Degrees in Preferential Attachment Networks with an Anomaly

Qiu Liang$^{(\boxtimes)}$, Remco van der Hofstad , and Nelly Litvak

Department of Mathematics and Computer Science, Eindhoven University
of Technology, Eindhoven, The Netherlands
{q.liang2,r.w.v.d.hofstad,n.v.litvak}@tue.nl

Abstract. We consider a preferential attachment model that incorporates an anomaly. Our goal is to understand the evolution of the network before and after the occurrence of the anomaly by studying the influence of the anomaly on the structural properties of the network. The anomaly is such that after its arrival it attracts newly added edges with fixed probability. We investigate the growth of degrees in the network, finding that the anomaly's degree increases almost linearly. We also provide a heuristic derivation for the exponent of the limiting degree distributions of ordinary vertices, and study the degree growth of the oldest vertex. We show that when the anomaly enters early, the degree distribution is altered significantly, while a late anomaly has minimal impact. Our analysis provides deeper insights into the evolution of preferential attachment networks with an anomalous vertex.

Keywords: Preferential attachment network · Anomaly · Degree Structure · Dynamic network

1 Introduction

Dynamic network models, where nodes and edges appear or disappear over time, attracted a lot of attention in the network science literature. Among these models, the Preferential Attachment (PA) network, as presented by Barabási and Albert [2], has been particularly influential, because it explains the emergence of scale-free property in networks through the 'rich-get-richer' phenomenon. Specifically, the probability that a new vertex connects to an existing vertex is proportional to the degree of the existing vertex.

One extension of the standard PA network is the *superstar model* [7], which is used to analyze key features of retweet networks. In this model, the initial vertex in the network is a *superstar*. At each time step, a newly added vertex connects to the superstar with probability p, or to one of the non-superstar vertices with probability $1 - p$ according to the preferential attachment rule. The superstar alters the dynamics and the resulting degree distribution. It was shown in [7] that the non-superstar vertices follow a power-law degree distribution, though with a modified exponent $3 + \frac{p}{1-p}$. Additionally, the maximal

degree of the non-superstar vertices is likewise affected; it grows slower with the network size. In this paper, we consider a PA model with an anomaly. Our anomaly is similar to the *superstar*, but it may arrive at any point of time.

Since the anomaly alters the network dynamics, our work is closely related to the line of research on PA models with a *change point*, where a parameter of the PA model changes at some point of time. For instance, [1,6] investigate methods to identify the change point in preferential attachment trees via embedding the discrete time tree in a continuous-time branching process and studying the proportion of leaves, while [5] proposes an approach based on the fraction of vertices with minimal degree to detect a late change point. Furthermore, [8] applies the likelihood ratio technique to estimate the change point in a PA model, and extend the method to detect multiple change points via screening and ranking, as well as binary segmentation.

While *change point* detection addresses abrupt changes in the parameters of the network dynamics, here we instead focus on how structural properties evolve when a single anomalous vertex enters a PA network. To explore this, we suggest a new model incorporating an anomaly. Our model is an extension of the *superstar model*, and the two models are highly similar when the anomaly coincides with the *initial* vertex. Our main contributions are as follows:

▷ We propose a PA network with an anomaly. The network evolves according to the standard preferential attachment rule until the anomaly enters. Once the anomaly appears in the network, it attracts newly added edges with a fixed probability, plus a probability that depends on its current degree as in the normal PA dynamics.

▷ We compute the mean degree of the anomaly as a function of network size, and study the mean degree of other vertices and their convergence.

▷ We provide a heuristic derivation of the limiting degree distribution of the ordinary vertices when the anomaly arrives in various different stages of the network's evolution.

Our results serve as a first step towards the understanding of anomaly detection in PA networks. In the final section, we provide an outlook on this detection problem.

2   Model Description

We start by introducing the preferential attachment network without an anomaly. The model constructs a graph sequence $(G_t)_{t\geq 2}$ such that each graph G_t is formed by adding one new vertex and m edges connecting the new vertex to existing vertices, where $m \geq 1$. Let $G_t = (V_t, E_t)$, where $V_t = \{v_1, \cdots, v_t\}$ and $E_t \subseteq \{\{v,w\} : v,w \in V_t\}$. The *initial* graph G_1 consists of a single vertex v_1 and m self-loops. For $t > 1$, no self-loops are present. We treat the process of connecting each edge from a newly introduced vertex to an existing vertex in the network as an individual step. Specifically, let $G_{t,j}$ denote the network after the jth edge of a newly added vertex v_t connects to an existing vertex

$v_i \in \{v_1, v_2, \cdots, v_{t-1}\}$, where $j \in [m] \equiv \{1, \ldots, m\}$. We let $D_i(t,j)$ denote the degree of vertex i in $G_{t,j}$, and introduce a *fitness parameter* $\delta > -m$. Further, we define $G_t = G_{t,m} = G_{t+1,0}$, and $D_i(t) = D_i(t,m)$. For $t > 1$, $j \in [m]$, we define the attachment rule for the jth edge linking to the vertex $v_i \in \{v_1, v_2, \cdots, v_{t-1}\}$ as defined in [3,4], and studied further in [5]:

$$P(v_{t,j} \to v_i \mid G_{t,j-1}) = \frac{D_i(t,j-1)+\delta}{2m(t-1)+(t-1)\delta+(j-1)}. \qquad (1)$$

Assume that an anomaly occurs at some time τ satisfying $1 < \tau < t$, where the attachment rule changes after the anomaly has occurred. We denote the anomaly by the vertex v_τ. After τ, each new edge connects to the anomaly with probability

$$p \approx \frac{\beta}{2m+\beta+\delta}, \qquad (2)$$

where $\beta > 0$ is a parameter of the model. Otherwise, with complementary probability, the edge connects to any existing vertex, including the anomaly, by following the usual PA rule. Formally, the dynamics of the model *at step t* are as follows:

(I) If $t < \tau$, the anomalous vertex has not occurred in the graph G_t, the attachment rule of G_t is the same as (1).

