Properties for extreme-valued degrees in recursive trees

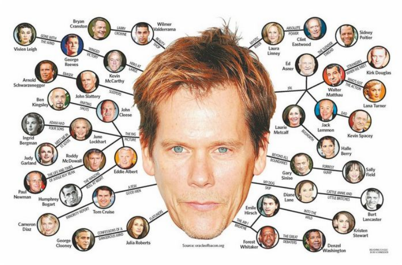
Laura Eslava

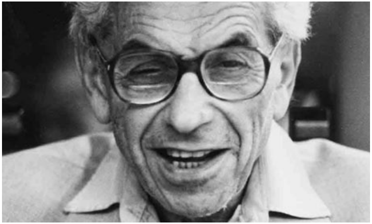
joint work with Louigi Addario-Berry

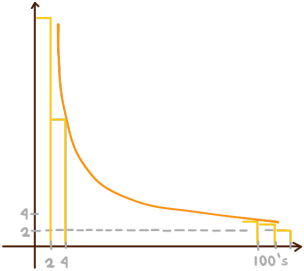
McGill University/Georgia Tech

AofA 2017

Motivation: Hubs in random networks

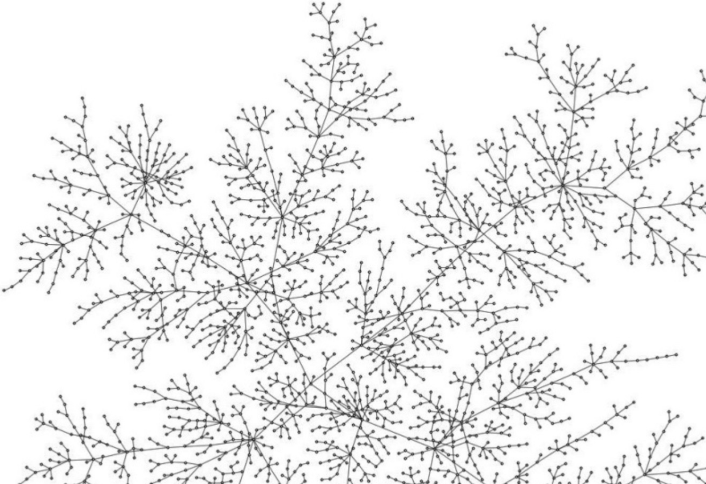






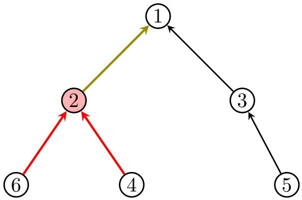
Tree growth processes

Image by Louigi A.-B.



Notation for trees: T

- Root / Leaves
- Children / Degree $\deg_T(\cdot)$
- Depth $\mathrm{ht}_T(\cdot)$ / Height
- Edges directed towards root.
- Vertices are labeled with $[n] = \{1, \ldots, n\}$.



$\deg_T(2) = 2$

$\mathrm{ht}_T(2) = 1$

Tree growth processes, $(T_n, n \in \mathbb{N})$

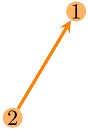
- T_1 is a single-vertex tree.
- For $n > 1$, build T_n from T_{n-1}

  adding: $\begin{cases} \text{vertex } n \\ \text{edge } n \to j \end{cases}$

$$\mathbb{P}(n \to j) = \frac{\beta \deg_{T_{n-1}}(j) + 1}{(\beta + 1)(n - 2) + 1}$$

Tree growth processes, $(T_n,\ n \in \mathbb{N})$

- T_1 is a single-vertex tree.
- For $n > 1$, build T_n from T_{n-1} adding: $\begin{cases} \text{vertex } n \\ \text{edge } n \to j \end{cases}$

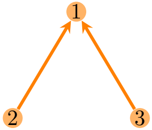


$$\mathbb{P}(n \to j) = \frac{\beta \deg_{T_{n-1}}(j) + 1}{(\beta + 1)(n - 2) + 1}$$

