Large degrees in weighted recursive graphs with bounded random weights

Bas Lodewijks

Institut Camille Jordan, Université Lyon 1, Université Saint-Étienne

Workshop on geometric random graph models

Joint work with Marcel Ortgiese and Laura Eslava

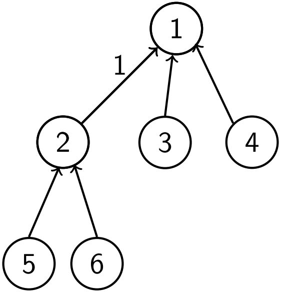
October 19, 2021

Motivation

- Random recursive tree, directed acyclic graph
- Egalitarian
- Introduce more heterogeneity
- Applicability
- More rich and diverse behaviour

Weighted recursive graph

Let $(W_i)_{i\in\mathbb{N}}$ be i.i.d. copies of a non-negative random variable W.
Assign vertex i (vertex-)weight W_i, $i \in \mathbb{N}$.



$$\mathbb{P}(n+1 \to i \mid (W_j)_{j\in\mathbb{N}}) = \frac{W_i}{\sum_{j=1}^{n} W_j}.$$

WRG: new vertices connect to $m \in \mathbb{N}$ predecessors.

Weighted recursive graph

If $W \leq x_0$ almost surely, without loss of generality $x_0 = 1$, as

$$\frac{W_i}{\sum_{j=1}^n W_j} = \frac{W_i/x_0}{\sum_{j=1}^n W_j/x_0} = \frac{\widetilde{W}_i}{\sum_{j=1}^n \widetilde{W}_j},$$

with $\widetilde{W}_i \leq 1$.

Degree evolutions

$\mathcal{Z}_n(i) :=$ in-degree of vertex i at step n.
How do the $(\mathcal{Z}_n(i))_{i\in[n]}$ behave as $n \to \infty$?

- Empirical degree distribution: $p_n(k) := \frac{1}{n}\sum_{i=1}^n \mathbb{1}_{\{\mathcal{Z}_n(i)=k\}}$, convergence and asymptotic behaviour.
- Maximum degree: growth rate & 'age'.

Results for the random recursive tree/directed acyclic graph/weighted recursive tree

- Degree distribution:
  - $p(k) := 2^{-(k+1)}, k \in \mathbb{N}_0$ (Meir, Moon, 1988)
  - Asymptotic normality of $\#$ degree k vertices, $k \in \mathbb{N}_0$ (Mahmoud, Smythe, 1992, Janson, 2005)
  - Degree distribution for large class of weighted evolving tree models (Iyer, 2020)
- Maximum degree:
  - $\max_{i\in[n]} \mathcal{Z}_n(i)/\log_{1+1/m}(n) \xrightarrow{a.s.} 1$, $m \in \mathbb{N}$ (Devroye, Lu, 1995)
  - $\max_{i\in[n]} \mathcal{Z}_n(i) - \lfloor \log_2(n) \rfloor$ converges in distribution along subsequences to a Poisson limit (Addario-Berry, Eslava, 2015)
  - Asymptotic normality of 'near-maximum' degrees: $\#$ vertices with degree $\lfloor \log_2 n \rfloor - i_n, \quad i_n \to \infty, i_n = o(\log n)$ (Addario-Berry, Eslava, 2015)
- Labels of maximum degree vertices:
  - Labels of maximum degree vertices in RRT are $n^{(1-1/(2\log 2))(1+o(1))}$ w.h.p. (Bannerjee, Bhamidi, 2020)

Results for the random recursive tree/directed acyclic graph/weighted recursive tree

- Degree distribution:
  - $p(k) := 2^{-(k+1)}, k \in \mathbb{N}_0$ (Meir, Moon, 1988)
  - Asymptotic normality of # degree k vertices, $k \in \mathbb{N}_0$ (Mahmoud, Smythe, 1992, Janson, 2005)
  - Degree distribution for large class of weighted evolving tree models (Iyer, 2020)
- Maximum degree:
  - $\max_{i\in[n]} \mathcal{Z}_n(i)/\log_{1+1/m}(n) \xrightarrow{a.s.} 1,\ m \in \mathbb{N}$ (Devroye, Lu, 1995)
  - $\max_{i\in[n]} \mathcal{Z}_n(i) - \lfloor \log_2(n) \rfloor$ converges in distribution along subsequences to a Poisson limit (Addario-Berry, Eslava, 2015)
  - Asymptotic normality of 'near-maximum' degrees: # vertices with degree $\lfloor \log_2 n \rfloor - i_n, \quad i_n \to \infty, i_n = o(\log n)$ (Addario-Berry, Eslava, 2015)
- Labels of maximum degree vertices:
  - Labels of maximum degree vertices in RRT are $n^{(1-1/(2\log 2))(1+o(1))}$ w.h.p. (Bannerjee, Bhamidi, 2020)

