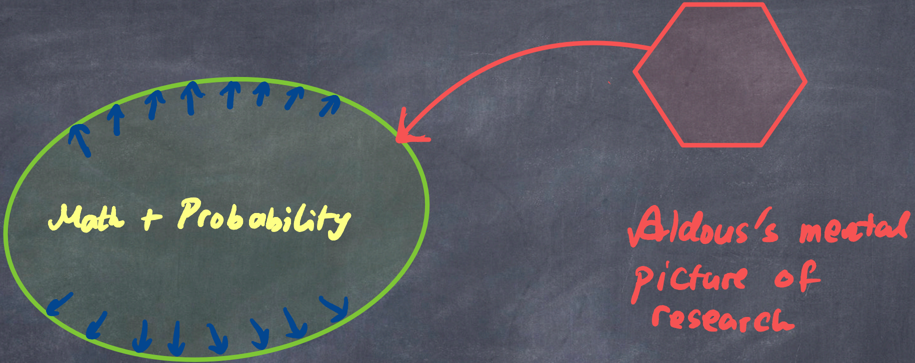
Dynamic network models and stochastic approximation

CNI, Indian Institute of Science, 2024.

Shankar Bhamidi
UNC Chapel Hill
Department of Statistics and OR



Bottom line: Area of dynamic networks needs mathematicians!

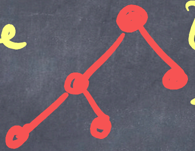
All of this work joint with

- Prof Sayan Banerjee [UNC]
- Prof Partha Dey [UIUC]
- Akshay Sakanaveeti [UNC]
- Zoe Huang [UNC]
- Vladas Pipiras [UNC]
- Nelson Antunes [University of Lisbon]
- Iain Carmichael [UNC]

(One) Math Punchline

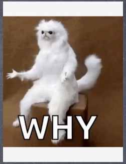
- Consider a sequence of growing network models $\{\mathcal{T}_n : n \geq 1\}$ in discrete time
- Fix your favorite empirical quantity of interest

e.g. $\dfrac{\#\text{ of vertices of degree} = 10}{n}$

− {# of vertices whose distance 2 neighborhood looks like  }

$\overline{}$

n



Turns out: In many many network models Continuous time branching processes naturally describe the limits of such objects.

OUTLINE

[1] Motivation from one area: Attributed network models

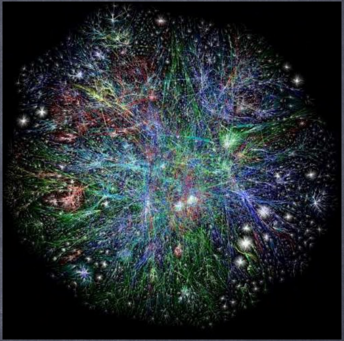
- Fundamental questions and hypothesis
- "News you can use"
- Propogation of chaos $\rightarrowtail$ CTBP $\rightarrowtail$ Math understanding

[2] Change point detection

[3] Network evolution with limited information

SUMMARY FINDINGS

1. Dynamic network models are truly complicated beasts. Simple rules give rise to complex phenomenon, quite often hard to predict even from simulation

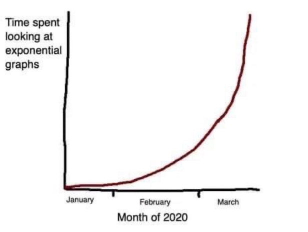


↳ "Internet" by Opte Project

2. Owing either explicitly (construction of model) or implicitly (propogation of chaos) dynamics often driven towards evolution mechanisms in continuous time branching processes.

SUMMARY FINDINGS CONTINUED

$\boxed{3}$ Continuous time branching processes grow exponentially (at some rate λ) while functionals of interest (e.g. degree distribution, Page rank scores) grow at a different rate ($\alpha_{\text{functional}}$). Asymptotics emerges from the interplay of these two rates.

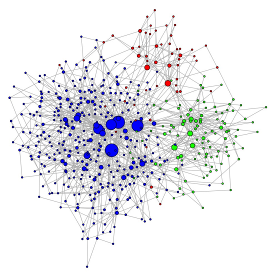


Found on Twitter
e.g. Michael Reuter

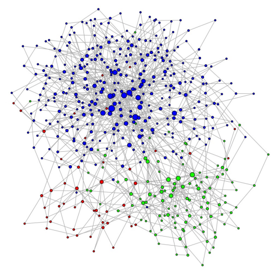
I. ATTRIBUTED NETWORK MODELS

Motivation

- Most social networks consist of vertices with attributes.
- $\mathcal{S}$ = attribute space. For talk $\mathcal{S} = \{1, 2, \ldots, k\}$
- Typically these networks are
    - Dynamic
    - Connections modulated by factors such as
        - heterogeneity of connection propensities across attributes
        - time and path dependent
        - Popularity bias



$\gamma = 1$



$\gamma = 0.2$

- Corresponding social networks play major role in diffusion of information
- Used by Companies via ranking/centrality algorithms to find influential nodes and pay such nodes to direct flow of information, effect perception of specific groups etc
- **Number of FOLK THEOREMS**

<u>Example</u>: Most centrality scores have similiar behavior for such networks

Related important question

- In many settings cannot directly observe network. Need to sample from network
- Perhaps interested in a "rare" minority
- e.g. Asian Immigrant populations in Research Triangle and Impact of COVID etc in early 2020 [GIOVANNA MERLI } Duke + Carolina Population Center]
  TED MOUW

WebRDS from University of Michigan Ann Arbor

Punchline - Has motivated a detailed development of network models that incorporate important functionals in their evolution

- Derive insight about various phenomenon from these models

"Mechanistic network models from domains of complexity science can enable researchers to consider various hypothetical scenarios ... This allows to evaluate robustness of algorithms with regards to different aspects concerning minorities, for example fairness or discrimination."

F. Karimi, M. Oliveira, and M. Strohmaier: arxiv 2206:07113

Main model in town

- Latent space $\mathcal{S} = [K] = \{1, 2, \ldots, K\}$.
- Fix a probability measure $\boldsymbol{\pi}$ on $\mathcal{S}$ (density of different types).
- Potentially asymmetric function $\kappa : \mathcal{S} \times \mathcal{S} \to \mathbb{R}_+$ (propensities of pairs of nodes to connect, based on their attributes).
- Preferential attachment parameter $\gamma \in [0, 1]$.

