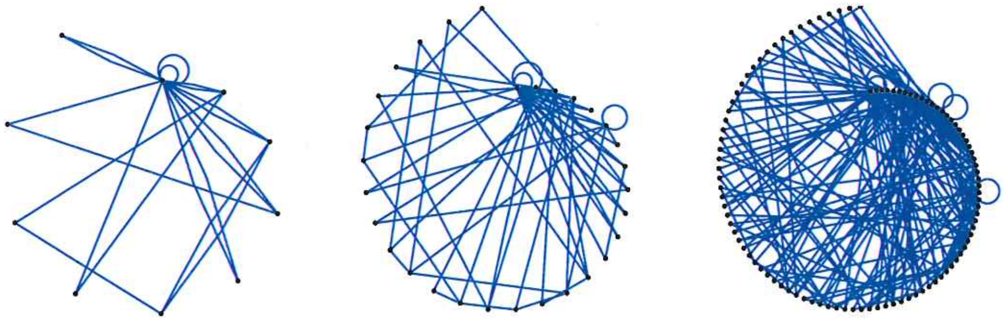

Figure 8.1: Preferential attachment random graph with $m = 2$ and $\delta = 0$ of sizes 10, 30 and 100.

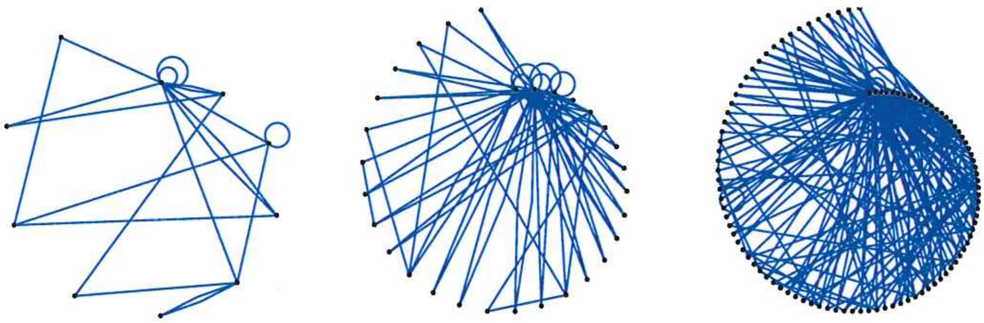

Figure 8.2: Preferential attachment random graph with $m = 2$ and $\delta = -1$ of sizes 10, 30 and 100.

since the probability of a self-loop in $\mathrm{PA}_t(m, \delta)$ is quite small when t is large. Thus, most of the results we shall prove in this chapter for $\big(\mathrm{PA}_t(m, \delta)\big)_{t\geq 1}$ shall also apply to $\big(\mathrm{PA}_t^{(b)}(m, \delta)\big)_{t\geq 2}$, but we do not state these extensions explicitly.

Interestingly, the above model with $\delta \geq 0$ can be viewed as an interpolation between the models with $\delta = 0$ and $\delta = \infty$. We show this for $m = 1$, the statement for $m \geq 2$ can again be seen by collapsing the vertices. We again let the graph at time 2 consist of two vertices with two edges between them. We fix $\alpha \in [0, 1]$. Then, we first draw a Bernoulli random variable I_{t+1} with success probability $1 - \alpha$. The random variables $(I_t)_{t=1}^{\infty}$ are independent. When $I_{t+1} = 0$, then we attach the $(t+1)$st edge to a *uniform* vertex in $[t]$. When $I_{t+1} = 1$, then we attach the $(t + 1)$st edge to vertex $i \in [t]$ with probability $D_i(t)/(2t)$. We denote this model by $\big(\mathrm{PA}_t^{(c)}(1, \alpha)\big)_{t\geq 1}$. When $\alpha \geq 0$ is chosen appropriately, then this is precisely the above preferential attachment model (see Exercise 8.7).

8.3 Degrees of fixed vertices

We start by investigating the degrees of given vertices. To formulate our results, we define the *Gamma-function* $t \mapsto \Gamma(t)$ for $t > 0$ by

$$\Gamma(t) = \int_0^\infty x^{t-1} \mathrm{e}^{-x} dx. \tag{8.3.1}$$

We also make use of the recursion formula (see e.g. Exercise 8.10)

$$\Gamma(t+1) = t\Gamma(t). \tag{8.3.2}$$

The main result in this section is the following:

Theorem 8.1 (Degrees of fixed vertices). *Fix $m = 1$ and $\delta > -1$. Then, $D_i(t)/t^{1/(2+\delta)}$ converges almost surely to a random variable ξ_i as $t \to \infty$, and*

$$\mathbb{E}[D_i(t) + \delta] = (1+\delta) \frac{\Gamma(t+1)\Gamma(i - 1/(2+\delta))}{\Gamma(t + \frac{1+\delta}{2+\delta})\Gamma(i)}. \tag{8.3.3}$$

In Section 8.7, we considerably extend the result in Theorem 8.1. For example, we also prove the almost sure convergence of *maximal* degree.

