Persistence and root finding algorithms in dynamic random networks

Sayan Banerjee

University of North Carolina, Chapel Hill

Based on joint works with Shankar Bhamidi (UNC, Chapel Hill) and Xiangying Huang (Duke)

Motivating question

Consider a dynamic network where nodes attach one by one via some stochastic mechanism that depends on relative age of vertices.

We observe a snapshot of the unlabelled network $\mathcal{T}_n$ when its size is n. Where is the root? (Online detection)

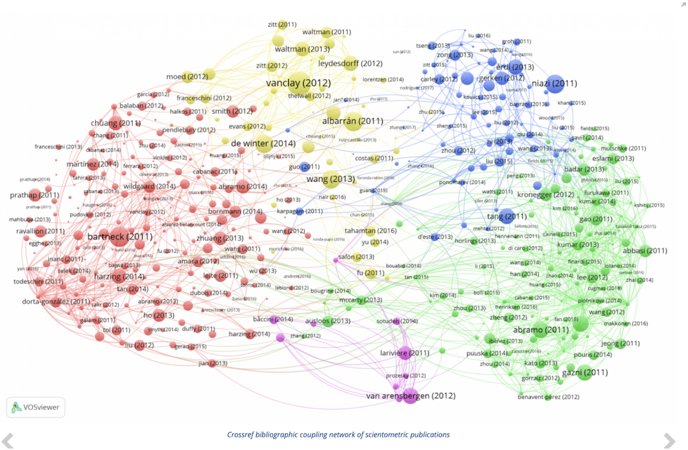


Figure: Visualizing freely available citation data using VOSviewer by Nees Jan van Eck, Ludo Waltman

Formalizing the question

- For any error tolerance $\epsilon \in (0, 1)$, produce a confidence set $S_n(\epsilon)$ based on $\mathcal{T}_n$ such that

$$\mathbb{P}\left(\text{root} \in S_n(\epsilon)\right) \geqslant 1 - \epsilon.$$

Formalizing the question

- For any error tolerance $\epsilon \in (0,1)$, produce a confidence set $S_n(\epsilon)$ based on $\mathcal{T}_n$ such that

$$\mathbb{P}\left(\text{root} \in S_n(\epsilon)\right) \geqslant 1 - \epsilon.$$

- Quantify dependence of $S_n(\epsilon)$ on ϵ, n and the geometry of $\mathcal{T}_n$.

Formalizing the question

- For any error tolerance $\epsilon \in (0,1)$, produce a confidence set $S_n(\epsilon)$ based on $\mathcal{T}_n$ such that

$$\mathbb{P}\left(\text{root} \in S_n(\epsilon)\right) \geqslant 1 - \epsilon.$$

- Quantify dependence of $S_n(\epsilon)$ on ϵ, n and the geometry of $\mathcal{T}_n$.

- When can the size of $S_n(\epsilon)$ be chosen independently of n (algorithm is stable in network size)?

Main Strategy

- Centrality measure: Devise a statistic $\Psi : V(\mathcal{T}_n) \to \mathbb{R}_+$ on the unlabelled network that gives a 'score' to each vertex that strongly correlates with age.

Main Strategy

- Centrality measure: Devise a statistic $\Psi : V(\mathcal{T}_n) \to \mathbb{R}_+$ on the unlabelled network that gives a 'score' to each vertex that strongly correlates with age.

- Confidence set: For any $\epsilon \in (0, 1)$, set $S_n(\epsilon)$ to be the set of $k(\epsilon)$ vertices with highest (or lowest) Ψ-scores in $\mathcal{T}_n$.

Main Strategy

- Centrality measure: Devise a statistic $\Psi : V(\mathcal{T}_n) \to \mathbb{R}_+$ on the unlabelled network that gives a 'score' to each vertex that strongly correlates with age.

- Confidence set: For any $\epsilon \in (0, 1)$, set $S_n(\epsilon)$ to be the set of $k(\epsilon)$ vertices with highest (or lowest) Ψ-scores in $\mathcal{T}_n$.