Degree distribution

Theorem (WRG degree distribution, L., Ortgiese, '20+)

Consider the WRG with i.i.d. bounded vertex-weights $(W_i)_{i\in\mathbb{N}}$ and out-degree $m \in \mathbb{N}$. Set $\theta_m := 1 + \mathbb{E}[W]/m$. Then, for any $k \in \mathbb{N}_0$,

$$p_n(k) \xrightarrow{a.s.} \mathbb{E}\left[\frac{\theta_m - 1}{\theta_m - 1 + W}\left(\frac{W}{\theta_m - 1 + W}\right)^k\right] =: p(k).$$

Set $m = 1$ (WRT). Assume $\mathbb{P}(W \geq w^*) = 1$ for some $w^* \in (0,1)$. Let

$$p_{\geq}(k) := \sum_{j=k}^{\infty} p(k) = \mathbb{E}\left[\left(\frac{W}{\theta - 1 + W}\right)^k\right]. \qquad (\theta := \theta_1)$$

Theorem (Eslava, L., Ortgiese, '21+)

Let $L \in \mathbb{N}, c \in (0, \theta/(\theta-1))$, $v_1, \ldots, v_L$ vertices selected u.a.r. from $[n]$. There exists a $\beta > 0$, such that for $k_1(n), \ldots, k_L(n) < c \log n$,

$$\mathbb{P}\left(\mathcal{Z}_n(v_\ell) \geq k_\ell, \ell \in [L]\right) = \prod_{\ell=1}^{L} p_{\geq}(k_\ell)\left(1 + o\left(n^{-\beta}\right)\right).$$

Maximum degree: first order

Theorem (WRG maximum degree, first order, L., Ortgiese, '20+)

Consider the WRG model with i.i.d. bounded vertex-weights $(W_i)_{i\in\mathbb{N}}$ *and out-degree* $m \in \mathbb{N}$*. Let* $\theta_m := 1 + \mathbb{E}[W]/m$*. Then,*

$$\max_{i\in[n]} \frac{\mathcal{Z}_n(i)}{\log_{\theta_m} n} \xrightarrow{a.s.} 1.$$

Theorem (WRG maximum degree location, L., '21+)

Consider the WRG model with i.i.d. bounded vertex-weights $(W_i)_{i\in\mathbb{N}}$ *and out-degree* $m \in \mathbb{N}$*. Let* $\theta_m := 1 + \mathbb{E}[W]/m$*, and set* $I_n := \inf\{i \in [n] : \mathcal{Z}_n(i) \geq \mathcal{Z}_n(j) \text{ for all } j \in [n]\}$*. Then,*

$$\frac{\log I_n}{\log n} \xrightarrow{a.s.} 1 - \frac{\theta_m - 1}{\theta_m \log \theta_m}. \qquad \left(I_n = n^{\left(1-\frac{\theta_m-1}{\theta_m \log \theta_m}\right)(1+o(1))}\right)$$

Maximum degree: higher order

Higher order corrections of maximum degree for WRT model ($m = 1$).
Additional (technical) assumption: $\mathbb{P}(W \geq w^*) = 1$ for some $w^* \in (0, 1)$.
Distinguish different classes of vertex-weight distributions.

Atom $\mathbb{P}(W = 1) = q_0 \in (0, 1]$.

Weibull $\mathbb{P}(W \geq 1 - 1/x) = \ell(x)x^{-(\alpha-1)}$, $x > 1$, $\alpha > 1$, ℓ slowly varying.