When $\beta > 0$: *The rich gets richer.*

Tree growth processes, $(T_n,\ n \in \mathbb{N})$

- T_1 is a single-vertex tree.
- For $n > 1$, build T_n from T_{n-1} adding: $\begin{cases} \text{vertex } n \\ \text{edge } n \to j \end{cases}$

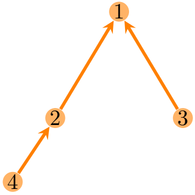


$$\mathbb{P}(n \to j) = \frac{\beta \deg_{T_{n-1}}(j) + 1}{(\beta + 1)(n - 2) + 1}$$

When $\beta > 0$: *The rich gets richer.*

Tree growth processes, $(T_n,\ n \in \mathbb{N})$

- T_1 is a single-vertex tree.
- For $n > 1$, build T_n from T_{n-1} adding: $\begin{cases} \text{vertex } n \\ \text{edge } n \to j \end{cases}$

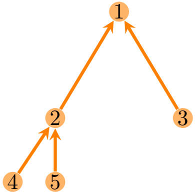


$$\mathbb{P}(n \to j) = \frac{\beta \deg_{T_{n-1}}(j) + 1}{(\beta + 1)(n-2) + 1}$$

When $\beta > 0$: *The rich gets richer.*

Tree growth processes, $(T_n,\ n \in \mathbb{N})$

- T_1 is a single-vertex tree.
- For $n > 1$, build T_n from T_{n-1}
  adding: $\begin{cases} \text{vertex } n \\ \text{edge } n \to j \end{cases}$



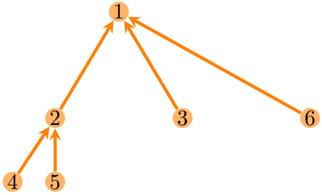
$$\mathbb{P}(n \to j) = \frac{\beta \deg_{T_{n-1}}(j) + 1}{(\beta+1)(n-2)+1}$$

When $\beta > 0$: *The rich gets richer.*

Tree growth processes, $(T_n,\ n \in \mathbb{N})$

- ▷ T_1 is a single-vertex tree.
- ▷ For $n > 1$, build T_n from T_{n-1} adding: $\begin{cases} \text{vertex } n \\ \text{edge } n \to j \end{cases}$

$$\mathbb{P}(n \to j) = \frac{\beta \deg_{T_{n-1}}(j) + 1}{(\beta+1)(n-2)+1}$$



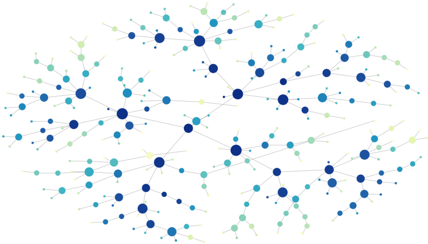
When $\beta > 0$: *The rich gets richer.*

$\alpha = 0$ 300 vertices

Recursive Trees (Uniform attach.)

$$\mathbb{P}(n \to j) = \frac{1}{n-1}$$

- Choices are **independent from the past**



$\alpha = 0$ 300 vertices $\alpha = 1$

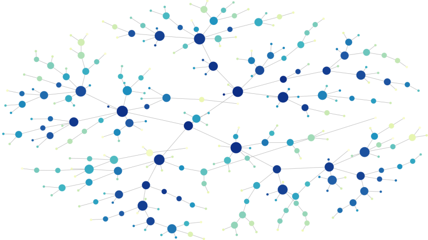
Recursive Trees (Uniform attach.)