- Persistence: For any $k \in \mathbb{N}$, the k vertices with the highest (or lowest) Ψ-scores eventually fixate as the network grows. Leads to algorithms stable in network size.

Examples of centrality measures

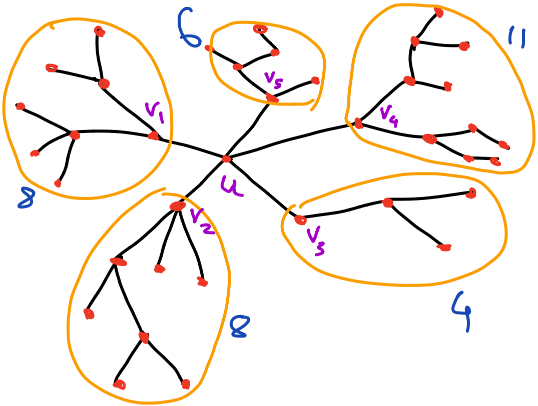
The following centrality measures will play a crucial role in this talk.

- Degree centrality (local): $\Psi_d(u, \mathcal{T}) := \deg(u)$.
- Subtree centrality (global, defined only for tree networks):

$$\Psi_s(u, \mathcal{T}) := \max_{v \sim u} |(\mathcal{T}, u)_{v\downarrow}|$$

where $(\mathcal{T}, u)_{v\downarrow}$ is the subtree rooted at v in the rooted tree (T, u). The centroid is the vertex that minimizes $\Psi_s(\cdot, \mathcal{T})$.

Centrality measures (contd.)



$$\Psi_d(u, \mathcal{T}) = 5$$

$$\Psi_s(u, \mathcal{T}) = |(\mathcal{T}, u)_{v_4\downarrow}| = 11.$$

Generalized Attachment (GA) networks

- At each time step n, a new vertex v_n connects to the existing network through m_n new edges attached one by one.

Generalized Attachment (GA) networks

- At each time step n, a new vertex v_n connects to the existing network through m_n new edges attached one by one.
- Given the current structure of the network, each new edge is attached to one of the existing vertices u with probability proportional to a positive function f of the degree of u.

$$\mathbb{P}\left((k+1)^{th} \text{ edge of } v_{n+1} \to u \mid \mathcal{T}_{n,k}\right) = \frac{f(\deg(u))}{\sum_{v \in \mathcal{T}_{n,k}} f(\deg(v))}.$$

where $\mathcal{T}_{n,k}$ is the network after adding n vertices and k edges of the $(n+1)$-th vertex.

Generalized Attachment (GA) networks

- At each time step n, a new vertex v_n connects to the existing network through m_n new edges attached one by one.

- Given the current structure of the network, each new edge is attached to one of the existing vertices u with probability proportional to a positive function f of the degree of u.

$$\mathbb{P}\left((k+1)^{th} \text{ edge of } v_{n+1} \to u \mid \mathcal{T}_{n,k}\right) = \frac{f(\deg(u))}{\sum_{v \in \mathcal{T}_{n,k}} f(\deg(v))}.$$

  where $\mathcal{T}_{n,k}$ is the network after adding n vertices and k edges of the $(n+1)$-th vertex.

- $\{m_n\}$ called the attachment sequence and f called the attachment function.

Generalized Attachment (GA) networks

- At each time step n, a new vertex v_n connects to the existing network through m_n new edges attached one by one.

- Given the current structure of the network, each new edge is attached to one of the existing vertices u with probability proportional to a positive function f of the degree of u.

$$\mathbb{P}\left((k+1)^{th} \text{ edge of } v_{n+1} \to u \mid \mathcal{T}_{n,k}\right) = \frac{f(\deg(u))}{\sum_{v \in \mathcal{T}_{n,k}} f(\deg(v))}.$$

  where $\mathcal{T}_{n,k}$ is the network after adding n vertices and k edges of the $(n+1)$-th vertex.