Gumbel Distinguish two sub-cases:

RV $\mathbb{P}(W \geq 1 - 1/x) = ax^b \mathrm{e}^{-(x/c_1)^\tau}(1 + o(1))$ as $x \to \infty$,
$a, c_1, \tau > 0, b \in \mathbb{R}$.

RaV $\mathbb{P}(W \geq 1 - 1/x) = a(\log x)^b \mathrm{e}^{-(\log(x)/c_1)^\tau}(1 + o(1))$ as $x \to \infty$,
$a, c_1 > 0, \tau > 1, b \in \mathbb{R}$.

RV: $\frac{1}{1-W}$ is e.g. |normal|, gamma, chi-squared, etc.
RaV: $\frac{1}{1-W}$ is e.g. log-normal, log chi-squared, etc.

Atom class

Theorem (Eslava, L., Ortgiese, '21+)

Consider the WRT *model with i.i.d. bounded vertex-weights* $(W_i)_{i\in\mathbb{N}}$ *in the* Atom *class* $(\mathbb{P}(W=1)=q_0\in(0,1])$. *Let* $\theta:=\theta_1=1+\mathbb{E}[W]$ *and fix* $i,j\in\mathbb{Z}, i<j$. *Define*

$$
\begin{aligned}
X_k^{(n)} &:= |\{i\in[n] : \mathcal{Z}_n(i)=\lfloor\log_\theta n\rfloor+k\}|, \quad k\in\mathbb{Z}\\
X_{\geq k}^{(n)} &:= |\{i\in[n] : \mathcal{Z}_n(i)\geq\lfloor\log_\theta n\rfloor+k\}|, \quad k\in\mathbb{Z}\\
\epsilon_n &:= \log_\theta n-\lfloor\log_\theta n\rfloor.
\end{aligned}
$$

Then, if $(n_\ell)_{\ell\in\mathbb{N}}$ *such that* $\epsilon_{n_\ell}\to\epsilon$ *as* $\ell\to\infty$,

$$
(X_i^{(n_\ell)},X_{i+1}^{(n_\ell)},\ldots,X_{j-1}^{(n_\ell)},X_{\geq j}^{(n_\ell)})\xrightarrow{d}(P_i^\epsilon,P_{i+1}^\epsilon,\ldots,P_{j-1}^\epsilon,P_{\geq j}^\epsilon)\ \text{as}\ \ell\to\infty,
$$

where $P_i^\epsilon\sim Poi(q_0(1-\theta^{-1})\theta^{-i+\epsilon}), P_{\geq j}^\epsilon\sim Poi(q_0\theta^{-j+\epsilon})$.

Weibull class

Theorem (Eslava, L., Ortgiese, '21+)

Consider the WRT *model with i.i.d. bounded vertex-weights* $(W_i)_{i\in\mathbb{N}}$ *in the* Weibull *class* ($\mathbb{P}\left(W \geq 1-1/x\right) = \ell(x)x^{-(\alpha-1)}, x>1$)*. Let* $\theta := \theta_1 = 1+\mathbb{E}\left[W\right]$*. Then,*

$$\max_{i\in[n]} \frac{\mathcal{Z}_n(i)-\log_\theta n}{\log_\theta \log_\theta n} \xrightarrow{\mathbb{P}} -(\alpha-1).$$

We expect a random third order, similar to the Atom class.

Gumbel class

Theorem (Eslava, L., Ortgiese, '21+)

Consider the WRT model with i.i.d. bounded vertex-weights $(W_i)_{i\in\mathbb{N}}$ in the Gumbel class. Let $\theta := \theta_1 = 1 + \mathbb{E}[W]$. If the weights satisfy the RV sub-case, i.e.

$$\mathbb{P}(W \geq 1 - 1/x) = ax^b \mathrm{e}^{-(x/c_1)^\tau}(1 + o(1)) \text{ as } x \to \infty, a, c_1, \tau > 0, b \in \mathbb{R}.$$

Then, with $\gamma := 1/(\tau+1)$,

$$\max_{i\in[n]} \frac{\mathcal{Z}_n(i) - \log_\theta n}{(\log_\theta n)^{1-\gamma}} \xrightarrow{\mathbb{P}} -\frac{\tau^\gamma}{(1-\gamma)\log\theta}\Big(\frac{1-\theta^{-1}}{c_1}\Big)^{1-\gamma}.$$

We expect $\lceil \tau \rceil$ higher order terms and then a random order term, similar to the Atom class.

