MURI Update

Tiandong Wang

School of ORIE, Cornell University

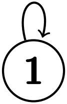
May 12th, 2017

Undirected Preferential Attachment Model

Notations:

- $G(n) :=$ the random graph after n-steps.
- $V(n) := \{1, 2, \ldots, n\}$, set of nodes in $G(n)$.
- $D_i(n) :=$ Degree of node $i \in V_n = \{1, 2, \ldots, n\}$.
- $\delta > -1$, parameter.

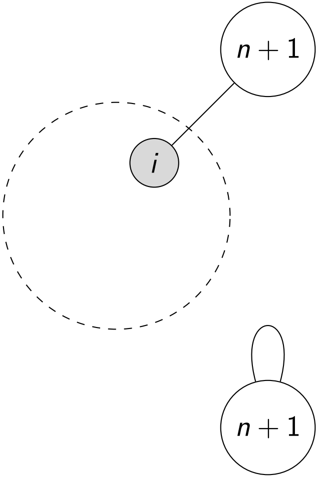
Initialize with a single node having a self loop.



This node is considered as having degree 2, i.e.

$$D_1(1) = 2.$$

From $G(n)$ to $G(n+1)$, consider two scenarios:



Assume **linear** preferential attachment function: $f(i) = i + \delta$.

(i) The new node $n+1$ attaches to node $i \in V_n$ with probability

$$\frac{D_i(n) + \delta}{(2+\delta)n + 1 + \delta},$$

and $D_{n+1}(n+1) = 1$.

(ii) The new node $n+1$ is born with a self loop and this happens with probability

$$\frac{1 + \delta}{(2+\delta)n + 1 + \delta},$$

and $D_{n+1}(n+1) = 2$.

What Is Known

Define $N_i(n) := \sum_{j=1}^{n} \mathbf{1}_{\{D_j(n)=i\}}$, then

(1) The degree counts converge to some deterministic limit:
As $n \to \infty$, $N_i(n)/n \to p_i$ with

$$p_i \sim C(\delta) i^{-3-\delta} \quad \text{for } i \to \infty.$$

(2) The degree sequence converges to some positive random variable:

$$\frac{D_i(n)}{n^{1/(2+\delta)}} \xrightarrow{a.s.} \xi_i.$$

(3) The maximum of the degree sequence also converges:

$$\max_{1 \le i \le n} \frac{D_i(n)}{n^{1/(2+\delta)}} \xrightarrow{a.s.} \max_{i \ge 1} \xi_i < \infty.$$

New Approach

Traditional approach: Use martingale convergence theorem.

Drawbacks:

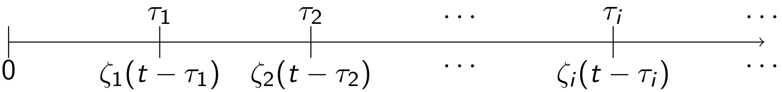
- Not much information on the limit quantity ξ_i has been provided.
- The traditional approach relies heavily on the preferential attachment function.
- For models with sub-linear preferential attachment functions, e.g. $f(i) = (i + \delta)^{-a}$, what will be the corresponding convergence results as in (1)(2)(3)??
- Direct calculations/traditional approaches do not provide much help.

Goal: Find a unified approach such that it can accommodate different types of preferential attachment assumptions.

Tentative solution: Model the degree growth of each single node as a birth process with immigration and find convergence results from the branching process.

Birth Process with Immigration

Let $0 < \tau_1 < \tau_2 < \dots$ be points of a homogeneous Poisson process of rate λ, then we initiate independent linear birth processes at $\tau_1, \tau_2, \dots$ as below: (cf. Tavaré(1987))



where $\zeta_i(\cdot)$ are independent linear birth processes with transition rate

$$q_{j,j+1} = j, \quad j = 1, 2, \dots.$$

Then

$$BI(t) = \sum_{i=1}^{\infty} \zeta_i(t-\tau_i)\mathbf{1}_{\{t \geq \tau_i\}}$$

is a birth process with immigration (BI process) and the transition rate is $q_{j,j+1} = j + \lambda$.