[Jonathan Jordan 2013
Lot of activity in applied
Community ≈ 2018+]

Model class $\mathscr{P}(\gamma, \boldsymbol{\pi}, \kappa)$

- Vertices enter the system sequentially for $n \geq 1$ starting with a base connected graph $\tilde{\mathcal{G}}_0$. Write v_n for the vertex that enters at time n; every vertex v_n has attribute distribution $a(v_n) \sim \boldsymbol{\pi}$ independent of $\left\{\tilde{\mathcal{G}}_s : 0 \leq s \leq n-1\right\}$.
- For $v \in \tilde{\mathcal{G}}_n$, let $\deg(v, n) =$ degree of v at time n.
- Conditional on $\tilde{\mathcal{G}}_n$ the probability that v_{n+1} connects to $v \in \tilde{\mathcal{G}}_n$ is given by:

$$\mathbb{P}(v_{n+1} \rightsquigarrow v | \tilde{\mathcal{G}}_n, a(v_{n+1}) = a^\star) = \frac{\kappa(a(v), a^\star)[\deg(v, n)]^\gamma}{\sum_{v' \in \tilde{\mathcal{G}}_n} \kappa(a(v'), a^\star)[\deg(v', n)]^\gamma}$$

Will restrict to $\gamma = 1$ in this talk. Will view as directed graphs with edges pointing from children to parent.

Interpretation of Kernel

$\mathcal{K}($  ,  $)$ = Propensity of new vertex (type = CORGI) to connect to existing vertex (type = Shibu)

Functionals of interest

- **Degree distribution of the graph:** Fix $k \geq 1$. $N_n(k) = \#$ of vertices of degree k in $\mathcal{G}_n$. $\mathbf{p}_n = \{N_n(k)/n : k \geq 0\}$ = empirical probability mass function.
- **Joint distribution of attributes and types:** $\varkappa_n(\cdot) = \frac{1}{n}\sum_{v \in \mathcal{V}_n} \delta_{(\deg(v), a(v))}$.
- **Page rank scores for directed graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ with damping factor $c \in (0,1)$=** stationary distribution $(\mathfrak{R}_{v,c} : v \in \mathcal{G})$ of following random walk: at each step, with probability c, follow an outgoing edge (uniform amongst available choices) from current location in the graph. With probability $1-c$, restart at uniformly selected vertex in entire graph. Given by linear system of equations:

$$\mathfrak{R}_{v,c} = \frac{1-c}{n} + c \sum_{u \in \mathcal{N}^-(v)} \frac{\mathfrak{R}_{u,c}}{d^+(u)} \tag{1}$$

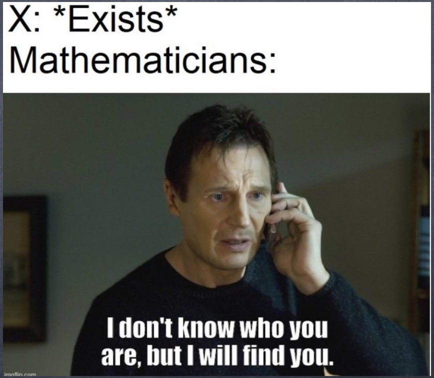
where $\mathcal{N}^-(v)$ is the set of vertices with edges pointed at v and $d^+(u)$ is the out-degree of vertex u. [Can similiarly look at joint distn between attribute + page rank]

Methodological questions: how do centrality measures (degree centrality; page rank scores) vary by attribute type?

Main issue: math tractability for functional of interest

$\mathcal{P}$ = "Please analyze this model"

$\mathcal{U}$ = "Useful (maybe)"



Model inputs

Kernel κ and weight measure $\boldsymbol{\nu}$.

Attributed network model $\{\tilde{\mathcal{G}}_n : n \geq 0\}$

$$\mathbb{P}\left(a(v_{n+1}) = a^\star, v_{n+1} \rightsquigarrow v | \tilde{\mathcal{G}}_n\right) := \frac{\kappa(a(v), a^\star)\nu(a^\star)[\deg(v,n)]^\gamma}{\sum_{a\in[K]}\sum_{v'\in\tilde{\mathcal{G}}_n}\nu(a)\kappa(a(v'), a)[\deg(v',n)]^\gamma}.$$

YUCK!

Seems like a mess: types of new vertices tightly coupled with the evolution of the entire process.

- *$\mathscr{U}$ can be simulated via dynamics where every vertex essentially* **behaves independently**
- Suppose one wanted to simulate model class $\mathscr{U}$ starting from one vertex of type a, then:
- Every vertex v that enters the system (starting with the root of type a) gives birth in continuous time independently to child nodes with attributes, connected to the vertex.
- For a node of type a, conditional on its degree d, the rate of reproduction of a child node of type a' is $\nu(a)\kappa(a, a')d^{\alpha}$.

Write $\{\mathrm{BP}(t) : t \geq 0\}$ for the (continuous time) process. For $n \geq 1$, T_n be the (random) time such that the size $|\mathrm{BP}(T_n)| = n$. Then easy to check that $\{\mathrm{BP}(T_n) : 1 \leq n \leq N\}$ has the same distribution as $\left\{\tilde{\mathcal{G}}_n : 1 \leq n \leq N\right\} \sim \mathscr{U}(\gamma, \nu, \kappa)$.

Math curiosity question

Suppose we can choose $\boldsymbol{\nu}$ such that "asymptotically" composition of population is approximately $\boldsymbol{\pi}$. Are the two model classes $\mathscr{P}$ and $\mathscr{U}$ "similar"?

$$\tilde{\pi}(t) = \frac{1}{|BP(t)|} \sum_{v \in BP(t)} \delta_{a(v)} \to \pi$$

Answer to math curiosity question = YES. Can carry out the entire program, so that asymptotics of all functionals of interest derivable from the "easier to simulate" model class $\mathscr{U}$.

$$
\begin{tikzcd}
\mathscr{P}(1, \boldsymbol{\pi}, \kappa) \arrow[r] \arrow[d, dashed, "??"'] & \boldsymbol{\nu}(1, \boldsymbol{\pi}, \kappa) \arrow[d] \\
\text{Asymptotics} & \mathscr{U}(1, \boldsymbol{\nu}, \kappa) \arrow[l]
\end{tikzcd}
$$

 Goal: model class $\mathcal{P}$

Inputs: π and κ

Let $\mathcal{P}(\mathcal{S})$ denote the space of all probability measures on $\mathcal{S}$. Define (in the interior of $\mathcal{P}(\mathcal{S})$) the function:

$$V_{\boldsymbol{\pi}}(\mathbf{y}) := 1 - \frac{1}{2}\sum_{j\in\mathcal{S}}\pi_j\left(\log(y_j) + \log\Big(\sum_{k\in\mathcal{P}} y_k\kappa_{k,j}\Big)\right)$$

Fundamental Lemma (Jordan (2013), EJP)

Under above Assumptions, $V_{\boldsymbol{\pi}}(\cdot)$ has a *unique* minimizer $\boldsymbol{\eta} := \boldsymbol{\eta}(\boldsymbol{\pi}) = (\eta_1(\boldsymbol{\pi}), \ldots, \eta_K(\boldsymbol{\pi}))$ in the interior of $\mathcal{P}(\mathcal{S})$.

$$\boxed{\nu_b := \frac{\pi_b}{\sum_{l=1}^K \kappa_{l,b}\eta_l},} \qquad \phi_{a,b} := \kappa_{a,b}\nu_b, \qquad \phi_a := \sum_{b=1}^K \phi_{a,b} = 2 - \frac{\pi_a}{\eta_a},$$

Algorithm

- Consider Model class $\mathcal{U}$ with parameters ν and κ
- Easy to simulate as a branching process (in continuous time). Individuals behave independently
- Anything else ??