- $\{m_n\}$ called the attachment sequence and f called the attachment function.

- Network is a tree if $m_n \equiv 1$.

Linear and Uniform attachment networks

When $f(k) = k + \beta$ we get the celebrated (Affine) Linear Preferential Attachment (LPA) network which exhibits scale-free and small-world properties.

When $f \equiv 1$, we get the Uniform Attachment (UA) network.

These are more tractable due to a deterministic denominator in the attachment probability.

Questions

- How much does degree correlate with age? Can we use it to obtain degree-based root finding algorithms? For what attachment functions are they stable in the network size?

Questions

- How much does degree correlate with age? Can we use it to obtain degree-based root finding algorithms? For what attachment functions are they stable in the network size?

- In the absence of stable degree-based root finding algorithms, how much of the connectivity information do we need to find the root?

Degree centrality: Persistence and root finding algorithms

Advantages

- Computationally more efficient.

- Works on non-tree like networks.

Previous work

- Dereich and Mörters (2009) investigated persistence of the maximal degree vertex for a related directed random graph model. They showed that persistence holds if and only if the associated attachment function $f(k)$ grows sufficiently fast with k (captured through a summability condition).

- Galashin (2013) showed that the maximal degree vertex in a GA network persists when $f(k) = k$ or f is a convex function.

Two regimes: persistent and non-persistent

We show that persistence of the maximal degree vertex is completely characterized by the attachment sequence and the following two quantities:

$$\Phi_1(n) := \sum_{k=1}^{n} \frac{1}{f(k)}, \quad \Phi_2(n) := \sum_{k=1}^{n} \frac{1}{f^2(k)}.$$

Theorem (B. and Bhamidi (2020a))

Assume f *is non-decreasing and there exists* $C_f > 0$ *such that* $f(i) \leqslant C_f i$ *for all* $i \geqslant 1$*. Also, suppose* $\Phi_2(\infty) < \infty$ *and that, almost surely,*

$$\limsup_{n\to\infty} \frac{\Phi_1(m_n)}{\log\left(\sum_{i=1}^{n} m_i\right)} \leqslant \frac{1}{8C_f}.$$

Then the maximal degree vertex persists eventually.
If $\Phi_2(\infty) = \infty$*, the maximal degree vertex does not persist.*

Breaking persistence in the persistent regime

Theorem (B. and Bhamidi (2020a))

Assume that f *is non-decreasing,* $\Phi_2(\infty) < \infty$ *and* $\{m_i\}_{i \geqslant 1}$ *form an i.i.d. sequence supported on* $\mathbb{N}$*.*
(i) Suppose there exists $C_f > 0$ *such that* $f(i) \leqslant C_f i$ *for all* $i \geqslant 1$*. If there exist positive constants* D, z_0 *and* $\theta > 1$ *such that*

$$\mathbb{P}\left(\Phi_1(m_1) > z\right) \leqslant e^{-Dz^{\theta}}, \; z \geqslant z_0,$$

then almost surely maximal degree vertex persists eventually.

(ii) Suppose $\mathbb{E}(m_1) < \infty$*. If there exist positive constants* D', z_0' *and* $\theta' \in (0, 1)$ *such that*

$$\mathbb{P}\left(\Phi_1(m_1) > z\right) \geqslant e^{-D' z^{\theta'}}, \; z \geqslant z_0',$$

then, almost surely, there is no persistence.

GA trees: Scaling of maximum degree in persistent regime

In this case, we do not need monotonicity assumptions on f and can obtain precise asymptotics for the maximum degree and age of the maximal degree vertex.
A key role is played by the Malthusian rate λ^* which is the unique solution to

$$\sum_{k=1}^{\infty} \prod_{i=1}^{k} \frac{f(i)}{\lambda + f(i)} = 1.$$

Theorem (B. and Bhamidi (2020a))

Let $\Phi_2(\infty) < \infty$ and $f(i) \leqslant C_f i$ for all $i \geqslant 1$. Then there is persistence and the maximum degree exhibits the following asymptotics:

$$d_{max}(n) = \Phi_1^{-1}\left(\frac{1}{\lambda^*}\log n + X_n^*\right),$$

where X_n^* converges almost surely to some random variable X^* as $n \to \infty$.