(II) If $1 < \tau \leq t$, the evolution rule changes as follows:

$$P(v_{t,j} \to v_i \mid G_{t,j-1}) = \begin{cases} \dfrac{D_i(t,j-1)+\delta}{(t-1)(2m+\beta+\delta)+j-1} & \text{if } i \neq \tau, \\[2ex] \dfrac{(t-1)\beta+D_\tau(t,j-1)+\delta}{(t-1)(2m+\beta+\delta)+j-1} & \text{if } i = \tau. \end{cases} \qquad (3)$$

To explain the rationale behind (3), suppose $m=1$, hence $j=1$ is the only edge of vertex v_t. Then

$$\frac{(t-1)\beta+D_\tau(t-1)+\delta}{(t-1)(2+\beta+\delta)} = p+(1-p)\frac{D_\tau(t-1)+\delta}{(t-1)(2+\delta)},$$

with $p = \frac{\beta}{2+\beta+\delta}$. Hence, we can think of our connection rule as connecting with probability p to the anomaly, and proportional to the degrees (including the anomaly) with probability $1-p$. This explains the choice in (2), which is *exactly* correct for $m=1$. To make (2) *exactly* correct for $m > 1$, we have to choose $p = p_{\beta,t,j}$ slightly differently, so that

$$p_{\beta,t,j} + (1-p_{\beta,t,j})\frac{D_\tau(t,j-1)+\delta}{(t-1)(2m+\delta)+j-1}$$
$$= \frac{p_{\beta,t,j}[(t-1)(2m+\delta)+(j-1)]+(1-p_{\beta,t,j})(D_\tau(t,j-1)+\delta)}{(t-1)(2m+\delta)+j-1}$$
$$= \frac{(t-1)\beta+D_\tau(t,j-1)+\delta}{(t-1)(2m+\beta+\delta)+j-1},$$

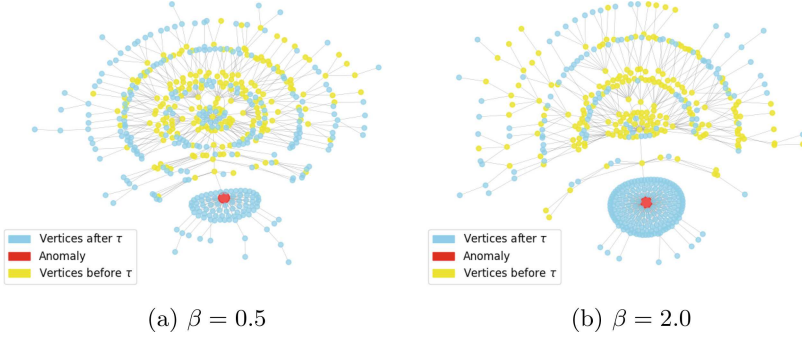(a) $\beta = 0.5$ (b) $\beta = 2.0$

Fig. 1. Examples of PA networks with an anomaly. Here $t = 500$, $\tau = 200$, $\delta = 0$, $m = 1$.

which leads to $p_{\beta,t,j} \approx p$ in (3). Since the precise form of (3) is a little simpler, we choose to work with this parameterization instead to simplify the formulas.

Figure 1 illustrates an example of a PA network with an anomaly. We see that a large number of edges connect to the anomaly, yet a positive proportion of the edges of vertices arriving after τ are attached to ordinary vertices. Our goal is to study the structural properties of a PA network with an anomaly and the asymptotic degree distribution for different types of vertices.

3   The Growth of the Degrees in G_t

In this section, we analyze the growth of the degrees in G_t, both for the anomaly (Sect. 3.1) and the ordinary vertices (Sects. 3.2, 3.3). We follow the approach in [9, Chapter 8], adapted to our model with an anomaly.

3.1   Expected Degree of the Anomaly

We start by investigating the expected degree of the anomaly:

Proposition 1 (Degree of the anomaly). *Consider an anomaly that occurs at time τ, where $1 < \tau \ll t$, and follows the attachment rule (3) with given $\delta > -m$, and $\beta > 0$. Then*

$$\mathbb{E}[D_\tau(t) + \delta] = \frac{m\beta t}{m + \beta + \delta} + c_0 \frac{\Gamma(t + \frac{m}{2m+\beta+\delta})\Gamma(\tau)}{\Gamma(t)\Gamma(\tau + \frac{m}{2m+\beta+\delta})}. \tag{4}$$

where $c_0 = m + \delta - \frac{m\beta\tau}{m+\beta+\delta}$. In particular,

$$\lim_{\beta \to \infty} \mathbb{E}[D_\tau(t) + \delta] = (t - \tau + 1)m + \delta. \tag{5}$$

Proof. Recall that $\mathbb{E}[D_\tau(\tau)+\delta]=m+\delta$, for $k<\tau<t$. Based on ([3](#3)), the expected degree of the anomaly satisfies the recursion

$$\mathbb{E}[D_\tau(t,m)+\delta|D_\tau(t,m-1)]$$
$$=D_\tau(t,m-1)+\delta+\mathbb{E}[D_\tau(t,m)-D_\tau(t,m-1)|G_{t,m-1}]$$
$$=D_\tau(t,m-1)+\delta+\frac{(t-1)\beta+D_\tau(t,m-1)+\delta}{(t-1)(2m+\beta+\delta)+(m-1)}$$
$$=(D_\tau(t,m-1)+\delta)\left(1+\frac{1}{(t-1)(2m+\beta+\delta)+(m-1)}\right)$$
$$+\frac{(t-1)\beta}{(t-1)(2m+\beta+\delta)+(m-1)}.$$