Main result 1 (all joint work with Nelson Antunes, Sayan Banerjee, Vladas Pipiras)



Theorem (2023) for $\gamma = 1$

Asymptotics for all "local" functionals of model class $\mathscr{P}$ can be obtained from model class $\mathscr{U}$ with above choice of $\boldsymbol{\nu}$. For example, pick a vertex at random in $\mathcal{G}_n \sim \mathscr{P}$ and consider the descendant subtree of that vertex. Then the distribution of this descendant subtree converges to the following:

- Pick $A \sim \boldsymbol{\pi}$.
- Start a branching process simulating model class $\mathscr{U}(1, \boldsymbol{\nu}, \kappa)$ starting from a single vertex of type A.
- Run this simulation for $\tau =$ Exponential random variable with rate $= 2$.

Under the hood: associated branching process $\mathscr{U}$ grows at rate $\lambda = 2$: Simulation takes $\approx \frac{1}{2} \log n$ in the computer to generate network of size n.

Implications: Asymptotics for degree distribution

Branching process grows like e^{2t}. For a vetex of type a, Number of children = degree+1 grows like $e^{\phi_a t}$. Interplay gives the following:

Degree distribution

For each $a \in [K]$, $\mathbf{p}_n^a \to \mathbf{p}_\infty^a$ where the tail pmf is given by

$$\bar{\mathbf{p}}_\infty^a(k) = \frac{\Gamma\left(1 + \frac{2}{\phi_a}\right)\Gamma(k+1)}{\Gamma\left(k+1+\frac{2}{\phi_a}\right)}, \qquad k \geq 0.$$

In particular $\mathbf{p}_\infty^a(k) \sim k^{1+2/\phi_a}$ as $k \to \infty$.

Previous derived in 2013 by Jordan using stochastic approximation techniques. Part of the methodological contribution of our work is to show, stochastic approximation techniques can be used to track evolution of motif counts. → e.g. trees of any shape and attribute structure

Degree distribution tails does depend on the attribute type. Thus potentially, degree centrality scores depend in a non-trivial manner on the type of a vertex.

Implications 2: Page rank scores for model class $\mathscr{P}$

- Recall $\mathcal{G}_n$ is directed with edges from child to parent. For $v \in \mathcal{G}_n$, let $P_l(v, n)$ denote the number of *directed* paths of length l that end at v in $\mathcal{G}_n$. Since $\mathcal{G}_n$ is a directed tree, easy to check PageRank scores have the explicit formulae:

$$\mathfrak{R}_{v,c}(n) = \frac{(1-c)}{n}\left(1 + \sum_{l=1}^{\infty} c^l P_l(v, n)\right).$$

- *Stare at this formula: suggests connection to percolation, where each edge retained with probability c, deleted with probability* $1 - c$*.*
- Easier to formulate results in terms of the *graph normalized* PageRank scores $\{R_{v,c}(n) : v \in \mathcal{G}_n\} = \{n\mathfrak{R}_{v,c}(n) : v \in \mathcal{G}_n\}$.
- Empirical distribution of normalized PageRank scores,

$$\hat{\mu}_{n,\mathrm{PR}} := n^{-1} \sum_{v \in \mathcal{G}_n} \delta\{R_{v,c}(n)\}.$$

Algorithm

- → Go back to model class $\mathcal{U}$
- → Consider percolation on $\mathcal{U}$
- → Turns out: This can again be viewed as a different Branching process. "Easy" to analyze.
- → **Punchline:** Asymptotics about $\wp$ follow from $\mathcal{U}$.

- Consider $\mathrm{BP}_a(\cdot)$, branching process started with one vertex of type a.
- $\mathcal{R}_{\emptyset,c}(t) = (1-c)\left(1+\sum_{l=1}^{\infty} c^l P_{l,\emptyset}(t)\right)$.
- Define "limit" $\mathcal{R}_{\emptyset,c} = \mathcal{R}_{\emptyset,c}(\tau) = (1-c)\left(1+\sum_{l=1}^{\infty} c^l P_{l,\emptyset}(\tau)\right)$.
- As before τ is an exponential rate two random variable.

Weird matrix associated with $\mathscr{U}$

$$\mathbf{M}^{(c)} = \left(\mathbf{M}^{(c)}_{(a,b)} := c\phi_{a,b} + \phi_a \mathbb{1}\{a=b\}\right)_{a,b\in[K]}.$$

λ_c= Perron-Frobenius eigen-value of $\mathbf{M}^{(c)}$.

Fix $a \in [K]$ and damping factor $c \in (0,1)$. For any $t \geq 0$, write $\mathrm{BP}_a^c(t)$ for the connected cluster of the root (which is also a tree) when we retain each edge $e \in \mathrm{BP}_a(t)$ with probability c and delete with probability $(1-c)$, independently across edges. Write $\{\mathrm{BP}_a^c(t) : t \geq 0\}$ for the corresponding non-decreasing rooted tree value process. Let $\mathcal{Z}_a^c(t) = |\mathrm{BP}_a^c(t)|$ for the size of the cluster at time t.

Turns out: $\mathrm{BP}_a^c(\cdot)$ *is also a branching process.* λ_c *is the rate of growth of* $\mathrm{BP}_a^c(\cdot)$ *i.e.* $|\mathrm{BP}_a^c(\cdot)| \approx e^{\lambda_c t}$

For model class $\mathcal{P}$:

Page rank asymptotics

For every continuity point r of the distribution of $\mathcal{R}_{\emptyset,c}$ under $\mathbb{P}_a$

$$n^{-1} \sum_{v \in \mathcal{G}_n} 1\{a(v) = a, R_{v,c}(n) > r\} \xrightarrow{\mathrm{P}} \pi_a \mathbb{P}_a(\mathcal{R}_{\emptyset,c} > r).$$ [→ from $\mathcal{U}$]

Further there exists constants $B_1 < B_2 < \infty$ such that for any attribute:

$$B_1 r^{-2/\lambda_c} \leq \mathbb{P}_a(\mathcal{R}_{\emptyset,c} > r) \leq B_2 r^{-2/\lambda_c}$$

News you can use: Page rank score distributions do **not** *depend on the attribute type. Negates some of the standard assumptions in social networks.*

1. **Uniform node sampling** ($\mathfrak{U}$)**:** Here one picks a vertex uniformly at random from $\mathcal{G}_n$.
2. **Sampling proportional to degree** ($\mathfrak{D}$)**:** Pick a vertex uniformly at random and then pick a neighbor of this vertex uniformly at random.
3. **Sampling proportional to in-degree** ($\mathfrak{I}\mathfrak{D}$)**:** Pick a vertex at random and then select the parent; by convention, if the root is picked (which happens with probability $o_{\mathbb{P}}(1)$ as $n \to \infty$) then select the root.