Proof. Fix $m = 1$ and let $t \geq i$. We compute that

$$\begin{aligned} \mathbb{E}[D_i(t+1) + \delta \mid D_i(t)] &= D_i(t) + \delta + \mathbb{E}[D_i(t+1) - D_i(t) \mid D_i(t)] \\ &= D_i(t) + \delta + \frac{D_i(t) + \delta}{(2+\delta)t + 1 + \delta} \\ &= (D_i(t) + \delta) \frac{(2+\delta)t + 2 + \delta}{(2+\delta)t + 1 + \delta} \\ &= (D_i(t) + \delta) \frac{(2+\delta)(t+1)}{(2+\delta)t + 1 + \delta}. \end{aligned} \tag{8.3.4}$$

(In fact, here we rely on Exercise 8.9.) Using also that

$$\begin{aligned} \mathbb{E}[D_i(i) + \delta] &= 1 + \delta + \frac{1+\delta}{(2+\delta)(i-1) + 1 + \delta} \\ &= (1+\delta) \frac{(2+\delta)(i-1) + 2 + \delta}{(2+\delta)(i-1) + 1 + \delta} \\ &= (1+\delta) \frac{(2+\delta)i}{(2+\delta)(i-1) + 1 + \delta}, \end{aligned} \tag{8.3.5}$$

we obtain (8.3.3). In turn, again using (8.3.4), the sequence $(M_i(t))_{t \geq i}$ given by

$$M_i(t) = \frac{D_i(t) + \delta}{1 + \delta} \prod_{s=i-1}^{t-1} \frac{(2+\delta)s + 1 + \delta}{(2+\delta)(s+1)} \tag{8.3.6}$$

is a non-negative martingale with mean 1. As a consequence of the martingale convergence theorem (Theorem 2.24), as $t \to \infty$, $M_i(t)$ converges almost surely to a limiting random variable M_i.

We next extend this result to almost sure convergence of $D_i(t)/t^{1/(2+\delta)}$. For this, we compute that

$$\prod_{s=i-1}^{t-1} \frac{(2+\delta)s+1+\delta}{(2+\delta)s+2+\delta} = \prod_{s=i-1}^{t-1} \frac{s+\frac{1+\delta}{2+\delta}}{s+1} = \frac{\Gamma(t+\frac{1+\delta}{2+\delta})\Gamma(i)}{\Gamma(t+1)\Gamma(i-1/(2+\delta))}. \tag{8.3.7}$$

Using Stirling's formula, it is not hard to see that (see Exercise 8.11)

$$\frac{\Gamma(t+a)}{\Gamma(t)} = t^a(1+O(1/t)). \tag{8.3.8}$$

Therefore, $D_i(t)/t^{1/(2+\delta)}$ converges in distribution to a random variable ξ_i. In particular, the degrees of the first i vertices at time t is at most of order $t^{1/(2+\delta)}$. Note, however, that we do not yet know whether $\mathbb{P}(\xi_i = 0) = 0$ or not! □

We can extend the above result to the case where $m \geq 1$ by using the relation between $\mathrm{PA}_t(m,\delta)$ and $\mathrm{PA}_{mt}(1,\delta/m)$. This implies in particular that

$$\mathbb{E}_m^{\delta}[D_i(t)] = \sum_{s=1}^{m} \mathbb{E}_1^{\delta/m}[D_{m(i-1)+s}(mt)], \tag{8.3.9}$$

where we have added a subscript m and a superscript δ to the expectation to denote the values of m and δ involved. Exercises 8.12 and 8.13 investigate $m \geq 2$ in more detail, Exercise 8.14 studies extensions to model (b).

We close this section by giving a heuristic explanation for the occurrence of a power-law degree sequence in preferential attachment models. Theorem 8.1 in conjunction with Exercise 8.13 implies that there exists an $a_{m,\delta}$ such that, for i, t large, and any $m \geq 1$,

$$\mathbb{E}_m^{\delta}[D_i(t)] \sim a_{m,\delta}\big(t/i\big)^{1/(2+\delta/m)}. \tag{8.3.10}$$

When the graph indeed has a power-law degree sequence, then the number of vertices with degrees at least k will be close to $ctk^{-(\tau-1)}$ for some $\tau > 1$ and some $c > 0$. The number of vertices with degree at least k at time t is equal to $N_{\geq k}(t) = \sum_{i=1}^t \mathbb{1}_{\{D_i(t)\geq k\}}$. Now, assume that in the above formula, we are allowed to replace $\mathbb{1}_{\{D_i(t)\geq k\}}$ by $\mathbb{1}_{\{\mathbb{E}_m^{\delta}[D_i(t)]\geq k\}}$ (there is a big leap of faith here). Then we would obtain that

$$\begin{aligned} N_{\geq k}(t) &\sim \sum_{i=1}^{t} \mathbb{1}_{\{\mathbb{E}_m^{\delta}[D_i(t)]\geq k\}} \sim \sum_{i=1}^{t} \mathbb{1}_{\{a_{m,\delta}(t/i)^{1/(2+\delta/m)}\geq k\}} \\ &= \sum_{i=1}^{t} \mathbb{1}_{\{i \leq t a_{m,\delta}^{2+\delta/m} k^{-(2+\delta/m)}\}} = t a_{m,\delta}^{2+\delta/m} k^{-(2+\delta/m)}, \end{aligned} \tag{8.3.11}$$

so that we obtain a power-law with exponent $\tau - 1 = 2 + \delta/m$. This suggests that the preferential attachment model has a power-law degree sequence with exponent τ satisfying $\tau = 3 + \delta/m$. The above heuristic is made precise in the following section, but the proof is quite a bit more subtle than the above heuristic!