GA trees: Age and degree of maximal degree vertex in non-persistent regime

Theorem (B. and Bhamidi (2020a))

Assume $\Phi_2(\infty) = \infty$. *Under some regularity assumptions on* f *and assuming* $f(k) \to \infty$ *as* $k \to \infty$, *the age* $\mathcal{A}_n$ *of the maximal degree vertex when the network is of size* n *exhibits the following asymptotics:*

$$\frac{\log \mathcal{A}_n}{\Phi_2 \circ \Phi_1^{-1}\left(\frac{1}{\lambda^*}\log n\right)} \xrightarrow{P} \frac{\lambda^{*2}}{2}, \quad \text{as } n \to \infty.$$

Moreover, the maximum degree satisfies

$$\frac{\Phi_1(d_{max}(n)) - \frac{1}{\lambda^*}\log n}{\Phi_2 \circ \Phi_1^{-1}\left(\frac{1}{\lambda^*}\log n\right)} \xrightarrow{P} \frac{\lambda^*}{2}, \quad \text{as } n \to \infty.$$

Example: Sublinear attachment functions

- When $f(k) = k^{\alpha}$, then persistent regime corresponds to $\alpha \in (1/2, 1]$.
- For $\alpha = 1$, $d_{max}(n) \sim \sqrt{n}$ and $\mathcal{A}_n = O(1)$.
- For $\alpha \in (1/2, 1)$, then $d_{max}(n) \sim (\log n)^{1/(1-\alpha)}$ and $\mathcal{A}_n = O(1)$.
- For $\alpha \in (0, 1/2)$, then $d_{max}(n) \sim (\log n)^{1/(1-\alpha)}$ and

$$\mathcal{A}_n \sim \exp\left\{ C(\log n)^{(1-2\alpha)/(1-\alpha)} \right\}.$$

For given error tolerance $\epsilon \in (0,1)$, let $K_d(\epsilon)$ denote the least number of maximal degree vertices required to form the confidence set for the root.

In the persistent regime, $K_d(\epsilon)$ can be chosen independent of network size.

Theorem (B. and Huang (2021))

Suppose $f(i) = i + \beta, i \geqslant 1,$ *for some* $\beta \geqslant 0$, *and* $m_i \equiv m \geqslant 1$. *There exist positive constants* C_1, C_1', C_2, *depending on* m, β, *such that for any* $\epsilon \in (0,1)$,

$$\frac{C_1'}{\epsilon^{\frac{2m+\beta}{m(m+\beta)}}} \leqslant K_d(\epsilon) \leqslant \frac{C_1}{\epsilon^{\frac{2m+\beta}{m(m+\beta)}}} \exp\left(\sqrt{C_2 \log \frac{1}{\epsilon}}\right).$$

Root finding algorithms in the persistent regime (contd.)

Write $f_* := \inf_{k \geqslant 1} f(k)$.

Theorem (B. and Huang (2021))

Suppose the attachment function f *satisfies* $\Phi_2(\infty) < \infty$ *and some regularity assumptions.*
(i) Suppose $m_i \equiv m = 1$. *For any fixed* $\delta \in (0,1)$, *there exist positive constants* C_1 *(not depending on* δ*) and* C_δ *(depending on* δ*) such that for all* $\epsilon \in (0,1)$,

$$\frac{C_1}{\epsilon^{\frac{\lambda^*}{f_*}}} \leqslant K_d(\epsilon) \leqslant \frac{C_\delta}{\epsilon^{\frac{\lambda^*}{(1-\delta)f_*}}}.$$