Network sampling schemes contd

4. **Sampling proportional to Page rank ($\mathfrak{PR}_c$):** Fix a damping factor c and sample a vertex with probability proportional to the page rank scores $\{\mathfrak{R}_{v,c} : v \in \mathcal{G}_n\}$. In the context of the (tree) network model $\{\mathcal{G}_n : n \geq 1\}$ starting with a single root at time zero, by work of Chebolu+Melsted: this can be accomplished by the following "local" algorithm:
   1. Pick a vertex uniformly V at random from $\mathcal{G}_n$.
   2. Independently let $G \sim \mathrm{Geom}(1-c) - 1$ (here $\mathrm{Geom}(\cdot)$ is a Geometric random variable with prescribed parameter with support starting at one).
   3. Starting from V Traverse G steps towards the root (i.e. using the directions of edges in $\mathcal{G}_n$ from child to parent), stopping at the root, if the root is reached before G steps.
4. **Fixed length sampling ($\mathfrak{PR}_M$):** Fix $M \geq 0$. Consider the same implementation of the page rank scheme but here the halting distribution is taken to be $G \equiv M$. Abusing notation, we use $\mathfrak{PR}_M$ to denote this sampling scheme.

Will skim next two slides

Bottom line: Get explicit formulae for bias of various network sampling schemes.

→ All using $\mathcal{U}$

Define matrix

$$\mathbf{M} = \left(\mathbf{M}_{(a,b)} := \frac{\phi_{a,b}}{2-\phi_a}\right)_{a,b\in[K]}.$$

Turns out this has Perron-Frobenius eigen-value =1. Let $\Psi = (\Psi_1, \Psi_2, \ldots, \Psi_K)$ denote the corresponding right eigen-vector, normalized so that $\sum_{a\in[K]} \pi_a \Psi_a = 1$. Consider the Markov chain $\mathbf{S} := \{S_n : n \geq 0\}$ on $[K]$ with transition probability matrix

$$\mathbb{P}_i^{\mathbf{S}}(S_1 = j) := \mathbb{P}^{\mathbf{S}}(S_1 = j | S_0 = i) = \frac{\mathbf{M}_{i,j}\Psi_j}{\Psi_i}, \qquad j \in [K].$$

Write $\mathbb{E}_i^{\mathbf{S}}$ for the expectation operator under $\mathbb{P}_i^{\mathbf{S}}$.

1. Under uniform sampling $\mathbb{P}_{\mathfrak{U}}(a(V_n) = b|\mathcal{G}_n) \xrightarrow{a.s.} \pi_b$.
2. Under sampling proportional to degree $\mathbb{P}_{\mathfrak{D}}(a(V_n) = b|\mathcal{G}_n) \xrightarrow{a.s.} \eta_b$.
3. Under sampling proportional to in-degree,

$$\mathbb{P}_{\mathfrak{I}\mathfrak{D}}(a(V_n) = b|\mathcal{G}_n) \xrightarrow{a.s.} \eta_b\phi_b = \pi_b \frac{\phi_b}{2-\phi_b} = \pi_b\Psi_b\mathbb{E}_b^{\mathbf{S}}\left[\frac{1}{\Psi_{S_1}}\right].$$

4. Under sampling proportional to Page-Rank, letting $G \sim \text{Geom}(1-c) - 1$ independent of $\mathbf{S}$,

$$\mathbb{P}_{\mathfrak{P}\mathfrak{R}_c}(a(V_n) = b|\mathcal{G}_n) \xrightarrow{a.s.} \pi_b\Psi_b\mathbb{E}_b^{\mathbf{S}}\left[\frac{1}{\Psi_{S_G}}\right].$$

   Since $\mathbf{S}$ has stationary distribution $\{\pi_a\Psi_a : a \in [K]\}$,

$$\lim_{c\uparrow 1}\lim_{n\to\infty} \mathbb{P}_{\mathfrak{P}\mathfrak{R}_c}(a(V_n) = b|\mathcal{G}_n) \xrightarrow{a.s.} \pi_b\Psi_b.$$

5. Under fixed length walk sampling,

$$\mathbb{P}_{\mathfrak{P}\mathfrak{R}_M}(a(V_n) = b|\mathcal{G}_n) \xrightarrow{a.s.} \pi_b\Psi_b\mathbb{E}_b^{\mathbf{S}}\left[\frac{1}{\Psi_{S_M}}\right].$$

Why care? Sampling of rare minorities

Consider the specific case of model class $\mathscr{P}$ with two classes 1, 2 with,

$$\boldsymbol{\kappa} = (\kappa(i,j))_{1\le i,j\le 2} = \begin{array}{c} \begin{matrix} 1 & \;\; 2 \end{matrix} \\ \begin{pmatrix} 1 & 1 \\ a & 1 \end{pmatrix} \end{array}, \qquad \boldsymbol{\pi} = \frac{1}{1+\theta}(\theta, 1). \tag{2}$$

We will be interested in the specific case where $\theta \to 0$, more specifically in the setting

$$\theta := \theta(a) = D\sqrt{a},$$

where $D > 0$ is a fixed constant and where $a \downarrow 0$. Thus,

1. Type 1 vertices are relatively *rare* compared to type 2 vertices; we will often refer to type 1 vertices as minorities and type 2 as majorities.
2. Newly entering majority vertices into the population have equal propensity to connect to minority or majority vertices. Minorities have (relatively) **much higher** propensity to connect to other minority vertices, as compared to majority vertices.