Taking expectation, and solving the recursion, gives that

$$\mathbb{E}[D_\tau(t,m)+\delta]=(m+\delta)\prod_{k_2=\tau}^{t-1}\prod_{k_1=0}^{m-1}\left(1+\frac{1}{k_2(2m+\beta+\delta)+k_1}\right)+C^m(\beta,\delta,t)$$
$$=(m+\delta)\prod_{k_2=\tau}^{t-1}\frac{k_2+\frac{m}{2m+\beta+\delta}}{k_2}+C^m(\beta,\delta,t)$$
$$=(m+\delta)\frac{\Gamma(t+\frac{m}{2m+\beta+\delta})\Gamma(\tau)}{\Gamma(\tau+\frac{m}{2m+\beta+\delta})\Gamma(t)}+C^m(\beta,\delta,t), \tag{6}$$

where $C^m(\beta,\delta,t)$ is a function of β, t and δ, that, with $c=2m+\beta+\delta$, equals

$$C^m(\beta,\delta,t)=\sum_{k=0}^{m-1}\left[\frac{(t-1)\beta}{(t-1)c+k}\times\frac{(t-1)c+m}{(t-1)c+k+1}\right]$$
$$+\frac{(t-1)c+m}{(t-1)c}\times\sum_{k=0}^{m-1}\left[\frac{(t-2)\beta}{(t-2)c+k}\times\frac{(t-2)c+m}{(t-2)c+k+1}\right]+\cdots$$
$$+\prod_{t_1=\tau+1}^{t-1}\frac{t_1c+m}{t_1c}\times\sum_{k=0}^{m-1}\left[\frac{\tau\beta}{\tau c+k}\times\frac{\tau c+m}{\tau c+k+1}\right].$$

For each time step, by the telescoping sum identity,

$$\sum_{k=0}^{m-1}\frac{t_1\beta(t_1c+m)}{(t_1c+k)(t_1c+k+1)}=t_1\beta(t_1c+m)\sum_{k=0}^{m-1}\left[\frac{1}{(t_1c+k)}-\frac{1}{(t_1c+k+1)}\right]$$
$$=t_1\beta(t_1c+m)\left(\frac{1}{t_1c}-\frac{1}{t_1c+m}\right)=\frac{m\beta}{c}.$$

Thus,

$$C^m(\beta\delta t) = \frac{m\beta}{c}\left[1 + \frac{(t-1)c+m}{(t-1)c} + \cdots + \prod_{t_1=\tau+1}^{t-1}\frac{t_1c+m}{t_1c}\right]$$

$$= \frac{m\beta}{c}\sum_{t_1=\tau+1}^{t}\frac{\Gamma(t+\frac{m}{c})\Gamma(t_1)}{\Gamma(t_1+\frac{m}{c})\Gamma(t)}.$$

Using properties of the gamma function, we can rewrite

$$C^m(\beta\delta t) = \frac{m\beta\Gamma(t+\frac{m}{c})}{c\Gamma(t)}\left(\frac{1}{\frac{m}{c}-1}\right)\times\sum_{t_1=\tau+1}^{t}\left(\frac{\Gamma(t_1)}{\Gamma(t_1-1+\frac{m}{c})} - \frac{\Gamma(t_1+1)}{\Gamma(t_1+\frac{m}{c})}\right)$$

$$= \frac{m\beta}{m+\beta+\delta}\frac{\Gamma(t+\frac{m}{2m+\beta+\delta})}{\Gamma(t)}\left(\frac{\Gamma(t+1)}{\Gamma(t+\frac{m}{2m+\beta+\delta})} - \frac{\Gamma(\tau+1)}{\Gamma(\tau+\frac{m}{2m+\beta+\delta})}\right)$$

$$= \frac{m\beta t}{m+\beta+\delta} - \frac{m\beta\tau}{m+\beta+\delta}\frac{\Gamma(t+\frac{m}{2m+\beta+\delta})\Gamma(\tau)}{\Gamma(t)\Gamma(\tau+\frac{m}{2m+\beta+\delta})}. \tag{7}$$

Substituting (7) into (6), we get (4). In particular, $\frac{m\beta}{m+\beta+\delta}\to m$ and $\frac{m}{2m+\beta+\delta}\to 0$ as $\beta\to\infty$, so that

$$\lim_{\beta\to\infty}\mathbb{E}[D_\tau(t)+\delta] = mt + (m+\delta-m\tau),$$

which implies that all the incoming edges are expected to connect to the anomaly, and we get (5).

For general $j\in[m]$, we can extend (4) to find $\mathbb{E}[D_\tau(t,j)+\delta]$. Applying the recursive approach,

$$\mathbb{E}[D_\tau(t,j)+\delta|D_\tau(t,j-1)]$$

$$= D_\tau(t,j-1)+\delta + \frac{(t-1)\beta+D_\tau(t,j-1)+\delta}{(t-1)(2m+\beta+\delta)+j-1}$$

$$= (D_\tau(t-1,m)+\delta)\times\frac{t-1+\frac{j}{2m+\beta+\delta}}{t-1} + \frac{\beta j}{2m+\beta+\delta},$$

and taking expectation, we have

$$\mathbb{E}[D_\tau(t,j)+\delta] = \mathbb{E}[D_\tau(t-1,m)+\delta]\left(1 + \frac{j}{(t-1)(2m+\beta+\delta)}\right) + \frac{\beta j}{2m+\beta+\delta}.$$

We obtain (4) by solving this recursion.   □

Figure 2 shows the growth of $D_\tau(t)$ over time, with our theoretical result for $\mathbb{E}[D_\tau(t)+\delta]$ depicted as yellow line. It is very interesting that the coefficient of the linear growth, $\frac{\beta}{m+\beta+\delta}$ is larger than the probability $\frac{\beta}{2m+\beta+\delta}$ in (3) that an edge attaches itself to an anomaly. Indeed, our assumption that vertices may attach to the anomaly also through the PA mechanism, has increased the rate of growth
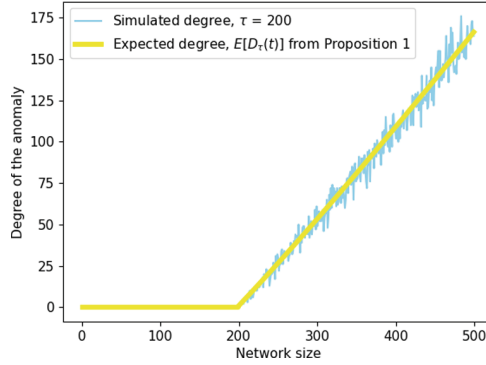
Fig. 2. The degree of the anomaly as a function of time t. The parameters used are $\tau = 200, \delta = 0, \beta = 2.0, \text{m} = 1$.