Gumbel class

Theorem (Eslava, L., Ortgiese, '21+)

Under the same assumptions, if the weights satisfy the RaV sub-case, i.e.

$$\mathbb{P}\left(W \geq 1 - 1/x\right) = a(\log x)^b \mathrm{e}^{-(\log(x)/c_1)^\tau}(1 + o(1)) \text{ as } x \to \infty,$$

$a, c_1 > 0, \tau > 1, b \in \mathbb{R}$. *Then, with*

$$C_1 := (\log\theta)^{\tau-1} c_1^{-\tau}, \quad C_2 := \tau(\tau - 1)C_1,$$

$$C_3 := \big(\log_\theta(\log\theta)(\tau - 1)\log\theta - \log(\mathrm{e}c_1^\tau(1 - \theta^{-1})/\tau)\big)\tau(\log\theta)^{\tau-2}c_1^{-\tau},$$

we have

$$\max_{i\in[n]} \frac{\mathcal{Z}_n(i) - (\log_\theta n - C_1(\log_\theta \log_\theta n)^\tau + C_2(\log_\theta \log_\theta n)^{\tau-1} \log_\theta \log_\theta \log_\theta n)}{(\log_\theta \log_\theta n)^{\tau-1}}$$

$$\xrightarrow{\mathbb{P}} C_3.$$

Idea of the proofs: first order

Set $S_n := \sum_{j=1}^n W_j$, and

$$f(x) := \frac{1}{\log\theta_m}\Big(\frac{(1-x)\log\theta_m}{\theta_m-1} - 1 - \log\Big(\frac{(1-x)\log\theta_m}{\theta_m-1}\Big)\Big), \qquad x\in(0,1).$$

- f has unique fixed point $\gamma_m := 1-\frac{\theta_m-1}{\theta_m\log\theta_m}$.
- $f(x) > x$ for all $x\in(0,1)\backslash\{\gamma_m\}$.

$$\begin{aligned}\mathbb{P}\left(\mathcal{Z}_n(i)\geq \log_{\theta_m} n \,|\, (W_k)_{k\in\mathbb{N}}\right) &\leq \mathrm{e}^{-t\log_{\theta_m} n}\prod_{j=i+1}^{n}\mathbb{E}\left[\mathrm{e}^{t\mathbb{1}_{j\to i}}\,|\,(W_k)_{k\in\mathbb{N}}\right]^m\\ &\leq \mathrm{e}^{-t\log_{\theta_m} n}\prod_{j=i}^{n-1}\exp\Big(m\big(\mathrm{e}^t-1\big)\frac{W_i}{S_j}\Big)\\ &= \mathrm{e}^{-\log n(u_i-1-\log u_i)/\log\theta_m},\end{aligned}$$

where

$$u_i := \frac{mW_i}{\log_{\theta_m} n}\sum_{j=i}^{n-1}\frac{1}{S_j}.$$

Idea of the proofs: first order

$$\mathbb{P}\left(\mathcal{Z}_n(i) \geq \log n \,\middle|\, (W_k)_{k\in\mathbb{N}}\right) \leq \mathrm{e}^{-\log n(u_i-1-\log u_i)/\log\theta_m},$$

where

$$u_i := \frac{mW_i}{\log_{\theta_m} n}\sum_{j=i}^{n-1}\frac{1}{S_j}.$$

Consider $i \sim n^\beta$ for some $\beta \in (0,1)$. Then, almost surely,

$$\begin{aligned} u_i &\leq \frac{m}{\log_{\theta_m} n}\sum_{j=i}^{n-1}\frac{1}{S_j} = \frac{m\log\theta_m}{\log n}\frac{\log(n/i)}{\mathbb{E}\left[W\right]}(1+o(1)) \\ &= \frac{\log\theta_m}{\theta_m-1}\frac{\log(n/n^\beta)}{\log n}(1+o(1)) = \frac{(1-\beta)\log\theta_m}{\theta_m-1}(1+o(1)). \end{aligned}$$

$$\begin{aligned} \mathbb{P}\left(\max_{i\sim n^\beta}\mathcal{Z}_n(i) \geq \log_{\theta_m} n \,\middle|\, (W_k)_{k\in\mathbb{N}}\right) &\leq n^\beta \mathrm{e}^{-f(\beta)\log n(1+o(1))} \\ &= \mathrm{e}^{\log n(\beta-f(\beta))(1+o(1))}. \end{aligned}$$