Implications

As $a \downarrow 0$:

1. Under uniform node sampling,

$$\mathbb{P}_{\mathfrak{U}}(a(V_n) = 1 | \mathcal{G}_n) \xrightarrow{a.s.} D\sqrt{a} + O(a).$$

2. For sampling proportional to degree,

$$\mathbb{P}_{\mathfrak{D}}(a(V_n) = 1 | \mathcal{G}_n) \xrightarrow{a.s.} 2D\sqrt{a} - (4D^2 + \frac{1}{2})a + O(a^{3/2}).$$

3. For random in-degree based sampling,

$$\mathbb{P}_{\mathfrak{ID}}(a(V_n) = 1 | \mathcal{G}_n) \xrightarrow{a.s.} 3D\sqrt{a} + O(a).$$

4. For Page-rank based sampling (both Geometric and fixed node implementations):

$$\lim_{c \uparrow 1} \lim_{n \to \infty} \mathbb{P}_{\mathfrak{PR}_c}(a(V_n) = 1 | \mathcal{G}_n) = \frac{2D^2 - \frac{1}{2} + \sqrt{\left((2D^2 - \frac{1}{2})^2 + 4D^2\right)}}{2D^2 + \frac{1}{2} + \sqrt{\left((2D^2 - \frac{1}{2})^2 + 4D^2\right)}} + O(\sqrt{a})$$

$$= \lim_{M \uparrow \infty} \lim_{n \to \infty} \mathbb{P}_{\mathfrak{PR}_M}(a(V_n) = 1 | \mathcal{G}_n)$$

Insight gleaned from above analysis

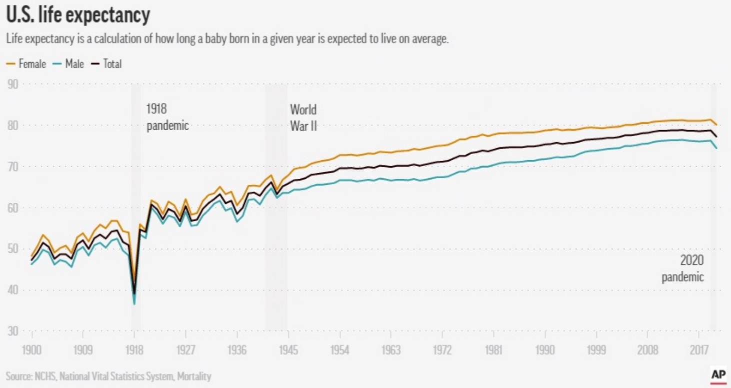
- In the "natural" time-scale of the above models, processes grow exponentially
- "Should imply": Signature of the seed of the network should "persist" for a long time.

→ Should make "estimating" initial seed when one has no temporal information "doable"

→ Should make change point detection "harder"

Long range dependence!

Change Point Detection



Source: Associated Press

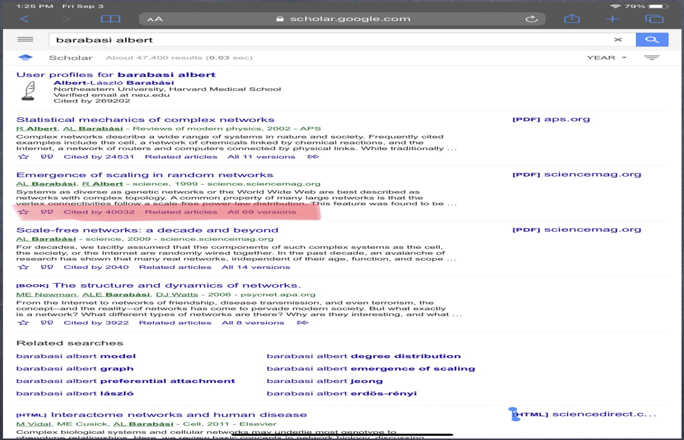
Our motivation in words

- Suppose you have temporal network data.
  - Ex: Adjacency matrix at all or sub-sample of time points
  - Ex: Time series observations at each node etc
- Suppose network experiences a shock at some point.
- Can we detect this change point from observations?
- Changes in structural properties of the system?

Recall: Probabilistic foundations

- Network model: Fix attachment function f. Start with single seed.
- At each stage new vertex v enters system. Connects to one pre-existing vertex
- Probability connecting to a vertex u in the system proportional to $f(\text{degree}(u))$.
- $\mathcal{T}_n$ = network of size n

Example $f(k) = k + \alpha$ Preferential attachment



1:25 PM Fri Sep 3 79%

AA scholar.google.com

barabasi albert

Scholar About 47,400 results (0.03 sec) YEAR

User profiles for **barabasi albert**

Albert-László **Barabási**
Northeastern University, Harvard Medical School
Verified email at neu.edu
Cited by 269202

Statistical mechanics of complex networks [PDF] aps.org
R **Albert**, AL **Barabási** - Reviews of modern physics, 2002 - APS
Complex networks describe a wide range of systems in nature and society. Frequently cited examples include the cell, a network of chemicals linked by chemical reactions, and the Internet, a network of routers and computers connected by physical links. While traditionally …
Cited by 24531 Related articles All 11 versions

Emergence of scaling in random networks [PDF] sciencemag.org
AL **Barabási**, R **Albert** - science, 1999 - science.sciencemag.org
Systems as diverse as genetic networks or the World Wide Web are best described as networks with complex topology. A common property of many large networks is that the vertex connectivities follow a scale-free power-law distribution. This feature was found to be …
Cited by 40032 Related articles All 69 versions

Scale-free networks: a decade and beyond [PDF] sciencemag.org
AL **Barabási** - science, 2009 - science.sciencemag.org
For decades, we tacitly assumed that the components of such complex systems as the cell, the society, or the Internet are randomly wired together. In the past decade, an avalanche of research has shown that many real networks, independent of their age, function, and scope …
Cited by 2040 Related articles All 14 versions

[BOOK] The structure and dynamics of networks.
ME Newman, ALE **Barabási**, DJ Watts - 2006 - psycnet.apa.org
From the Internet to networks of friendship, disease transmission, and even terrorism, the concept--and the reality--of networks has come to pervade modern society. But what exactly is a network? What different types of networks are there? Why are they interesting, and what …
Cited by 3922 Related articles All 8 versions

Related searches

barabasi albert **model**
barabasi albert **graph**
barabasi albert **preferential attachment**
barabasi albert **lászló**
barabasi albert **degree distribution**
barabasi albert **emergence of scaling**
barabasi albert **jeong**
barabasi albert **erdös-rényi**

[HTML] Interactome networks and human disease [HTML] sciencedirect.c…
M Vidal, ME Cusick, AL **Barabási** - Cell, 2011 - Elsevier
Complex biological systems and cellular networks may underlie most genotype to

Known results for $f(k)=k+\alpha$

- $N_k(n)$ = # of vertices of degree k in $\mathcal{T}_n$

$$\frac{N_k(n)}{n} \longmapsto p_k$$

- $p_k \sim \dfrac{C}{k^{\alpha+3}}$ Degree exponent $= \alpha+3$

- max-degree $= M_n \sim n^{\frac{1}{\alpha+2}}$

Example of Standard change point model

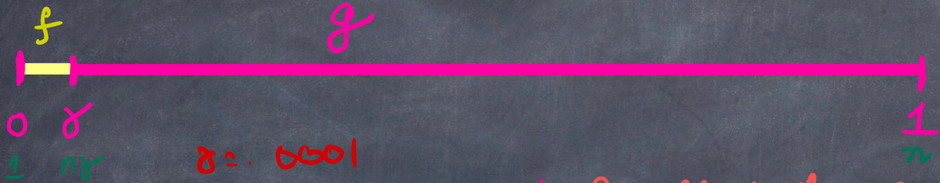
- Fix $\gamma \in (0,1)$.
- For $t \in [1, n\gamma]$, network uses attachment function

$$f(k) = k + \alpha$$

- For $t \in [n\gamma+1, n]$, network uses

$$g(k) = k + \beta$$

Any guesses on the degree exponent?



