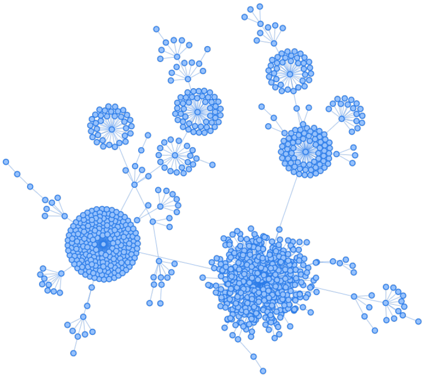
Random friend trees

Based on joint work with Louigi Addario-Berry, Simon Briend, Luc Devroye, Céline Kerriou and Gabor Lugosi

January 2024

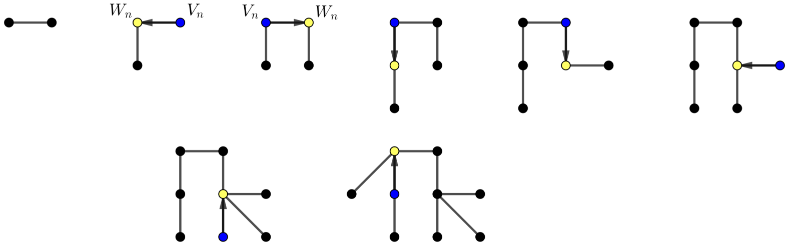
Introduction

- We study a random recursive tree model called *the random friend tree*
- Attachment via complete redirection
- 'Rich-getting-richer' dynamics
- Interesting emergent properties
- Local attachment rule



Definition of the model

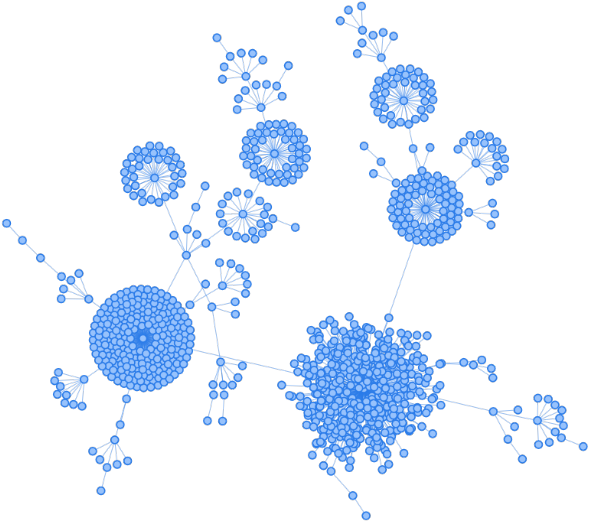
- Let T_2 be the tree consisting of vertices 1 and 2 connected via an edge.
- Given T_n, let V_n be a uniform vertex in T_n
- Let W_n be a uniform neighbour of V_n in T_n
- Include the vertex $n+1$ and the edge $\{n+1, W_n\}$ in T_n to obtain T_{n+1}



Earlier work

- 'Introduced' by Saramäki and Kaski (2004)
- Cannings and Jordan show that T_n contains $n - o(n)$ leaves almost surely (2014)
- More specific properties of the degree sequence were studied non-rigorously in the physics literature by Karpivsky and Redner (2017)

Intuition for dynamics



Results I - Hubs

For $u \in [n]$, let $N_n(u)$ be the number of neighbours of vertex u in T_n and let $L_n(u)$ be the number of leaf neighbours of vertex u in T_n.

Convergence of normalised degrees ⋆

There exists a random variable Z_u on $[0, \infty)$ such that

$$\left(\frac{L_n(u)}{n}, \frac{N_n(u)}{n} \right) \to (Z_u, Z_u) \text{ almost surely}$$

We call a vertex u a *hub* if $Z_u > 0$.

Results I - Hubs

For $u \in [n]$, let $N_n(u)$ be the number of neighbours of vertex u in T_n and let $L_n(u)$ be the number of leaf neighbours of vertex u in T_n.

Convergence of normalised degrees $\star$

There exists a random variable Z_u on $[0, \infty)$ such that

$$\left(\frac{L_n(u)}{n}, \frac{N_n(u)}{n} \right) \to (Z_u, Z_u) \text{ almost surely}$$

We call a vertex u a *hub* if $Z_u > 0$.