Idea of the proofs: first order

$$\mathbb{P}\left(\max_{i\sim n^\beta}\mathcal{Z}_n(i)\geq \log_{\theta_m} n \,\middle|\, (W_k)_{k\in\mathbb{N}}\right)\leq \mathrm{e}^{\log n(\beta-f(\beta))(1+o(1))}.$$

- f has unique fixed point $\gamma_m := 1-\frac{\theta_m-1}{\theta_m\log\theta_m}$.
- $f(\beta)>\beta$ for all $\beta\in(0,1)\backslash\{\gamma_m\}$, decreasing on $(0,1)$.

Probability $\downarrow 0$ for $\beta\neq\gamma_m$, almost surely.

Idea of the proofs: higher order

Recall

$$p_{\geq}(k) := \mathbb{E}\left[\left(\frac{W}{\theta - 1 + W}\right)^k\right].$$

Theorem (WRT degree distribution, Eslava, L., Ortgiese, '21+)

Let $L \in \mathbb{N}, c \in (0, \theta/(\theta-1))$, $v_1, \ldots, v_L$ vertices selected u.a.r. from $[n]$. There exists a $\beta > 0$, such that for $k_1(n), \ldots, k_L(n) < c \log n$,

$$\mathbb{P}\left(\mathcal{Z}_n(v_\ell) \geq k_\ell, \ell \in [L]\right) = \prod_{\ell=1}^{L} p_{\geq}(k_\ell)\big(1 + o\big(n^{-\beta}\big)\big).$$

Maximum degree d_n satisfies $p_{\geq}(d_n) \approx \frac{1}{n}$. More precisely,

$$\max_{i\in[n]} \mathcal{Z}_n(i) \geq d_n \;\Leftarrow np_{\geq}(d_n) \to \infty,$$

$$\max_{i\in[n]} \mathcal{Z}_n(i) \leq d_n \;\Leftarrow np_{\geq}(d_n) \to 0.$$

Idea of the proofs: higher order

- **Weibull**: $\underline{L}(k)k^{-(\alpha-1)}\theta^{-k} \leq p_{\geq}(k) \leq \overline{L}(k)k^{-(\alpha-1)}\theta^{-k}$, $\underline{L}, \overline{L}$ slowly varying.
- **Gumbel**:

$$\text{RV} \quad p_{\geq}(k) = \exp\Big(-\tfrac{\tau^\gamma}{1-\gamma}\Big(\tfrac{(1-\theta^{-1})k}{c_1}\Big)^{1-\gamma}(1+o(1))\Big)\theta^{-k}.$$

$$\text{RaV} \quad p_{\geq}(k) = \exp\Big(-\Big(\tfrac{\log k}{c_1}\Big)^{\tau}\Big(1-\tau(\tau-1)\tfrac{\log\log k}{\log k}+\tfrac{K_{\tau,c_1,\theta}}{\log k}(1+o(1))\Big)\theta^{-k},$$

$$K_{\tau,c_1,\theta} := \tau\log(\mathrm{e}c_1^\tau(1-\theta^{-1})/\tau).$$

Method of Moments for Poisson limits in **Atom** case.
Atom: $p_{\geq}(k) = q_0\theta^{-k}(1+o(1))$.
For any $i<j\in\mathbb{Z}$ and $K\in\mathbb{N}_0$, and any $a_i,\ldots,a_j$ such that $\sum_{k=i}^j a_k = K$,

$$\mathbb{E}\left[\left(X_{\geq j}^{(n_\ell)}\right)_{a_j}\prod_{k=i}^{j-1}\left(X_k^{(n_\ell)}\right)_{a_k}\right] \to \mathbb{E}\left[\left(P_{\geq j}^{\epsilon}\right)_{a_j}\prod_{k=i}^{j-1}\left(P_k^{\epsilon}\right)_{a_k}\right],$$

when $\epsilon_{n_\ell}\to\epsilon\in[0,1]$.

Thank you for your attention!