(ii) Suppose $m_i \equiv m > 1$. *For any* $\delta \in (0,1)$, *there exist positive constants* C_1 *(not depending on* δ*) and* $\bar{C}_\delta$ *(depending on* δ*) such that for all* $\epsilon \in (0,1)$,

$$\frac{C_1}{\epsilon^{\frac{f_*}{mf(m)}}} \leqslant K_d(\epsilon) \leqslant \frac{\bar{C}_\delta}{\epsilon^{\frac{2C_f}{(1-\delta)f_*}}}, \ \epsilon \in (0,1).$$

Root finding in the non-persistent regime (tree case)

Let $f(k) = k^\alpha$ for some $\alpha \in (0, 1/2]$. For $v \in V(\mathcal{T}_n)$, define

$$\Psi_n(v) := \max\{\deg(u) : \text{dist}(u, v) \leqslant C_1 (\log n)^{(1-2\alpha)/(1-\alpha)}\}.$$

Theorem (B. and Huang (2021))

Suppose $m_i \equiv 1$. *Let* S_n *be the set of*

$$\exp\left\{C_2 (\log n)^{(1-2\alpha)/(1-\alpha)} \log\log n\right\}$$

vertices with the highest Ψ_n*-scores. Then, for any* $\epsilon \in (0, 1)$,

$$\mathbb{P}\left(\text{root} \in S_n\right) \geqslant 1 - \epsilon$$

for all $n \geqslant \exp\left\{C_3 [\log(1/\epsilon)]^{(1-\alpha)/(1-2\alpha)}\right\}$.

Stable root finding algorithms in non-persistent regime?

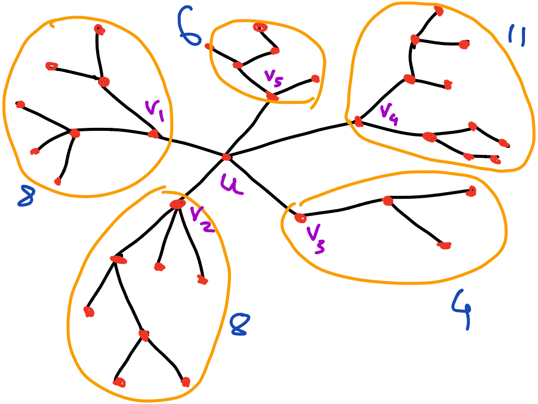
- Can we obtain a persistent centrality measure even in the (degree) non-persistent regime by using the global network geometry to compute the centrality score of any vertex?

Stable root finding algorithms in non-persistent regime?

- Can we obtain a persistent centrality measure even in the (degree) non-persistent regime by using the global network geometry to compute the centrality score of any vertex?
- Associated root finding algorithms will be stable in the network size at the cost of added computational complexity.

Subtree centrality: Persistence and root finding in tree networks

Recall



$$\psi_d(u, T) = 5$$

$$\psi_s(u, T) = |(T, u)_{v_4\downarrow}| = 11.$$

Previous work

- Bubeck, Devroye and Lugosi (2015) used the subtree centrality measure to obtain confidence sets for the root in the LPA and UA case that grow polynomially in ϵ^{-1}.

Previous work

- Bubeck, Devroye and Lugosi (2015) used the subtree centrality measure to obtain confidence sets for the root in the LPA and UA case that grow polynomially in ϵ^{-1}.

- Loh and Jog (2015) showed that the subtree centrality measure persists for LPA and UA trees.

Previous work

- Bubeck, Devroye and Lugosi (2015) used the subtree centrality measure to obtain confidence sets for the root in the LPA and UA case that grow polynomially in ϵ^{-1}.

- Loh and Jog (2015) showed that the subtree centrality measure persists for LPA and UA trees.