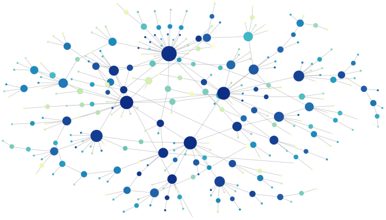
$$\mathbb{P}(n \to j) = \frac{1}{n-1}$$

Linear Pref. Attachment

$$\mathbb{P}(n \to j) = \frac{\deg_{T_{n-1}}(j) + 1}{2n-3}$$

- Choices are **independent from the past**

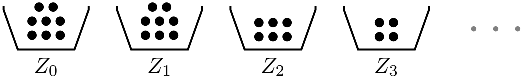




- **The rich gets richer**

Asymptotic normality ($n \to \infty$)

Degree distribution [Janson 05]:



$$Z_d(n) = \#\text{Vertices with degree } d \text{ in } T_n \sim N(c_d\, n, \sigma_d^2\, n)$$

Insertion depth [Devroye 88, Mahmoud 91-92]:

$$\mathrm{ht}_{T_n}(n) \sim N(c \ln n, c \ln n)$$

$$\alpha = 0 \to \begin{cases} \frac{1}{c_d} & = 2^{d+1} \\ c & = 1 \end{cases} \qquad \alpha = 1 \to \begin{cases} \frac{1}{c_d} & = (d+1)(d+2)(d+3) \\ c & = \frac{1}{2} \end{cases}$$

Maximum degree

Maximum degree Δ_n [Devroye, Lu 95, Mori 05]**:**

$$\alpha = 0: \quad \lim_{n\to\infty} \frac{\Delta_n}{\log n} = 1 \quad a.s. \quad \longleftarrow \log n \approx 1.4 \ln n$$

$$\alpha = 1: \quad \lim_{n\to\infty} \frac{\Delta_n}{\sqrt{n}} = D \quad a.s. \quad \longleftarrow \text{Continuous dist.}$$

Maximum degree

Maximum degree Δ_n [Devroye, Lu 95, Mori 05]**:**

$$\alpha = 0: \quad \lim_{n\to\infty} \frac{\Delta_n}{\log n} = 1 \quad a.s. \quad \longleftarrow \log n \approx 1.4 \ln n$$

$$\alpha = 1: \quad \lim_{n\to\infty} \frac{\Delta_n}{\sqrt{n}} = D \quad a.s. \quad \longleftarrow \text{Continuous dist.}$$

[Goh, Schmutz 2002] If T_n is a recursive tree, $n = 2^k$. For $i \in \mathbb{N}$ fixed

$$\mathbb{P}(\Delta_n - \log n < i) = \exp\{-2^{-i}\} + o(1).$$

More details for Linear Pref. Attachement

[Mori2005, Pekös, Röllin, Ross 2016]

Let $E_i \stackrel{\mathcal{L}}{=} \mathrm{Exp}(1)$ be iid. If T_n is a linear perf. attachment tree. As $n \to \infty$,

$$\left(\frac{\deg_{T_n}(i)}{\sqrt{n}}, \, i \geq 1 \right) \xrightarrow{\mathcal{L}} \left(B_i, i \geq 1 \right),$$

where for all $k \geq 2$,

$$\sum_{i=1}^{k} B_i \stackrel{\mathcal{L}}{=} \left(\sum_{i=1}^{k} E_i \right)^{1/2}$$

More details for Linear Pref. Attachement

[Mori2005, Pekös, Röllin, Ross 2016]

Let $E_i \overset{\mathcal{L}}{=} \mathrm{Exp}(1)$ be iid. If T_n is a linear perf. attachment tree. As $n \to \infty$,

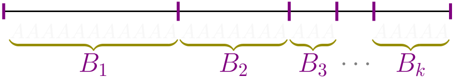
$$\left(\frac{\deg_{T_n}(i)}{\sqrt{n}}, \, i \geq 1 \right) \xrightarrow{\mathcal{L}} (B_i, i \geq 1) \, ,$$

where for all $k \geq 2$,

$$\sum_{i=1}^{k} B_i \overset{\mathcal{L}}{=} \left(\sum_{i=1}^{k} E_i \right)^{1/2}$$

What is the proportion for each B_j?