Model Construction

We formulate the linear preferential attachment model as below:

t

$T_1 = 0$
$BI_1(0) = 2$
$BI_2(0) = 1$

$T_2 = T_1 + m_2$
$\sum_{i=1}^{2} BI_i(T_2) = 4$

- At $T_1 = 0$, initiate two independent BI processes such that $BI_1(0) = 2$, $BI_2(0) = 1$ and transition rate $q_{j,j+1} = j + \delta$.
- Each process has an exponential clock which determines the time of their next jump: $\tau_1^{(1)} \sim \text{Expo}(2+\delta)$, $\tau_2^{(1)} \sim \text{Expo}(1+\delta)$, and $\tau_1^{(1)} \perp \tau_2^{(1)}$.
- Define $m_2 := \min\{\tau_1^{(1)}, \tau_2^{(1)}\}$, so $m_2 \sim \text{Expo}(3 + 2\delta)$.

$T_1 = 0$, $BI_1(0) = 2$, $BI_2(0) = 1$

$T_2 = T_1 + m_2$, $\sum_{i=1}^{2} BI_i(T_2) = 4$, $BI_3(0) = 1$

$T_3 = T_2 + m_3$, $\sum_{i=1}^{3} BI_i(T_3 - T_{i-1}) = 6$, $BI_4(0) = 1$

$\ldots$ t

- At $T_2 = T_1 + m_2$, one of BI_1 and BI_2 jumps so that the sum increases to 4.
- At the same time, initiate a new, independent BI process with $BI_3(0) = 1$ and transition rate $q_{j,j+1} = j + \delta$.

The construction process then goes on following this dynamic:
Given $G(n)$, we have (with the convention that $T_0 = T_1 = 0$)

$$\tau_i^{(n)} \sim \text{Expo}(BI_i(T_n - T_{i-1}) + \delta), \quad \tau_{n+1}^{(n)} \sim \text{Expo}(1 + \delta),$$

$$m_n = \bigwedge_{i=1}^{n+1} \tau_i^{(n)} \sim \text{Expo}\left(\sum_{i=1}^{n+1} (BI_i(T_n - T_{i-1}) + \delta)\right),$$

$$T_n = \sum_{i=2}^{n} m_i, \quad \{m_i, i \geq 2\} \text{ are all independent from each other.}$$

$\cdots$

$T_1 = 0$, $BI_1(0) = 2$, $BI_2(0) = 1$

$T_2 = T_1 + m_2$, $\sum_{i=1}^{2} BI_i(T_2) = 4$, $BI_3(0) = 1$

$T_3 = T_2 + m_3$, $\sum_{i=1}^{3} BI_i(T_3 - T_{i-1}) = 6$, $BI_4(0) = 1$

$\cdots t$

Define the degree sequence as

$$D_i(n) := \begin{cases} BI_i(T_n - T_{i-1}), & 1 \le i \le n \\ 0, & \text{otherwise.} \end{cases}$$

Check:

- $\sum_{i=1}^{n} D_i(n) = 2n$.
- Attachment probability:
  1. Node $i \in V_n$ is chosen:

$$\mathbb{P}\left(\tau_i^{(n)} < \bigwedge_{1\le k\le n+1, k\ne i} \tau_k^{(n)} \,\middle|\, G(n)\right) = \frac{BI_i(T_n - T_{i-1}) + \delta}{\sum_{i=1}^{n+1}(BI_i(T_n - T_{i-1}) + \delta)}$$

$$= \frac{D_i(n) + \delta}{(2+\delta)n + 1 + \delta}.$$

$\cdots$

$\mathrm{T}_1 = 0$, $\mathrm{BI}_1(0) = 2$, $\mathrm{BI}_2(0) = 1$

$\mathrm{T}_2 = T_1 + m_2$, $\sum_{i=1}^{2} BI_i(T_2) = 4$, $\mathrm{BI}_3(0) = 1$