(rather than, say, giving rise to an extra polynomial term as we conjectured at the beginning). To further explore the relation between $D_\tau(t)$ and the network size, we assume that $t = a\tau, a \geq 1$. By Stirling's formula,

$$\frac{\Gamma(t+a)}{\Gamma(t)} = t^a(1+O(1/t)), \tag{8}$$

when $t \to \infty$ and a is fixed (see, e.g., [9, (8.3.9)]). This approximation allows us to simplify the formula of expected degree of v_τ. Indeed, applying it to (4) gives

$$f(a) = \lim_{\tau \to \infty} \frac{\mathbb{E}[D_\tau(a\tau) + \delta]}{a\tau} = \frac{m\beta}{m+\beta+\delta}\left(1 - a^{-\frac{m+\beta+\delta}{2m+\beta+\delta}}\right),$$

and

$$f(1) = 0, \qquad \lim_{a \to \infty} f(a) = \frac{m\beta}{m+\beta+\delta},$$

$$f'(a) = \frac{m\beta}{2m+\beta+\delta} a^{-\frac{m+\beta+\delta}{2m+\beta+\delta}-1}, \qquad f'(1) = \frac{m\beta}{2m+\beta+\delta}.$$

We see that larger values of a, corresponding to an earlier anomaly, result in a larger deviation from the linear growth of $\mathbb{E}[D_\tau(a\tau)]$.

3.2   Expected Degree of the Ordinary Vertices

Despite the presence of an anomaly, ordinary vertices continue to receive edges based on the preferential attachment mechanism. We apply the recursive approach from [9, Section 8.3] to calculate their expected degrees, as the anomaly does not alter the edge assignment mechanism for these vertices. Regarding the attachment function, it differs for $i \leq t \leq \tau$ and $t > \tau$. Therefore, we will analyze the following two scenarios:

(1) If $i < \tau$, for $t < \tau$, the anomaly has not yet occurred, the expected degree of v_i is

$$
\mathbb{E}[D_i(G_t) + \delta] = \begin{cases} (2m+\delta) \ \dfrac{\Gamma(t + \frac{m}{2m+\delta})}{\Gamma(1 + \frac{m}{2m+\delta})\Gamma(t)}, & i = 1, \\[3mm] (m+\delta) \ \dfrac{\Gamma(t + \frac{m}{2m+\delta})\Gamma(i)}{\Gamma(i + \frac{m}{2m+\delta})\Gamma(t)}, & 1 < i < \tau. \end{cases}
\tag{9}
$$

See, e.g., [9, Exercise 8.14] for the case where $m = 1$. Remarkably, for the model chosen here, the formula for $m > 1$ is actually quite nice.

For $t > \tau$, the attachment rule is changed after the anomaly appears. Then, the expected degree of v_i changes into

$$
\mathbb{E}[D_i(t) + \delta] = \begin{cases} (2m+\delta) \ \dfrac{\Gamma(\tau + \frac{m}{2m+\delta})\Gamma(t + \frac{m}{2m+\beta+\delta})}{\Gamma(1 + \frac{m}{2m+\delta})\Gamma(\tau + \frac{m}{2m+\beta+\delta})\Gamma(t)}, & i = 1, \\[3mm] (m+\delta) \ \dfrac{\Gamma(\tau + \frac{m}{2m+\delta})\Gamma(t + \frac{m}{2m+\beta+\delta})\Gamma(i)}{\Gamma(i + \frac{m}{2m+\delta})\Gamma(\tau + \frac{m}{2m+\beta+\delta})\Gamma(t)}, & 1 < i < \tau. \end{cases}
\tag{10}
$$

(2) If $i > \tau$, for $t > \tau$, the expected degree of v_i is

$$
\mathbb{E}[D_i(t) + \delta] = (m+\delta) \ \frac{\Gamma(t + \frac{m}{2m+\beta+\delta})\Gamma(i)}{\Gamma(i + \frac{m}{2m+\beta+\delta})\Gamma(t)}.
\tag{11}
$$

3.3   Convergence of Degrees for Ordinary Vertices

In [9, Section 8.3], it is proved that the degree of vertices in a standard PA network scales as $t^{\frac{1}{2+\delta}}$ when $m = 1$ and as $t^{\frac{1}{2+\delta/m}}$ when $m \geq 2$ (see [9, Exercise 8.13]). In our model, if $t < \tau$, the anomaly is not included in G_t, so that again the degrees of vertices are of the order $t^{\frac{1}{2+\delta/m}}$. When $t > \tau$, it is necessary to analyze the convergence of the degrees of vertices separately.