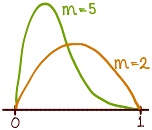
Stick-breaking process



$\sum_{i=1}^{k} B_i \overset{\mathcal{L}}{=} (\mathrm{Exp}(k))^{1/2}$

$$\frac{B_2}{B_1 + B_2} \overset{\mathcal{L}}{=} \mathrm{Beta}(2,2)$$

$$\frac{B_j}{B_1 + \cdots B_j} \overset{\mathcal{L}}{=} \mathrm{Beta}(2, 3j-4)$$



$\mathrm{Beta}(2, m)$

pdf: $x(1-x)^{m-1}$

- Each piece is broken independently

Extreme-valued degrees in recursive trees

Through Kingman's Coalescent

For a uniformly chosen vertex v, let $B_i \overset{\mathcal{L}}{=} Ber(2/i)$ be independent,

$$\mathcal{S} \overset{\mathcal{L}}{=} \sum_{i=2}^{n} B_i,$$

and draw $\mathcal{S}$ independent fair coin flips.

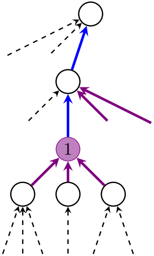
Proposition.

Depth = Total # unfavourable flips.

$$\mathrm{ht}_{F_n}(v) \overset{\mathcal{L}}{=} \mathrm{Bin}(|\mathcal{S}|, 1/2).$$

Degree = First streak fav. flips.

$$\deg_{F_n}(v) \overset{\mathcal{L}}{=} \min\{Geo(1/2), |\mathcal{S}|\}.$$



Summary of results

- **Poisson Point Process for near-maximum degree vertices**

  Number and their depth

- Conditional depth of high-degree vertices
- Tighten tails for maximum degree distribution ($\sim$Gumbel)

- **CLT's -rates of converge ($1 < c < \log e$)**

$$X_c = \{v \in [n], \deg_{T_n}(v) \geq c \ln n\}$$

Summary of results

- **Poisson Point Process for near-maximum degree vertices**

  Number and their depth

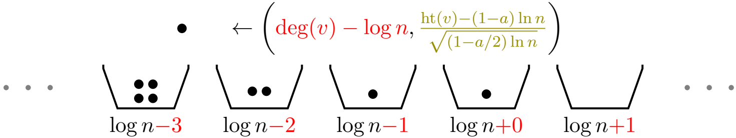
- Conditional depth of high-degree vertices
- Tighten tails for maximum degree distribution ($\sim$Gumbel)

- **CLT's -rates of converge ($1 < c < \log e$)**

$$X_c = \{v \in [n], \deg_{T_n}(v) \geq c \ln n\}$$

- High degrees of random recursive trees (joint with Louigi Addario-Berry). RSA 2017^{+}
- Depth of vertices with high degree in random recursive trees.
- Extremal values in recursive trees via a new tree growth process.

The marked point process



- Natural
  - # vertices with same excess degree k is Poisson(2^{-k-1})
  - Depths have Gaussian fluctuations
- Good news
  - '# vertices'/depths become independent variables
- Surprising
  - Process of **vertex-overtakes** to become max-degree

Comparing...

Recursive Trees [Addario-Berry,E. 17, E. 17$^+$]

"Maximum-degree vertices are in a constant race"

$$\bullet \qquad \leftarrow \left(\deg(v) - \log n, \frac{\mathrm{ht}(v) - (1-a)\ln n}{\sqrt{(1-a/2)\ln n}}\right)$$

There are $\approx n^a$ vertices with depth $\approx (1-a)\ln n$

Linear Pref. Attachment [Mori 05, Pekös, Röllin,Ross 16]

"Maximum degree vertices are immediately established"

$$\lim_{n\to\infty}\left(\frac{\deg_{T_n}(1)}{\sqrt{n}}, \frac{\deg_{T_n}(2)}{\sqrt{n}}, \frac{\deg_{T_n}(3)}{\sqrt{n}}, \ldots\right) = (B_1, B_2, B_3, \ldots)$$

Thanks!

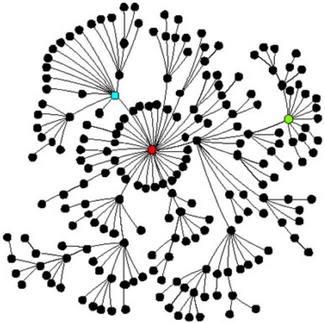


Image from scalefreenetworks, Flickr