All edges contain a hub $\star$

For any edge $\{m+1, W_m\}$, it holds that

$$Z_{m+1} + Z_{W_m} > 0 \text{ almost surely}$$

Results II - Frozen degrees

Vertices never acquire a new neighbour with positive probability

For each k, there is a p_k so that for any n, any degree k vertex present at time n has probability at least p_k to *never* acquire another neighbour.

Results II - Frozen degrees

Vertices never acquire a new neighbour with positive probability

For each k, there is a p_k so that for any n, any degree k vertex present at time n has probability at least p_k to *never* acquire another neighbour.

Most leaves stay leaves forever

Suppose v is a hub. Then, then the number of leaves adjacent to v at time n that *ever* acquire another neighbour is tight.

Results III - Distances

Diameter $\star$

Let D_n be the diameter of T_n. Then,

$$1 \leq \liminf \frac{D_n}{\log n} \leq \limsup \frac{D_n}{\log n} \leq 4e.$$

Results III - Distances

Diameter ⋆

Let D_n be the diameter of T_n. Then,

$$1 \leq \liminf \frac{D_n}{\log n} \leq \limsup \frac{D_n}{\log n} \leq 4e.$$

Leaf depth

Let M_n be the maximal distance of any vertex to its nearest leaf in T_n. Then,

$$M_n = \Theta\left(\frac{\log n}{\log\log n}\right) \text{ in probability.}$$

Results IV - Degree sequence

Let $X_n^{\geq k}$ be the number of vertices in T_n with degree at least k.

Large-ish degrees

For any sequence $(m_n)_{n\geq 1}$ with $m_n = o(n)$,

$$\lim_{n\to\infty} X_n^{\geq m_n} = \infty.$$

Results IV - Degree sequence

Let $X_n^{\geq k}$ be the number of vertices in T_n with degree at least k.

Large-ish degrees

For any sequence $(m_n)_{n\geq 1}$ with $m_n = o(n)$,

$$\lim_{n\to\infty} X_n^{\geq m_n} = \infty.$$

Small degrees

There exist constants $0.1 < \alpha < \beta < 0.9$ such that for any $k \geq 2$, as $n \to \infty$,

$$\frac{X_n^{\geq k}}{n^\alpha} \to \infty \text{ and}$$
$$\frac{X_n^{\geq k}}{n^\beta} \to 0.$$

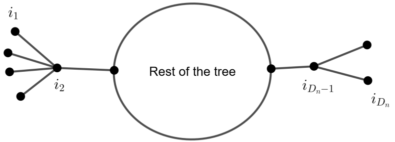
It is a slippery model to study!

- In preferential attachment trees and uniform attachment trees, the development of the sequence of vertex degrees, or even of a single vertex degree, can be studied without keeping track of the geometry of the tree.

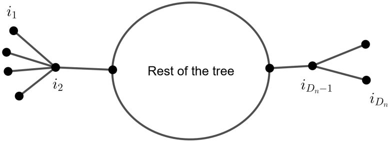
It is a slippery model to study!

- In preferential attachment trees and uniform attachment trees, the development of the sequence of vertex degrees, or even of a single vertex degree, can be studied without keeping track of the geometry of the tree.
- In the random friend tree, this is not possible.
- To know $\mathbb{E}[\Delta N_n(v) \mid T_n]$, you need to know the degrees of the neighbours of v. But to sample their development, you need to know the degrees of *their* neighbours, etc.
- More generally, in the random friend tree, the global structure affects local properties.

Proof lower bound diameter $\liminf D_n/n \geq 1$

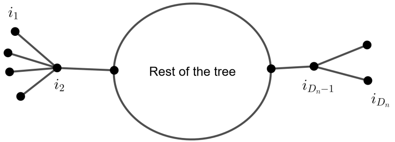


Proof lower bound diameter $\liminf D_n/n \geq 1$



$$\mathbb{P}(\Delta D_n = 1 \mid T_n) \geq \mathbb{P}(V_n = i_2, W_n \neq i_3) + \mathbb{P}(V_n = i_{D_n-1}, W_n \neq i_{D_n-2})$$
$$\geq \frac{1}{n} \times \frac{1}{2} + \frac{1}{n} \times \frac{1}{2} = \frac{1}{n}.$$