- Loh and Jog (2016) considered the case when $f(k) = k^{\alpha}, \alpha \in (0, 1)$ and showed that there is a unique terminal centroid, namely, a vertex v^* such that for any other fixed vertex u, there exists N_u such that

$$\Psi_s(v^*, \mathcal{T}_n) < \Psi_s(u, \mathcal{T}_n), \ \forall\, n \geqslant N_u.$$

  They used this weak persistence to show existence of the confidence set $S(\epsilon)$ although their methods did not give explicit quantitative bounds.

Main results: Persistence

Theorem (B. and Bhamidi (2020b))

Suppose the attachment function f *satisfies* $\inf_{k \geqslant 1} f(i) = f_* > 0$ *and* $\lim_{k \to \infty} \frac{f(k)}{k}$ *exists and is finite. Then the subtree centrality measure is persistent (in the strong sense).*

Confidence set upper bounds

Let $k_s(\epsilon)$ denote the smallest positive integer such that $S(\epsilon)$ comprises the vertices with the least $k_s(\epsilon)$ Ψ_s-scores.

Theorem (B. and Bhamidi (2020a))

Suppose the attachment function f *satisfies the assumptions of the previous theorem.*

- *Suppose for some* $\overline{C}_f > 0$, $\beta \geqslant 0$, f *satisfies* $f_* \leqslant f(i) \leqslant \overline{C}_f \cdot i + \beta$ *for all* $i \geqslant 1$. *Then* $\exists$ *positive constants* C_1, C_2 *such that for any error tolerance* $0 < \varepsilon < 1$,

$$k_s(\epsilon) \leqslant \frac{C_1}{\varepsilon^{(2\overline{C}_f+\beta)/f_*}} \exp(\sqrt{C_2 \log 1/\varepsilon}).$$

- *If further the attachment function* f *is in fact bounded with* $f(i) \leqslant f^*$ *for all* $i \geqslant 1$ *then one has for any error tolerance* $0 < \varepsilon < 1$,

$$k_s(\epsilon) \leqslant \frac{C_1}{\varepsilon^{f^*/f_*}} \exp(\sqrt{C_2 \log 1/\varepsilon}).$$

Theorem (B. and Bhamidi (2020a))

Suppose the attachment function f *satisfies the assumptions of the previous theorems.*

- *If* $\exists\ \underline{C}_f > 0$ *and* $\beta \geqslant 0$ *such that* $f(i) \geqslant \underline{C}_f \cdot i + \beta$ *for all* $i \geqslant 1$ *then* $\exists$ *a positive constant* C_1' *such that for any error tolerance* $0 < \varepsilon < 1$,

$$k_s(\epsilon) \geqslant \frac{C_1'}{\varepsilon^{(2\underline{C}_f + \beta)/f(1)}}.$$

- *For general* f *one has for any error tolerance* $0 < \varepsilon < 1$,

$$k_s(\epsilon) \geqslant \frac{C_1'}{\varepsilon^{f_*/f(1)}}.$$

Special cases

Corollary (B. and Bhamidi (2020a))

- **Uniform attachment** $f(k) \equiv 1$**:**

$$\frac{C_1'}{\varepsilon} \leqslant k_s(\epsilon) \leqslant \frac{C_1}{\varepsilon} \exp\Big(\sqrt{C_2 \log \frac{1}{\varepsilon}}\Big).$$

- **Affine LPA** $f(k) = k + \beta,\ \beta \geqslant 0$**:**

$$\frac{C_1'}{\varepsilon^{\frac{2+\beta}{1+\beta}}} \leqslant k_s(\epsilon) \leqslant \frac{C_1}{\varepsilon^{\frac{2+\beta}{1+\beta}}} \exp\Big(\sqrt{C_2 \log \frac{1}{\varepsilon}}\Big).$$

- **Sublinear preferential attachment** $f(k) = k^\alpha, \alpha \in (0,1)$**:**

$$\frac{C_1'}{\varepsilon} \leqslant k_s(\epsilon) \leqslant \frac{C_1}{\varepsilon^2} \exp\Big(\sqrt{C_2 \log \frac{1}{\varepsilon}}\Big).$$

Proof outline: Key technical ingredients

Point processes

Point processes

- Let $\{E_k : k \geqslant 0\}$ be sequence of independent exponential random variables with E_k having rate $f(k)$. View above as the inter-arrival times of point process ξ_f i.e. writing

$$L_i = E_1 + \cdots + E_i, \qquad i \geqslant 1,$$

$\xi_f[0, t] := \#\{i : L_i \leqslant t\}$. Intensity measure $\mu_f[0, t] := \mathbb{E}(\xi_f[0, t])$.