For the vertices added before τ, we consider the sequence $(M_i^{(1)}(t))_{t \geq i}$ given by

$$
M_i^{(1)}(t) = \frac{D_i(t) + \delta}{\mathbb{E}[D_i(t) + \delta]}.
$$

It is easy to see that $(M_i^{(1)}(t))_{t \geq i}$ is a non-negative martingale. Indeed, since $m \leq D_i(t) < 2mt$, we have $\mathbb{E}[|M_i^{(1)}(t)|] < \infty$. Computing the conditional expectation, we get

$$\mathbb{E}[M_i^{(1)}(t+1)|M_i^{(1)}(t)] = \mathbb{E}[M_i^{(1)}(t+1)|D_i(t)]$$

$$= \frac{\mathbb{E}[D_i(t+1)+\delta|D_i(t)]}{\mathbb{E}[D_i(t+1)+\delta]}$$

$$= \frac{(D_i(t)+\delta)}{\mathbb{E}[D_i(t+1)+\delta]} \prod_{j=1}^{m}\left(1+\frac{1}{t(2m+\beta+\delta)+j-1}\right)$$

$$= \frac{D_i(t)+\delta}{\mathbb{E}[D_i(t+1)+\delta]} \frac{t+\frac{m}{2m+\beta+\delta}}{t}$$

$$= \frac{D_i(t)+\delta}{\mathbb{E}[D_i(t)+\delta]} = M_i^{(1)}(t),$$

since also

$$\mathbb{E}[D_i(t+1)+\delta] = \mathbb{E}[D_i(t)+\delta]\frac{t+\frac{m}{2m+\beta+\delta}}{t}. \tag{12}$$

Thus, $(M_i^{(1)}(t))_{t\geq i}$ is a non-negative martingale with respect to $(G_t)_{t\geq i}$. According to the martingale convergence theorem, $M_i^{(1)}(t)$ converges almost surely to a limiting random variable as $t\to\infty$ [9, Theorem 2.24], consequently, the result can be extended to establish the convergence of degrees when $i < \tau$. By the Stirling's formula, for sufficiently large t and τ,

$$\frac{D_i(t)+\delta}{\left(\frac{t}{\tau}\right)^{\frac{m}{2m+\beta+\delta}}\tau^{\frac{m}{2m+\delta}}} = M_i^{(1)}(t)\frac{(d+\delta)\Gamma(i)}{\Gamma(i+\frac{m}{2m+\delta})}(1+o(1))$$

$$\xrightarrow{a.s.} \frac{(d+\delta)\Gamma(i)}{\Gamma(i+\frac{m}{2m+\delta})}\xi_i^{(1)},$$

where $\xi_i^{(1)}$ is the almost sure limit of $M_i^{(1)}(t)$, and $d=2m$ for $i=1$ and $d=m$ for $i>1$. Thus, when $i<\tau$, $\frac{D_i(t)+\delta}{\left(\frac{t}{\tau}\right)^{\frac{m}{2m+\beta+\delta}}\tau^{\frac{m}{2m+\delta}}}$ converges almost surely as $t\to\infty$ and $\tau\to\infty$.

Similarly, for the vertices added after τ, let $(M_i^{(2)}(t))_{t\geq i}$ be given by

$$M_i^{(2)}(t) = \frac{D_i(t)+\delta}{\mathbb{E}[D_i(t)+\delta]}, \quad > \qquad i$$

Following the previously described steps again, we get $(M_i^{(2)}(t))_{t\geq i}$ is a non-negative martingale, and when t is large enough,

$$\frac{D_i(t)+\delta}{t^{\frac{m}{2m+\delta}}} = M_i^{(2)}(t)\frac{(m+\delta)\Gamma(i)}{\Gamma(i+\frac{m}{2m+\beta+\delta})}(1+o(1))$$

$$\xrightarrow{a.s.} \frac{(m+\delta)\Gamma(i)}{\Gamma(i+\frac{m}{2m+\beta+\delta})}\xi_i^{(2)},$$

where $\xi_i^{(2)}$ is the almost sure limit of $M_i^{(2)}(t)$. Thus, when $i>\tau$, $\frac{D_i(t)+\delta}{t^{\frac{m}{2m+\delta}}}$ converges almost surely.

4 Heuristic Derivation of the Limiting Degree Distribution for Ordinary Vertices

In this section, we aim to investigate the limiting degree distribution of the ordinary vertices in a PA network containing an anomaly. For the standard PA network, the exponent of power law degree is $3 + \frac{\delta}{m}$. The rigorous proof in [9, Chapter 8] strongly relies on the recursive relation between the fractions of vertices with degree k. However, it is hard to adapt this approach, due to the alteration in the recursion after the appearance of an anomaly. Consequently, the rigorous derivation remains an open problem for future research. Here, we present a heuristic argument [11] that yields results consistent with our numerical simulations.

The main idea is that if we can find the range of vertex index i such that the average degree of vertex v_i falls in $(k - 0.5, k + 0.5)$, then the limiting fraction p_k of vertices of degree k can be evaluated as that range divided by t, as $t \to \infty$. We will explore the power-law degree distribution in different scaling regimes for the arrival time of anomaly τ as a function of time t. Specifically, we consider three cases: when the anomaly arrives late, mid-way or early.

4.1 Late Anomaly: $\tau = t - t^\gamma$, $\gamma \in (0,1)$

If the arrival time of v_τ is quite late, for example when $\tau = t - t^\gamma$, where $\gamma \in (0,1)$, then only a vanishing fraction of vertices arrives after the anomaly. Therefore, the degree distribution will be defined by the vertices that arrived before the anomaly. When $i < \tau$ the expected degree of v_i is (10), here we only consider $i \gg 1$, as t and τ are large enough, by the Stirling formula, $\mathbb{E}[D_i(t) + \delta]$ can be approximated as

$$\mathbb{E}[D_i(t) + \delta] = (m + \delta) \left(\frac{t}{t - t^\gamma} \right)^{\frac{m}{2m + \beta + \delta}} \left(\frac{t - t^\gamma}{i} \right)^{\frac{m}{2m + \delta}} (1 + o(1)), \quad < i \quad (13)$$