$\mathrm{T}_3 = T_2 + m_3$, $\sum_{i=1}^{3} BI_i(T_3 - T_{i-1}) = 6$, $\mathrm{BI}_4(0) = 1$

$\cdots t$

2. Node $n+1$ is born with a self loop:

$$\mathbb{P}\left(\tau_{n+1}^{(n)} < \bigwedge_{1\le k\le n} \tau_k^{(n)} \middle| G(n)\right) = \frac{1+\delta}{\sum_{i=1}^{n+1}(BI_i(T_n - T_{i-1}) + \delta)}$$
$$= \frac{1+\delta}{(2+\delta)n + 1 + \delta}.$$

Asymptotics

Define the counting process

$$N(t) := \frac{1}{2}\sum_{i=1}^{\infty} BI_i(t - T_{i-1})\mathbf{1}_{\{t \geq T_i\}} - 1,$$

then it is a birth process with rate

$$q_{j,j+1} = (2+\delta)(j+1) + 1 + \delta,$$

and $N(0) = 0$.

Applying the result in Tavaré (1987) gives

$$e^{-(2+\delta)t} N(t) \xrightarrow{a.s.} \text{Gamma}\left(\frac{3+2\delta}{2+\delta}, 1\right) =: W,$$

so

$$n^{1/(2+\delta)} / e^{T_n} \xrightarrow{a.s.} W^{1/(2+\delta)}.$$

For each $i \geq 2$,

$$e^{-t} BI_i(t) \xrightarrow{a.s.} \text{Gamma}(1+\delta, 1) =: \sigma_i, \qquad t \to \infty.$$

Then

$$\frac{D_i(n)}{e^{T_n - T_{i-1}}} = \frac{BI_i(T_n - T_{i-1})}{e^{T_n - T_{i-1}}} \xrightarrow{a.s.} \sigma_i, \qquad n \to \infty.$$

Hence,

$$\frac{D_i(n)}{e^{T_n}} \xrightarrow{a.s.} \sigma_i e^{-T_{i-1}}, \qquad n \to \infty,$$

and $\sigma_i \perp T_{i-1}$.

Recall that $n^{1/(2+\delta)}/e^{T_n} \xrightarrow{a.s.} W^{1/(2+\delta)}$. Then it follows

$$\frac{D_i(n)}{n^{1/(2+\delta)}} \xrightarrow{a.s.} W^{-1/(2+\delta)} \sigma_i e^{-T_{i-1}}, \qquad n \to \infty.$$

Can also show that

$$\bigvee_{i=1}^{n} \frac{D_i(n)}{n^{1/(2+\delta)}} \xrightarrow{a.s.} W^{-1/(2+\delta)} \bigvee_{i\geq 1} \sigma_i e^{-T_{i-1}}, \qquad n \to \infty,$$

and

$$\sum_{i=1}^{n} \epsilon_{D_i(n)/n^{1/(2+\delta)}} \Rightarrow \sum_{i=1}^{\infty} \epsilon_{W^{-1/(2+\delta)}\sigma_i e^{-T_{i-1}}}$$

in $M_+(0,\infty]$.

$$\frac{D_i(n)}{n^{1/(2+\delta)}} \xrightarrow{a.s.} W^{-1/(2+\delta)} \sigma_i e^{-T_{i-1}}, \qquad n \to \infty.$$

Question:

(i) When $\delta = 0$, there is a representation in Bollobás (2003):

$$\frac{D_i(n)}{2n^{1/2}} \xrightarrow{a.s.} (Gamma(i,1))^{1/2} - (Gamma(i-1,1))^{1/2}.$$

How is it related to our result?

(ii) What does this tell about consistency of the Hill estimator?

(iii) Can we generalize this formulation to accommodate the directed linear PA model?
(Multi-type branching process??)

(iv) When the PA function is changed, how will it affect the asymptotic results?
(Will lead to a modification on the transition rate of the birth process.)