Key ingredients: Branching process embedding of network (Rudas, Toth and Valko, 2006)

Continuous time branching process (CTBP)

Fix attachment function f. CTBP driven by f, written as $\{\mathrm{BP}_f(t) : t \geqslant 0\}$: branching process started with one individual at time $t = 0$; every individual born into the system has an offspring distribution that is an independent copy of ξ_f.

Embedding

Fix attachment function f consider the sequence of random trees $\{\mathcal{T}_m : 2 \leqslant m \leqslant n\}$ constructed using attachment function f and $m_i \equiv 1$. Define for $m \geqslant 1$ the stopping times $T_m := \inf\{t \geqslant 0 : |\mathrm{BP}_f(t)| = m\}$. Then

$$\{\mathrm{BP}_f(T_m) : 2 \leqslant m \leqslant n\} \stackrel{d}{=} \{\mathcal{T}_m : 2 \leqslant m \leqslant n\}.$$

A continuous time embedding of the non-tree network into a collapsed branching process was obtained in B. and Huang (2021).

Malthusian rate as growth rate of BP

- Recall that the Malthusian rate λ^* is the (unique) solution to $\hat{\rho}(\lambda^*) = 1$, where
$$\hat{\rho}(\lambda) := \sum_{k=1}^{\infty} \prod_{i=1}^{k} \frac{f(i)}{\lambda + f(i)}.$$

Malthusian rate as growth rate of BP

- Recall that the Malthusian rate λ^* is the (unique) solution to $\hat{\rho}(\lambda^*) = 1$, where
$$\hat{\rho}(\lambda) := \sum_{k=1}^{\infty} \prod_{i=1}^{k} \frac{f(i)}{\lambda + f(i)}.$$

- $\hat{\rho}$ arises as the Laplace transform of the intensity measure μ_f of the point process ξ_f:
$$\hat{\rho}(\lambda) = \int_0^\infty e^{-\lambda t} \mu_f(dt).$$
Thus, λ^* is the unique positive λ which makes the measure $\theta_\lambda(dt) = e^{-\lambda t}\mu_f(dt)$ a probability measure.

Malthusian rate as growth rate of BP

- Recall that the Malthusian rate λ^* is the (unique) solution to $\hat{\rho}(\lambda^*) = 1$, where
$$\hat{\rho}(\lambda) := \sum_{k=1}^{\infty} \prod_{i=1}^{k} \frac{f(i)}{\lambda + f(i)}.$$

- $\hat{\rho}$ arises as the Laplace transform of the intensity measure μ_f of the point process ξ_f:
$$\hat{\rho}(\lambda) = \int_0^{\infty} e^{-\lambda t} \mu_f(dt).$$
Thus, λ^* is the unique positive λ which makes the measure $\theta_\lambda(dt) = e^{-\lambda t}\mu_f(dt)$ a probability measure.

- λ^* quantifies the 'rate of exponential growth' for the branching process:
$$e^{-\lambda^* t}|\mathsf{BP}(t)| \overset{a.s.}{\to} W.$$

Tail behavior of W and rate of convergence

W satisfies the following recursive distributional equation (RDE):

$$W \stackrel{d}{=} \sum_{i=1}^{\infty} e^{-\lambda^* L_i} W_i$$

where W_i are i.i.d. copies of W. A crucial technical advancement is the use of the above RDE in showing the following.