Assume that $\mathbb{E}[D_i(t) + \delta]$ falls in the interval $(k + \delta - 0.5, k + \delta + 0.5)$, then i should be in

$$\left((m + \delta)^{2 + \frac{\delta}{m}} \frac{(1 - t^{\gamma - 1})^{\frac{\beta}{2m + \beta + \delta}} t}{(k + \delta + 0.5)^{2 + \frac{\delta}{m}}}, (m + \delta)^{2 + \frac{\delta}{m}} \frac{(1 - t^{\gamma - 1})^{\frac{\beta}{2m + \beta + \delta}} t}{(k + \delta - 0.5)^{2 + \frac{\delta}{m}}} \right).$$

Let L_1 be the length of interval,

$$L_1 = (m + \delta)^{2 + \frac{\delta}{m}} (1 - t^{\gamma - 1})^{\frac{\beta}{2m + \beta + \delta}} t \left(\frac{1}{(k + \delta - 0.5)^{2 + \frac{\delta}{m}}} - \frac{1}{(k + \delta + 0.5)^{2 + \frac{\delta}{m}}} \right)$$

$$= (m + \delta)^{2 + \frac{\delta}{m}} (1 - t^{\gamma - 1})^{\frac{\beta}{2m + \beta + \delta}} t \frac{k^{2 + \frac{\delta}{m}} \left[\left(1 + \frac{\delta + 0.5}{k} \right)^{2 + \frac{\delta}{m}} - \left(1 - \frac{0.5 - \delta}{k} \right)^{2 + \frac{\delta}{m}} \right]}{k^{4 + \frac{2\delta}{m}} \left(1 + \frac{\delta + 0.5}{k} \right)^{2 + \frac{\delta}{m}} \left(1 - \frac{0.5 - \delta}{k} \right)^{2 + \frac{\delta}{m}}}.$$

Next we apply the binomial theorem to simplify the expression of L_1, to obtain, as $k \to \infty$,

$$L_1 = \left(2 + \frac{\delta}{m}\right)(m+\delta)^{2+\frac{\delta}{m}}(1-t^{\gamma-1})^{\frac{\beta}{2m+\beta+\delta}}tk^{-(3+\frac{\delta}{m})}(1+o(1)),$$

so that, for $k \to \infty$,

$$p_k^{(\text{late})} = \lim_{t \to \infty}\frac{L_1}{t} = \left(2 + \frac{\delta}{m}\right)(m+\delta)^{2+\frac{\delta}{m}}k^{-(3+\frac{\delta}{m})}(1+o(1)). \qquad (14)$$

Thus, the asymptotic degree distribution remains unchanged compared to the standard PA network.

4.2 Mid-Way Anomaly: $\tau = \alpha T, \alpha \in (0,1)$

Assume that the arrival time of v_τ scales linearly with t, that is $\tau = \alpha t$, where $\alpha \in (0,1)$. For $1 < i < \tau$, the average degree is approximated by

$$\mathbb{E}[D_i(t)+\delta] = (m+\delta)\alpha^{-\frac{m}{2m+\beta+\delta}}\left(\frac{\alpha t}{i}\right)^{\frac{m}{2m+\delta}}(1+o(1)), < i \qquad (15)$$

Similarly, for the vertex v_i with $i > \tau$,

$$\mathbb{E}[D_i(t)+\delta] = (m+\delta)\left(\frac{t}{i}\right)^{\frac{m}{2m+\beta+\delta}}(1+o(1)). \qquad (16)$$

We see that the average degree of the vertices born after τ grows slower than that of the vertices that arrived before τ. Moreover, the right-hand side of (16) is bounded, and thus it cannot fall into the interval $(k+\delta-0.5, k+\delta+0.5)$ for large k. Therefore, we perform derivations for the vertices that have arrived before time τ as the higher order probability to achieve a large degree k. Repeat the steps in Sect. 4.1, we get, as $k \to \infty$,

$$p_k^{(\text{mid}-\text{way})} = \left(2 + \frac{\delta}{m}\right)(m+\delta)^{2+\frac{\delta}{m}}\alpha^{\frac{\beta}{2m+\beta+\delta}}k^{-(3+\frac{\delta}{m})}(1+o(1)). \qquad (17)$$

From this formula we see that the vertices arriving before τ follow the power law distribution with exponent $3 + \frac{\delta}{m}$. This is the same power-law exponent as in the standard PA network. However, we get a factor $\alpha^{\frac{\beta}{2m+\beta+\delta}}$ in front. This conforms with the intuition that the anomaly slows down the degree growth of high-degree vertices. Moreover, this factor decreases with β, as a larger β increases the effect of the anomaly.

4.3 Early Anomaly: $\tau = t^\gamma, \gamma \in (0,1)$

Suppose that the anomaly arrives quite early at $\tau = t^\gamma$, where $\gamma \in (0,1)$. Then the fraction of vertices born before τ among all vertices is vanishing. Therefore, we investigate the behavior of vertices born after τ. Repeating the steps in

Sect. 4.1, we obtain an asymptotic degree distribution given by

$$p_k^{(\text{early})} = \left(2 + \frac{\beta}{m} + \frac{\delta}{m}\right)(m+\delta)^{2+\frac{\beta}{m}+\frac{\delta}{m}} k^{-(3+\frac{\beta}{m}+\frac{\delta}{m})}(1+o(1)), \; k \to \infty. \quad (18)$$

The power-law exponent of the degree distribution changes to $3 + \frac{\beta}{m} + \frac{\delta}{m}$. We see that the anomaly has altered the power-law exponent, as also observed in the superstar model [7].