8.11 Exercises for Chapter 8

Exercise 8.1 (A dynamic formulation of $\mathrm{ER}_n(\lambda/n)$). *Give a dynamical model for the Erdős-Rényi random graph, where at each time n we add a single individual, and where at time n the graph is equal to $\mathrm{ER}_n(\lambda/n)$. See also the dynamic description of the Norros-Reittu model on page 209.*

Exercise 8.2 (Non-negativity of $D_i(t)+\delta$). *Fix $m=1$. Verify that $D_i(t)\geq 1$ for all i and t with $t\geq i$, so that $D_i(t)+\delta\geq 0$ for all $\delta\geq -1$.*

Exercise 8.3 (Attachment probabilities sum up to one). *Verify that the probabilities in (8.2.1) sum up to one.*

Exercise 8.4 (Total degree). *Prove that the total degree of $\mathrm{PA}_t(m,\delta)$ equals $2mt$.*

Exercise 8.5 (Collapsing vs. growth of the PA model). *Prove that the alternative definition of $(\mathrm{PA}_t(m,\delta))_{t\geq 1}$ is indeed equal to the one obtained by collapsing m consecutive vertices in $(\mathrm{PA}_t(1,\delta/m))_{t\geq 1}$.*

Exercise 8.6 (Graph topology for $\delta=-1$). *Show that when $\delta=-1$, the graph $\mathrm{PA}_t(1,\delta)$ consists of a self-loop at vertex $v_1^{(1)}$, and each other vertex is connected to $v_1^{(1)}$ with precisely one edge. What is the implication of this result for $m>1$?*

Exercise 8.7 (Alternative formulation of $\mathrm{PA}_t(1,\delta)$). *For $\alpha=\frac{\delta}{2+\delta}$, show that the law of $\big(\mathrm{PA}_t^{(c)}(1,\alpha)\big)_{t\geq 2}$ is equal to the one of $\big(\mathrm{PA}_t^{(b)}(1,\delta)\big)_{t\geq 2}$. For the original PA model $\big(\mathrm{PA}_t(1,\delta)\big)_{t\geq 2}$ a similar identity holds, with the only difference that the coin probability $\alpha=\alpha_t=\delta(t+1)/[(2t+1)+\delta(t+1)]$ depends slightly on t. Note that, for large t, α_t is asymptotic to $\delta/(2+delta)$, as for $\big(\mathrm{PA}_t^{(b)}(1,\delta)\big)_{t\geq 2}$.*

Exercise 8.8 (Degrees grow to infinity a.s.). *Fix $m=1$ and $i\geq 1$. Prove that $D_i(t)\xrightarrow{a.s.}\infty$. Hint: use that, with $(I_t)_{t=i}^{\infty}$ a sequence of independent Bernoulli random variables with $\mathbb{P}(I_t=1)=(1+\delta)/(t(2+\delta)+1+\delta)$, we have that $\sum_{s=i}^{t} I_s\preceq D_i(t)$. What does this imply for $m>1$?*

Exercise 8.9 (Degree Markov chain). *Prove that the degree $(D_i(t))_{t\geq i}$ forms a (time-inhomogeneous) Markov chain. Compute its transition probabilities for $m=1$.*

Exercise 8.10 (Recursion formula for the Gamma function). *Prove (8.3.2) using partial integration, and also prove that $\Gamma(n)=(n-1)!$ for $n=1,2,\ldots$.*

Exercise 8.11 (Asymptotics for ratio $\Gamma(t+a)/\Gamma(t)$). *Prove (8.3.8), using Stirling's formula (see e.g. [129, 8.327])*

$$\mathrm{e}^{-t}t^{t+\frac{1}{2}}\sqrt{2\pi}\leq\Gamma(t+1)\leq\mathrm{e}^{-t}t^{t+\frac{1}{2}}\sqrt{2\pi}\mathrm{e}^{\frac{1}{12t}}. \tag{8.11.1}$$

Exercise 8.12 (Mean degree for $m\geq 2$). *Prove (8.3.9) and use it to compute $\mathbb{E}_m^{\delta}[D_i(t)]$.*

Exercise 8.13 (A.s. limit of degrees for $m\geq 2$). *Prove that, for $m\geq 2$ and any $i\geq 1$, $D_i(t)(mt)^{-1/(2+\delta/m)}\xrightarrow{a.s.}\xi_i'$, where*

$$\xi_i'=\sum_{j=(i-1)m+1}^{mi}\xi_j, \tag{8.11.2}$$