Theorem (B. and Bhamidi (2020a))

Suppose the attachment function f *satisfies* $\inf_{k \geqslant 1} f(i) = f_* > 0$ *and* $\lim_{k \to \infty} \frac{f(k)}{k}$ *exists and is finite. Then the distribution of* W *has exponential tails.*

This, along with a quantitative rate of convergence of $e^{-\lambda^* t}|\mathrm{BP}(t)|$ to W obtained in B., Bhamidi and Carmichael (2018), were two crucial technical ingredients in proving our results.

A convenient martingale

- The analysis of degree centrality crucially depends on the continuous time martingale
$$M(t) := \Phi_1(\xi_f(t)) - t,$$
where $\Phi_1(n) = \sum_{k=1}^{n} \frac{1}{f(k)}$ and $\xi_f(\cdot)$ is the point process used in the continuous time embedding.

A convenient martingale

- The analysis of degree centrality crucially depends on the continuous time martingale
$$M(t) := \Phi_1(\xi_f(t)) - t,$$
where $\Phi_1(n) = \sum_{k=1}^{n} \frac{1}{f(k)}$ and $\xi_f(\cdot)$ is the point process used in the continuous time embedding.

- The competition between degrees of two fixed vertices as the network grows can be encoded in terms of competition between two copies of the above martingale started from different points.

A convenient martingale

- The analysis of degree centrality crucially depends on the continuous time martingale
$$M(t) := \Phi_1(\xi_f(t)) - t,$$
where $\Phi_1(n) = \sum_{k=1}^{n} \frac{1}{f(k)}$ and $\xi_f(\cdot)$ is the point process used in the continuous time embedding.

- The competition between degrees of two fixed vertices as the network grows can be encoded in terms of competition between two copies of the above martingale started from different points.

- $\Phi_2(\infty) < \infty$ implies $M(\cdot)$ has finite quadratic variation. Showing persistence relies on concentration inequalities and large deviations for this martingale and the Borel-Cantelli lemma.

A convenient martingale

- The analysis of degree centrality crucially depends on the continuous time martingale
$$M(t) := \Phi_1(\xi_f(t)) - t,$$
where $\Phi_1(n) = \sum_{k=1}^{n} \frac{1}{f(k)}$ and $\xi_f(\cdot)$ is the point process used in the continuous time embedding.

- The competition between degrees of two fixed vertices as the network grows can be encoded in terms of competition between two copies of the above martingale started from different points.

- $\Phi_2(\infty) < \infty$ implies $M(\cdot)$ has finite quadratic variation. Showing persistence relies on concentration inequalities and large deviations for this martingale and the Borel-Cantelli lemma.

- When $\Phi_2(\infty) = \infty$, non-persistence is a consequence of a functional central limit theorem for the martingale.

Asymptotics for age and moderate deviations

Obtaining the age asymptotics for the maximal degree vertex in the non-persistent regime relies on obtaining moderate deviation principles for the martingale $M(\cdot)$ along with the continuous time embedding of the GA tree in a branching process.

Ongoing and future work

- Persistent local centrality scores in the (degree) non-persistent regime and connection to Google's PageRank (ongoing work with Mariana Olvera-Cravioto and Shankar Bhamidi).

Ongoing and future work

- Persistent local centrality scores in the (degree) non-persistent regime and connection to Google's PageRank (ongoing work with Mariana Olvera-Cravioto and Shankar Bhamidi).

- Source detection for epidemics on Galton-Watson trees conditioned to be infinite (ongoing work with Shankar Bhamidi and Sumit Kar).

Ongoing and future work

- Persistent local centrality scores in the (degree) non-persistent regime and connection to Google's PageRank (ongoing work with Mariana Olvera-Cravioto and Shankar Bhamidi).
- Source detection for epidemics on Galton-Watson trees conditioned to be infinite (ongoing work with Shankar Bhamidi and Sumit Kar).
- Exploration type algorithms for root finding (see Borgs et. al (2013) and Frieze and Pegden (2018) for such algorithms in the LPA case).

Thank You!