$\gamma \approx .0001$

Guesses?

Recall under no change

$f(k) = k + \alpha$

degree exponent $= \alpha + 3$

$g(k) = k + \beta$

degree exponent $= \beta + 3$

Punchline of the Theorems

Irrespective of how small γ is (e.g $\gamma = .01$ or $\gamma = .00000001$), the initializer function Always wins!

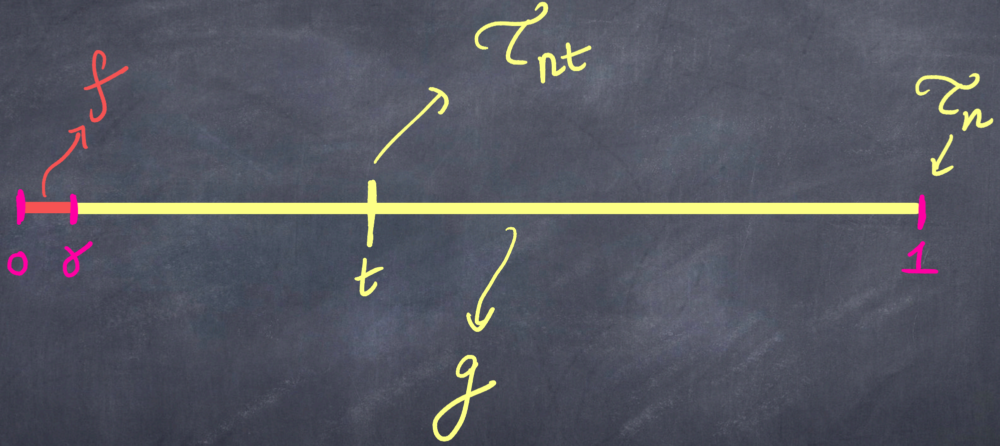
Standard change point model

- Fix $\gamma \in (0,1)$.
- For $t \in [1, n\gamma]$, network uses attachment function

$$f(k) = \text{general function}$$

- For $t \in [n\gamma+1, n]$, network uses

$$g(k) = \text{general function}$$


f
τ_{nt}
τ_n
0
γ
t
1
g

Fix $t \in [0,1]$. Let $N_k(nt)$ = # of vertices of degree k in $\mathcal{T}_{nt}$

<u>Theorem</u> [Banerjee, B, Carmichael]

Under conditions on f and g $\exists$ explicit probability mass functions $\{(p_k(t))_{k\geq 1} : t \in [0,1]\}$ such that

$$\sup_{t \in [0,1]} \left| \frac{N_k(nt)}{nt} - p_k(t) \right| \longrightarrow 0$$

__Theorem__ [Banerjee, B, Carmichael]

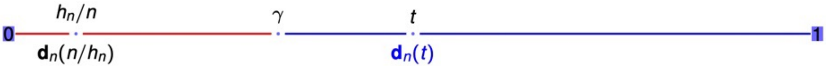
Under above technical conditions on f & g, irrespective of how small γ is f always wins!

- So if degree exponent with f and no change point is δ so is the model __with__ change point.

Change point estimator: For each $t \in (0,1)$ compare degree distn $\left(\frac{N_k(nt)}{nt}\right)_{k \geq 1}$ with the degree distribution when network is of size $\frac{n}{\ln n}$ (recall change point at $n\gamma$)

and become alarmed the first time there seems to be a big change in degree distn.

Change point estimator



Nonparametric change point estimator

Fix any two sequences $h_n \to \infty, b_n \to \infty$: $\frac{\log h_n}{\log n} \to 0$, $\frac{\log b_n}{\log n} \to 0$. Define

$$\hat{\mathcal{T}}_n = \inf \left\{ t \geq \frac{1}{h_n} : \sum_{k=0}^{\infty} 2^{-k} \left| \frac{D_n(k, \mathcal{T}^{\boldsymbol{\theta}}_{\lfloor nt \rfloor})}{nt} - \frac{D_n(k, \mathcal{T}^{\boldsymbol{\theta}}_{\lfloor n/h_n \rfloor})}{n/h_n} \right| > \frac{1}{b_n} \right\}.$$

Then $\hat{\mathcal{T}}_n \xrightarrow{\mathrm{P}} \gamma$.

Lots of open problems

Simulations

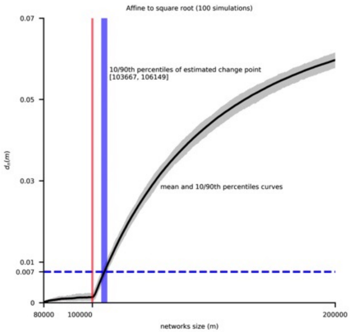


Figure: $n = 2 * 10^5$, $\gamma = 0.5$, $f_0(i) = i + 2$, $f_1(i) = \sqrt{i+2}$, $h_n = \log\log n$, $b_n = n^{1/\log\log n}$

$$d_n(m) := \sum_{k=0}^{\infty} 2^{-k} \left| \frac{D_n(k, \mathcal{T}_m^\theta)}{m} - \frac{D_n(k, \mathcal{T}_{\lfloor n/h_n \rfloor}^\theta)}{n/h_n} \right|, \qquad \frac{n}{\log\log n} < m \leq n.$$

The big bang model: What if the change happened very early in the system?

Figure: Big Bang: Getty images

Fix functions $f_0, f_1 : \{0, 1, 2, \ldots\} \to \mathbb{R}_+$ and $\gamma \in (0, 1)$. Let $\boldsymbol{\theta} = (f_0, f_1, \gamma)$.

Model

- **Time** $1 \leq m \leq n^\gamma$ Vertices perform attachment with probability proportional to $f_0(out - deg)$.
- **Time** $n^\gamma < m \leq n$ Vertices perform attachment with probability probability proportional to $f_1(out - deg)$.

Result 1

- Here change point at n^γ (e.g. $\sqrt{n}$).
- Here
$$\frac{N_n(k)}{n} \xrightarrow{\mathbb{P}} p_k^1$$
namely the degree distribution of the model run purely with attachment function f_1

So what changes?