4.4   Comparison to the Empirical Degree Distribution

Figure 3, 4 and 5 show the empirical and the theoretically predicted degree distributions for the PA network with the late anomaly, the mid-way anomaly, and the early anomaly, respectively.
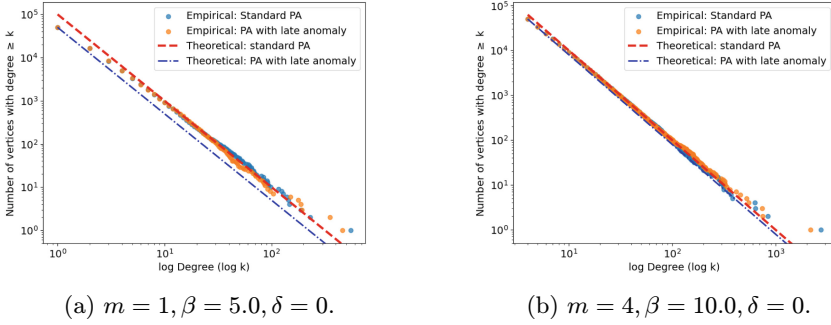

(a) $m = 1, \beta = 5.0, \delta = 0.$ (b) $m = 4, \beta = 10.0, \delta = 0.$

Fig. 3. The complementary cumulative degree distribution for PA network with late anomaly, parameters are $t = 50000, \tau = 49950, \gamma = 0.3615$.

Generally, we see that our computations correctly predict the slope, but the multiplicative factor may deviate from the experiments. In the future, a rigorous derivation for the mid-way and early anomaly is needed. The late anomaly is equivalent to standard PA network but we plot the line that we derived in (14) to show the difference between the correct multiplicative factor and that resulting from the heuristic derivation.

In Fig. 3 we see that when the anomaly arrives near the end of the network's growth, the proportion of vertices with degree at least k is close to that of the standard PA network. The straight line derived from (14) (dashed blue line) has the correct slope, but is slightly different from the empirical degree distribution. The dashed red line, based on the more precise formula (8.4.11) in [9, Chapter 8] for the standard preferential attachment model has both the slope and the multiplicative factor matching the experiments.

In Fig. 4 for the mid-way anomaly, the proportion of vertices with degree at least k is close to that for the standard PA network, but differs by a factor smaller
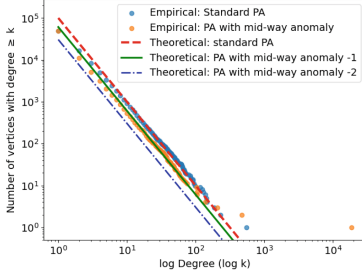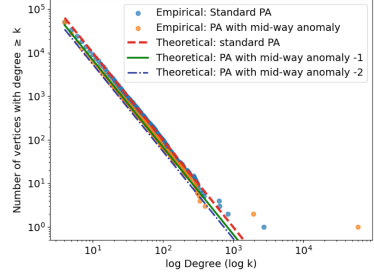
(a) $m = 1, \beta = 5.0, \delta = 0.$ (b) $m = 4, \beta = 10.0, \delta = 0.$

Fig. 4. The complementary cumulative degree distribution for PA network with midway anomaly, parameters are $t = 50000, \tau = 25000, \alpha = 0.5$. The green line uses the formula (8.4.11) in [9, Chapter 8], multiplied by factor $\alpha^{\frac{\beta}{2m+\beta+\delta}}$ as in (17). The dashed blue line follows (17). (Color figure online)
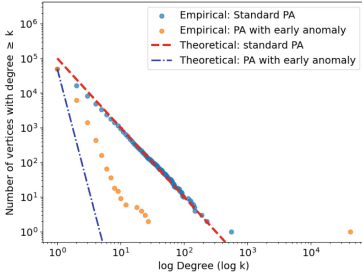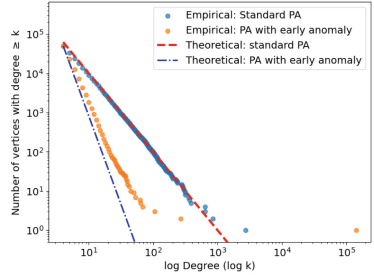

(a) $m = 1, \beta = 5.0, \delta = 0.$ (b) $m = 4, \beta = 10.0, \delta = 0.$

Fig. 5. The complementary cumulative degree distribution for PA network with early anomaly, parameters are $t = 50000, \tau = 50, \gamma = 0.3615.$

than 1 as predicted by (17). Interestingly, the experiments agree with formula (8.4.11) in [9, Chapter 8] multiplied by the factor $\alpha^{\frac{\beta}{2m+\beta+\delta}}$. Thus, our heuristic derivation correctly predicts the effect of a mid-way anomaly. The visible outlier in Fig. 4 is the anomaly itself.

In Fig. 5 for the early anomaly, the proportion of vertices with degree greater than k is significantly lower, as the anomaly attracts a fraction of edges from the very beginning. However, (18) predicts an even steeper slope. We also note that the tail of the distribution deviates strongly to the right from the straight line. We will come back to this phenomenon in the next section, where we analyze the behavior of the oldest ordinary vertex.

5   Degree Growth of the Oldest Vertex

Because the existing vertices connect to a new coming vertex with a probability proportional to its degree, it is more likely that the old vertices receive more

and more edges over time, and their degrees are of the order $t^{\frac{1}{2+\delta/m}}$ [9, Theorem 8.2]. We further explore how the degree of the oldest vertex behaves after the anomaly's occurrence, by computing the exponent of power-law degree of the initial vertex v_1. By Stirling's formula, as t and τ are large enough, the expected degree of the oldest vertex is

$$\mathbb{E}[D_1(t) + \delta] = (2m+\delta) \left(\frac{t}{\tau}\right)^{\frac{1}{2+\frac{\beta}{m}+\frac{\delta}{m}}} \tau^{\frac{1}{2+\frac{\delta}{m}}} (1+o(1)).$$

In different scenarios, as $t \to \infty$, the approximate expression are shown as follows:

Late anomaly: If $\tau = t - t^\gamma$, $\gamma \in (0,1)$,

$$\mathbb{E}[D_1(t) + \delta] = (2m+\delta)t^{\frac{1}{2+\frac{\delta}{m}}} (1+o(1)).$$