Proof lower bound diameter $\liminf D_n/n \geq 1$



$$\mathbb{P}(\Delta D_n = 1 \mid T_n) \geq \mathbb{P}(V_n = i_2, W_n \neq i_3) + \mathbb{P}(V_n = i_{D_n-1}, W_n \neq i_{D_n-2})$$
$$\geq \frac{1}{n} \times \frac{1}{2} + \frac{1}{n} \times \frac{1}{2} = \frac{1}{n}.$$

So D_n stochastically dominates $B_1 + \cdots + B_n$, for B_i a Bernoulli random variable with $\mathbb{P}(B_i = 1) = \frac{1}{i}$. Kolmogorov's strong law of large numbers implies that

$$\liminf \frac{D_n}{\log n} \geq 1 \text{ almost surely.}$$

Proof almost sure limit $L_n(v)/n, N_n(v)/n$.

$$\begin{aligned}\mathbb{E}[L_{n+1}(v) \mid T_n] &= L_n(v) - \mathbb{P}(V_n = v, W_n \text{ a leaf}) + \mathbb{P}(W_n = v) \\ &\geq L_n(v) - \mathbb{P}(V_n = v) + \mathbb{P}(W_n = v, V_n \text{ a leaf}) \\ &= L_n(v) - \frac{1}{n} + \frac{L_n(v)}{n}.\end{aligned}$$

Proof almost sure limit $L_n(v)/n, N_n(v)/n$.

$$\begin{aligned}\mathbb{E}[L_{n+1}(v) \mid T_n] &= L_n(v) - \mathbb{P}(V_n = v, W_n \text{ a leaf}) + \mathbb{P}(W_n = v) \\ &\geq L_n(v) - \mathbb{P}(V_n = v) + \mathbb{P}(W_n = v, V_n \text{ a leaf}) \\ &= L_n(v) - \frac{1}{n} + \frac{L_n(v)}{n}.\end{aligned}$$

Rearrangement yields that

$$\mathbb{E}\left[\frac{L_{n+1}(v) - 1}{n+1} \Big| T_n\right] \geq \frac{L_n(v) - 1}{n}.$$

This implies that $(L_n(v) - 1)/n$ is a submartingale

Proof almost sure limit $L_n(v)/n, N_n(v)/n$.

$$\begin{aligned}\mathbb{E}[L_{n+1}(v) \mid T_n] &= L_n(v) - \mathbb{P}(V_n = v, W_n \text{ a leaf}) + \mathbb{P}(W_n = v) \\ &\geq L_n(v) - \mathbb{P}(V_n = v) + \mathbb{P}(W_n = v, V_n \text{ a leaf}) \\ &= L_n(v) - \frac{1}{n} + \frac{L_n(v)}{n}.\end{aligned}$$

Rearrangement yields that

$$\mathbb{E}\left[\frac{L_{n+1}(v) - 1}{n+1} \middle| T_n\right] \geq \frac{L_n(v) - 1}{n}.$$

This implies that $(L_n(v) - 1)/n$ is a submartingale, so it has an almost sure limit Z_v. Then,

$$\frac{L_n(v)}{n} \leq \frac{N_n(v)}{n} \leq \frac{L_n(v) + X_n^{\geq 2}}{n}.$$

Joint convergence follows from $X_n^{\geq 2} = o(n^{0.9})$ almost surely.

Proof sketch $Z_u + Z_v > 0$ for $\{u, v\}$ an edge

Set $(u, v) = (m + 1, W_m)$ and for $n \geq m + 1$, write $N_n = N_n(u) + N_n(v)$ and $L_n = L_n(u) + L_n(v)$. We show that $\liminf L_n/n > 0$ almost surely.

Proof sketch $Z_u + Z_v > 0$ for $\{u, v\}$ an edge

Set $(u, v) = (m+1, W_m)$ and for $n \geq m+1$, write $N_n = N_n(u) + N_n(v)$ and $L_n = L_n(u) + L_n(v)$. We show that $\liminf L_n/n > 0$ almost surely. Observe that $\Delta(L_n, N_n) \in \{(1,1), (0,0), (-1,0)\}$