1. **Uniform $\rightsquigarrow$ Linear:** $f_0 \equiv 1$ whilst $f_1(k) = k + 1 + \alpha$ for fixed $\alpha > 0$. Then for $\omega_n \uparrow \infty$,
$$\frac{n^{\frac{1-\gamma}{2+\alpha}} \log n}{\omega_n} \ll M_n(1) \ll n^{\frac{1-\gamma}{2+\alpha}} (\log n)^2.$$
2. **Linear $\rightsquigarrow$ Uniform:** $f_0(k) = k + 1 + \alpha$ whilst $f_1(\cdot) \equiv 1$.
$$\frac{n^{\frac{\gamma}{2+\alpha}} \log n}{\omega_n} \ll M_n(1) \ll n^{\frac{\gamma}{2+\alpha}} (\log n)^2.$$
3. **Linear $\rightsquigarrow$ Linear:** $f_0(k) = k + 1 + \alpha$ whilst $f_1(k) = k + 1 + \beta$ where $\alpha \neq \beta$. Then $M_n(1)/n^{\eta(\alpha,\beta)}$ is tight where
$$\eta(\alpha, \beta) := \frac{\gamma(2+\beta) + (1-\gamma)(2+\alpha)}{(2+\alpha)(2+\beta)}. \quad (5)$$

Network delay and evolution

- Many probabilistic models of network evolution assume individuals entering the system have complete knowledge of the network
- Next frontier: Network evolution with limited information

Two major recent mechanisms

- Network evolution where individuals form connection via local exploration schemes
- Network delay where individuals have limited information when making descisions on where to connect

Network Co-evolution: Motivation

- Most real world networks support some particular purpose (e.g. diffusion of information on Twitter)
- Co-evolution: Network influences individuals and vice-versa

Motivation 2: More sophisticated models for PA

Motivations

Despite PA being heavily used, number of limitations

I. Assumes global knowledge of network. Each new vertex needs complete knowledge of network

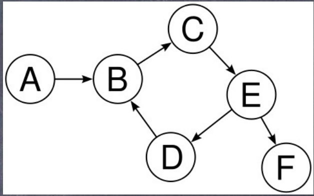
II. In principle attractiveness should not depend ONLY on degree but potentially on "attenuated" neighborhood features.

Example: Page rank score attachment scheme.

Defn [Page rank Scores] Fix "damping factor" c.

For directed graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, page rank score $(\pi_v : v \in \mathcal{V})$ is the stationary distⁿ of a random walk that at each step

- with prob c does usual random walk using outgoing edges
- with prob $1-c$ jumps to a randomly selected vertex uniformly at random

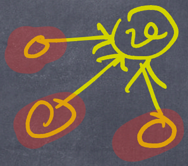


Computersciencewiki Mr. Mckenty

Thus (π_v) satisfies linear system of equations

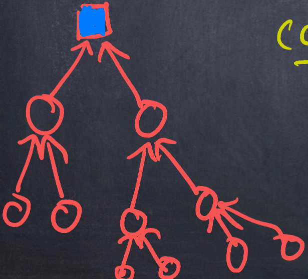
out-degree of u

$$\pi_v = \frac{1-c}{n} + c \sum_{u \in N^-(v)} \pi_u / d^+(u)$$



$N^-(v)$

Special case Directed tree, directions to the root

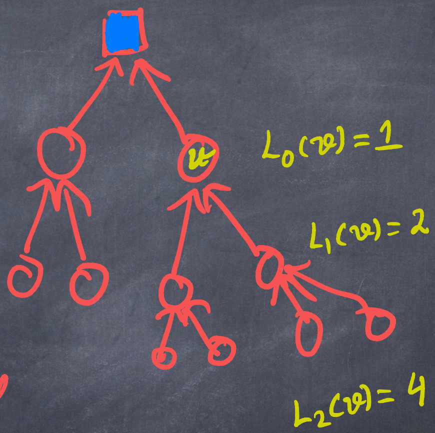


can check For $v \neq$ root

$$\pi_v = \frac{1-c}{n} \sum_{k=0}^{\infty} c^k L_k(v)$$

$L_k(v)$ = # of individuals at level k below v

Thus if one does Preferential attachment using Page rank scores, then one does attachment using more global attractiveness function



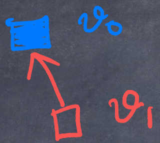
III Motivation which might be contradictory to the previous motivations: Local exploration based attachment schemes

- Might want network evolution schemes where vertices decide to attach to a previous vertex after exploring the network

"web - surfing"

for sometime.

Co-evolutionary network model ($\mathbf{P}$)

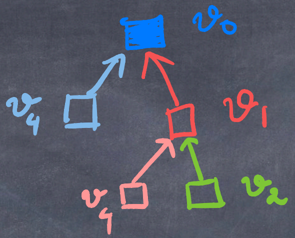
$T_1 =$ 

1. Having Constructed $\mathcal{T}_n$
2. At time $n+1$ a new vertex v_{n+1} enters system
3. Selects vertex V_n u.a.r in $\mathcal{T}_n$
4. Selects # of "exploration steps to root" variable

   $$Z_{n+1} \sim \mathbf{P}$$
5. Goes up that many steps and attaches, stopping at root if need be.

$$\mathbf{P} = \mathrm{pmf} = \{p_0, p_1, \ldots\} \qquad P(Z=i) = p_i \quad i \geq 0$$

Example



$C(2) = v_1, \quad Z_2 = 0$

$C(3) = v_2, \quad z_3 = 4$

$C(4) = v_2, \quad z_4 = 1$

$C(5) = v_0, \quad z_5 = 2$

$C(6) = v_2, \quad z_1 = 1$

$\vdots$

Special cases

① $p_0 \equiv 1$ $\rightarrow$ Random recursive tree (Uniform Attachment)

② $p_0 = p$, $p_1 = 1-p$ $\rightarrow$ Preferential attachment

$$f(k) = k + \frac{(1-2p)}{p}$$

③ $p_0 = p$, $p_1 = p(1-p)$, $p_2 = p(1-p)^2$, ...

"Page rank model"

Theorem [Chebolu + Melsted 200x]

③ $\equiv$ Page rank attachment scheme with $1-c = p$

Theorems [Chebolut Melsted] Phase transition !

- If $p \le \frac{1}{2}$

  $E(\text{degree of root}) = \tilde{\Theta}^*(n)$

- If $p > \frac{1}{2}$

  $E(\text{degree of root}) = \tilde{\Theta}(n^{4pq})$

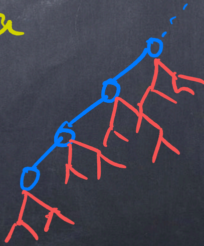
* $\tilde{\Theta} \Leftrightarrow O(n \log^k(n))$ $k \in \mathbb{Z}$

→ LOCAL WEAK CONVERGENCE

Theorem(s) [Banerjee, SB, Huang]

Let $Z \sim \mathbb{P}$

$\boxed{1}$ Assume $E(Z) < \infty$. Then the sequence of trees $\{\tau_n\}_{n \geq 1}$ converge in the local weak convergene sense to a limiting infinite sin-tree



sin: tree with single infinite path to ∞

⇒ for example for every fixed $k \geq 0$

$N_k(n)$ = # of vertices with k children

then $\dfrac{N_k(n)}{n} \overset{a.s}{\longmapsto} p_k \to \{p_k\}_{k \geq 0}$ = PMF