Mid-way anomaly: If $\tau = \alpha t$, $\alpha \in (0,1)$,

$$\mathbb{E}[D_1(t) + \delta] = (2m+\delta)\alpha^{\frac{1}{2+\frac{\delta}{m}} - \frac{1}{2+\frac{\beta}{m}+\frac{\delta}{m}}} t^{\frac{1}{2+\frac{\delta}{m}}} (1+o(1)).$$

Early anomaly: If $\tau = t^\gamma$, $\gamma \in (0,1)$,

$$\mathbb{E}[D_1(t) + \delta] = (2m+\delta)t^{\frac{\gamma}{2+\frac{\delta}{m}} + \frac{1-\gamma}{2+\frac{\beta}{m}+\frac{\delta}{m}}} (1+o(1)).$$

Interestingly, the mid-way and the early anomaly affect the mean degree of the oldest vertex in a different way. The mid-way anomaly reduces the mean degree by a constant factor, while an early anomaly changes the exponent of t. Moreover, in the case of an early anomaly, $\mathbb{E}[D_1(t) + \delta]$ grows more quickly than $t^{\frac{1}{2+\frac{\beta}{m}+\frac{\delta}{m}}}$. The latter expression would be consistent with the power-law distribution for an early anomaly in (18) meaning that the oldest ordinary vertices have higher degrees than predicted by (18); therefore, we see many outliers to the right in Fig. 5.

6   Conclusion and Further Research

We introduced a PA network incorporating an anomaly, where the attachment rule of the anomaly is vastly different from that of the ordinary vertices. We derived insights on the growth of the degrees and their distribution in this model. Below we list some directions for further research:

(1) **Concentration of the degree of an anomalous vertex.** We see in Fig. 2 that the degree of the anomaly is close to its mean. The martingale convergence theorem is a standard way to prove such concentration. However, we could not directly apply this method to the degree of the anomaly due to the presence of a linear term in its expression.

(2) **Convergence of degree sequences.** In [9, Chapter 8], the limiting degree distribution is derived from the recursive equations for the number of vertices of degree k. However, we could not use this method because the occurrence of an anomaly changes the recursion. In previous works [1, 6], the asymptotic degree distribution of PA networks with change points was rigorously derived using continuous-branching processes, which embed the growth of the PA network for $m=1$ in continuous time. Investigating whether these techniques can be adapted to our model is a promising direction for gaining deeper insights into the interplay between anomalies and degree distributions.

(3) **Different attachment mechanism.** In this paper, we consider only the case where the anomaly attracts new edges at a *constant* additional rate. An interesting direction for future research would be to explore scenarios where the anomaly attaches new edges at rates that change over time, such as an increasing rate or for a randomly determined duration.

(4) **Anomaly detection.** Building on our results, detecting anomalies is a natural next step. Existing work [13] has utilized Lyapunov-based method to detect certain anomalous events in PA network. Our model, where the anomaly arrives at a specific point and alters the attachment mechanisms, requires a different approach. While many studies on anomaly detection in dynamic networks focus on analyzing spatial and temporal [12, 14], or structural [10], features, we find it an interesting problem to detect the anomaly based only on the history of G_t. Our initial attempts show that this problem is more challenging than one could expect given how strong our anomaly is. We hope to report on the progress in the near future.

References

1. Banerjee, S., Bhamidi, S., Carmichael, I.: Fluctuation bounds for continuous time branching processes and evolution of growing trees with a change point. Ann. Appl. Probab. **33**(4), 2919–2980 (2023)
2. Barabási, A.L., Albert, R.: Emergence of scaling in random networks. Science **286**(5439), 509–512 (1999)
3. Berger, N., Borgs, C., Chayes, J., Saberi, A.: Asymptotic behavior and distributional limits of preferential attachment graphs. Ann. Probab. **42**(1), 1–40 (2014)
4. Berger, N., Borgs, C., Chayes, J., Saberi, A.: On the spread of viruses on the internet. In: SODA 2005: Proceedings of the Sixteenth Annual ACM-SIAM Symposium on Discrete Algorithms, pp. 301–310. Society for Industrial and Applied Mathematics, Philadelphia, PA, USA (2005)

5. Bet, G., Bogerd, K., Castro, R.M., van der Hofstad, R.: Detecting a late change-point in the preferential attachment model (2023). https://arxiv.org/abs/2310.02603
6. Bhamidi, S., Jin, J., Nobel, A.: Change point detection in network models: preferential attachment and long range dependence. Ann. Appl. Probab. **28**(1), 35–78 (2018)
7. Bhamidi, S., Steele, J.M., Zaman, T.: Twitter event networks and the Superstar model. Ann. Appl. Probab. **25**(5), 2462–2502 (2015)
8. Cirkovic, D., Wang, T., Zhang, X.: Likelihood-based inference for random networks with changepoints (2024). https://arxiv.org/abs/2206.01076
9. van der Hofstad, R.: Random Graphs and Complex Networks, vol. 1. Cambridge Series in Statistical and Probabilistic Mathematics, vol. 34. Cambridge University Press, Cambridge (2017)
10. Kipf, T.N., Welling, M.: Semi-supervised classification with graph convolutional networks. In: International Conference on Learning Representations (ICLR) (2017)
11. Litvak, N.: Randomness and structure in complex networks. Nieuw archief voor wiskunde **5**(24), 103–113 (2023)
12. Liu, Y., et al.: Anomaly detection in dynamic graphs via transformer. IEEE Trans. Knowl. Data Eng. **35**(12), 12081–12094 (2023)
13. Ruiz, D., Finke, J.: Lyapunov-based anomaly detection in preferential attachment networks. Int. J. Appl. Math. Comput. Sci. **29**(2), 363–373 (2019)
14. Zheng, L., Li, Z., Li, J., Li, Z., Gao, J.: Addgraph: anomaly detection in dynamic graph using attention-based temporal GCN. In: Proceedings of the Twenty-Eighth International Joint Conference on Artificial Intelligence, IJCAI 2019, pp. 4419–4425. International Joint Conferences on Artificial Intelligence Organization (2019)