Proof sketch $Z_u + Z_v > 0$ for $\{u, v\}$ an edge

Set $(u, v) = (m+1, W_m)$ and for $n \geq m+1$, write $N_n = N_n(u) + N_n(v)$ and $L_n = L_n(u) + L_n(v)$. We show that $\liminf L_n/n > 0$ almost surely. Observe that $\Delta(L_n, N_n) \in \{(1,1), (0,0), (-1,0)\}$ and

$$\mathbb{P}\left(\Delta(L_n, N_n) = (1,1)\right) = \mathbb{P}(W_n = u) + \mathbb{P}(W_n = v) \geq \frac{1}{n}\left(L_n + \frac{1}{N_n}\right).$$

Proof sketch $Z_u + Z_v > 0$ for $\{u, v\}$ an edge

Set $(u, v) = (m+1, W_m)$ and for $n \geq m+1$, write $N_n = N_n(u) + N_n(v)$ and $L_n = L_n(u) + L_n(v)$. We show that $\liminf L_n/n > 0$ almost surely. Observe that $\Delta(L_n, N_n) \in \{(1,1), (0,0), (-1,0)\}$ and

$$\mathbb{P}\left(\Delta(L_n, N_n) = (1,1)\right) = \mathbb{P}(W_n = u) + \mathbb{P}(W_n = v) \geq \frac{1}{n}\left(L_n + \frac{1}{N_n}\right).$$

We also have

$$\mathbb{P}\left(\Delta(L_n, N_n) = (-1,0)\right) \leq \frac{\min(2, L_n)}{n}.$$

Proof sketch $Z_u + Z_v > 0$ for $\{u, v\}$ an edge

Set $(u, v) = (m+1, W_m)$ and for $n \geq m+1$, write $N_n = N_n(u) + N_n(v)$ and $L_n = L_n(u) + L_n(v)$. We show that $\liminf L_n/n > 0$ almost surely. Observe that $\Delta(L_n, N_n) \in \{(1,1), (0,0), (-1,0)\}$ and

$$\mathbb{P}\left(\Delta(L_n, N_n) = (1,1)\right) = \mathbb{P}(W_n = u) + \mathbb{P}(W_n = v) \geq \frac{1}{n}\left(L_n + \frac{1}{N_n}\right).$$

We also have

$$\mathbb{P}\left(\Delta(L_n, N_n) = (-1,0)\right) \leq \frac{\min(2, L_n)}{n}.$$

For $X_n \geq 0$ a random variable with increments in $\{0, 1\}$, $\mathbb{P}(\Delta X_n = 1) \geq X_n/n$ implies that X_n can be bounded from below by the black balls in a Pólya urn, so $\liminf X_n/n > 0$ almost surely would follow!

Further directions I - Extend results

- Does the degree of an older vertex stochastically dominate the degree of a younger vertex in T_n?

Further directions I - Extend results

- Does the degree of an older vertex stochastically dominate the degree of a younger vertex in T_n?
- Is the number of vertices of degree k of the same order as the number of vertices of degree at least k?

Further directions I - Extend results

- Does the degree of an older vertex stochastically dominate the degree of a younger vertex in T_n?
- Is the number of vertices of degree k of the same order as the number of vertices of degree at least k? Does it hold that $\limsup_{n\to\infty} \frac{X_n^{\geq k}}{X_n^{\geq 2}} \to 0$ as $k \to \infty$?

Further directions I - Extend results

- Does the degree of an older vertex stochastically dominate the degree of a younger vertex in T_n?
- Is the number of vertices of degree k of the same order as the number of vertices of degree at least k? Does it hold that $\limsup_{n\to\infty} \frac{X_n^{\geq k}}{X_n^{\geq 2}} \to 0$ as $k \to \infty$? Does $\frac{X_n^{\geq k}}{X_n^{\geq 2}}$ have some almost sure limit? (Krapivsky and Redner conjecture the limit is $\Theta(k^{-\mu})$ for $\mu \approx 0.566$)

Further directions I - Extend results

- Does the degree of an older vertex stochastically dominate the degree of a younger vertex in T_n?
- Is the number of vertices of degree k of the same order as the number of vertices of degree at least k? Does it hold that $\limsup_{n\to\infty} \frac{X_n^{\geq k}}{X_n^{\geq 2}} \to 0$ as $k \to \infty$? Does $\frac{X_n^{\geq k}}{X_n^{\geq 2}}$ have some almost sure limit? (Krapivsky and Redner conjecture the limit is $\Theta(k^{-\mu})$ for $\mu \approx 0.566$)
- What does the graph restricted to the small degree vertices look like?

Further directions II - Extend results

- Is there a γ such that $n^{-\gamma}X_n^{\geq 2} \to X$ almost surely for X a random variable on $(0, \infty)$?

Further directions II - Extend results

- Is there a γ such that $n^{-\gamma} X_n^{\geq 2} \to X$ almost surely for X a random variable on $(0, \infty)$?

A heuristic

For $X_n \geq 0$ a random variable with increments in $\{0, 1\}$ and $0 < \gamma \leq 1$, $\mathbb{P}(\Delta X_n = 1) = \frac{\gamma}{n} X_n$ implies that $X_n = \Theta(n^{\gamma})$ almost surely.

Further directions II - Extend results

- Is there a γ such that $n^{-\gamma} X_n^{\geq 2} \to X$ almost surely for X a random variable on $(0, \infty)$?

A heuristic

For $X_n \geq 0$ a random variable with increments in $\{0, 1\}$ and $0 < \gamma \leq 1$, $\mathbb{P}(\Delta X_n = 1) = \frac{\gamma}{n} X_n$ implies that $X_n = \Theta(n^\gamma)$ almost surely.

$$\mathbb{P}(\Delta X_n^{\geq 2} = 1 \mid T_n) = \frac{1}{n} \sum_{v : N_n(v) \geq 2} \frac{L_n(v)}{N_n(v)}$$

Further directions II - Extend results

- Is there a γ such that $n^{-\gamma} X_n^{\geq 2} \to X$ almost surely for X a random variable on $(0, \infty)$?

A heuristic

For $X_n \geq 0$ a random variable with increments in $\{0, 1\}$ and $0 < \gamma \leq 1$, $\mathbb{P}(\Delta X_n = 1) = \frac{\gamma}{n} X_n$ implies that $X_n = \Theta(n^\gamma)$ almost surely.

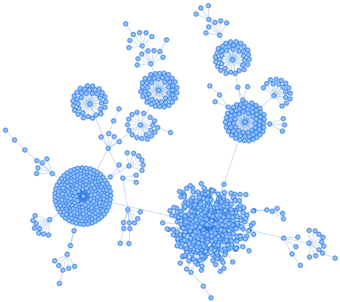
$$\mathbb{P}(\Delta X_n^{\geq 2} = 1 \mid T_n) = \frac{1}{n} \sum_{v : N_n(v) \geq 2} \frac{L_n(v)}{N_n(v)} = \frac{1}{n} \left(\frac{1}{X_n^{\geq 2}} \sum_{v : N_n(v) \geq 2} \frac{L_n(v)}{N_n(v)} \right) X_n^{\geq 2}.$$

Can we show that the average proportion of leaf neighbours across non-leaf vertices converges almost surely?

Krapivsky and Redner conjecture that $\gamma = \mu \approx 0.566$

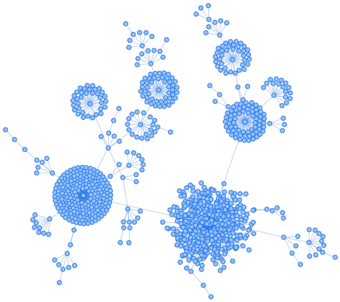
Further directions III - A potential master key

Intuitively, the different parts of the tree should 'decouple' at the hubs. The hubs cut the tree in (almost) independent parts that, up to a time change, evolve like rooted random friend trees.



Further directions III - A potential master key

Intuitively, the different parts of the tree should 'decouple' at the hubs. The hubs cut the tree in (almost) independent parts that, up to a time change, evolve like rooted random friend trees.



If we can make this formal, averaging across the tree allows for use of the law of large numbers, bringing many more complicated questions within reach.

Further directions IV - Generalisations

- W_n can be viewed as the end point of a random walk of length 1 on T_n started at V_n. What happens if we increase the length of the random walk to k?

Further directions IV - Generalisations

- W_n can be viewed as the end point of a random walk of length 1 on T_n started at V_n. What happens if we increase the length of the random walk to k?
- We may consider partial redirection, where $n+1$ connects to V_n instead of W_n with probability $p > 0$.

Further directions IV - Generalisations

- W_n can be viewed as the end point of a random walk of length 1 on T_n started at V_n. What happens if we increase the length of the random walk to k?
- We may consider partial redirection, where $n+1$ connects to V_n instead of W_n with probability $p > 0$.
- We can let new vertices connect to multiple vertices.