→ If $E(Z) \leq 1$ $\quad \sum_{k=0}^{\infty} k p_k = 1$

→ If $E(Z) > 1$ $\quad \sum_{k=0}^{\infty} k p_k < 1$ ⇝

→ Intuition for mass escaping to ∞. CONDENSATION

$\boxed{2}$ Let $\{Z_i\}_{i\geq 1} \overset{IID}{=} P$ $\qquad f(s) = \text{pgf} = \sum_{k=0}^{\infty} p_k s^k$

Let $S_n = S_0 + \sum_{k=1}^{n} (Z_i - 1)$ = random walk started at S_0

Assumptions

(1) $p_0 \in (0,1)$, $p_0 + p_1 < 1$ [else if $p_0 + p_1 = 1 \Leftrightarrow$ preferential attachment regime]

(2) $f(s)$ is analytic at $s = 1$

Assumptions $\Rightarrow$ by work of [Daley 69] with a few more technical assumptions*

$\Rightarrow$ if we let $T_0 = \inf\{n \geq 1 : S_n = 0\}$ then

- If $E(Z) \neq 1$ then

$$P_1(n < T_0 < \infty) \approx e^{-n \log R}$$

i.e. $S_0 = 1$

for $R = R(P) > 1$

- If $E(Z) = 1$ then $R = 1$

* Aperiodicity + analyticity of pgf at $s=1$.

Results [non-root degree]

Fix $k > 0$ and consider $D_{v_k}(n)$ = degree of v_k at time n

then $\forall \; \delta > 0$

$$\frac{D_{v_k}(n)}{n^{\frac{1}{k}-\delta}} \to \infty \qquad \qquad \frac{D_{v_k}(n)}{n^{1/k}(\log n)^{1+\delta}} \to 0$$

Results

If $E(Z) \leq 1$

Let D = random variable with limiting degree distrn

$$\lim_{k \to \infty} \frac{\log P(D \geq k)}{\log k} = -R$$

Intuitively $P(D \geq k) \simeq \dfrac{C}{k^{R}}$

↳ for $E(Z) > 1$ get upper + lower bounds for degree exponent

CONDENSATION

- Assume $E(Z) > 1$
- Few more technical conditions

Then $\dfrac{D_{v_0}(n)}{n} \xrightarrow{a.s.}$ limit random variable > 0

root

E.g. Random surfer model $p < \frac{1}{2}$ above is true!

Height scales like log(system size)

$$\mathcal{K}_0 = \inf_{s \in (0,1)} \frac{f(s)}{s \log(1/s)}$$

$$\frac{H_n}{\log n} \xrightarrow{P} \mathcal{K}_0$$

Connection between Random Walks + Trees

- What is the first step in studying such models?
- [Chebolu + Melsted idea for Page rank driven model]
- Fix a vertex $u \neq$ root
- Fix a vertex $t+1$ at time $\geq u$
- What is $P($ $t+1$ attaches to u | info till time $t)$?

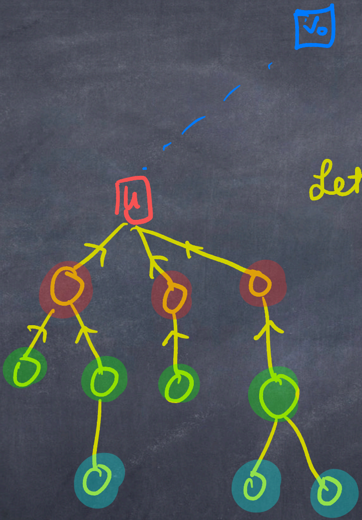
$L_0(t;u) = 1$

$L_1(t;u) = 3$

$L_2(t;u) = 4$

$L_3(t;u) = 3$

$\vdots$



let $L_i(t;u) = \#$ of vertices at distance i below vertex u at time t

$$P(t+1 \rightsquigarrow u \mid \text{information till time } t)$$

$$= \frac{L_0(t;u)}{t} p_0 + \frac{L_1(t;u)}{t} p_1 + \frac{L_2(t;u)}{t} p_2 + \cdots$$

<u>Check</u>: This leads to evolution equation for $\{L_k(t;u) : k \geq 0\}$ as

$L_k(t) = L_k(t; u)$

$$P\left(L_k(t+1) = L_k(t) + 1 \mid \underset{\sim}{L}(t)\right)$$

$$= p_0 \frac{L_{k-1}(t)}{t} + p_1 \frac{L_k(t)}{t} + \cdots$$

$$= \frac{\left[A \cdot \underset{\sim}{L}(t)\right]_k}{t}$$ where

mass escaping above u.

$$A = \begin{pmatrix} 0 & 0 & 0 & \cdots \\ p_0 & p_1 & p_2 & \cdots \\ 0 & p_0 & p_1 & \cdots \\ 0 & 0 & p_0 & \cdots \\ 0 & \cdots & & \end{pmatrix}.$$

→ Easier to do things in continuous time. Continuous time version of what is happening below a vertex $u \neq$ root.

Let $\mathbb{T}$ denote the space of rooted, directed, labelled trees. Let $\mathcal{T}^*(\cdot)$ be the continuous time process of growing trees started with $\mathcal{T}^*(0) = \{v_0\}$,where v_0 is the root of the tree. The vertices in $\mathcal{T}^*(\cdot)$ are labelled $v_0, v_1, v_2, \ldots$ in order of appearance. $\mathcal{T}^*(\cdot)$ is generated by the following procedure:

"vertex u"

Each vertex reproduces at rate 1. When vertex v reproduces, a random variable Z following the law F is sampled independently.

- If $Z \leq dist(v_0, v)$, then a new vertex $\tilde{v}$ is attached to the unique vertex u lying on the path between v and v_0 that satisfies $dist(v, u) = Z$ via a directed edge from $\tilde{v}$ to u.
- If $Z > dist(v_0, v)$, nothing occurs.

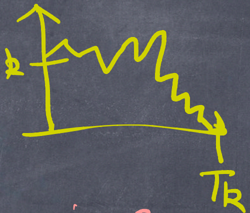
thus

probability of new vertex being born to a current vertex = uniform distr^n

Two at first disjoint objects

Random walk: $S_n = S_0 + \sum_{i=1}^{n} (Z_i - 1)$

$$T_k = \inf\{ n \geq 0 : S_n = 0 \mid S_0 = k \}$$



Branching process: Let $L_k(t)$ = # of vertices in generation k in the tree process described on previous page

Lemma:

$$E(L_k(t)) = \sum_{i=0}^{\infty} \frac{t^i}{i!} P(T_k = i)$$

Mechanism II : Network delay

1.1. **Network evolution with delay.** The model has three ingredients:

a) A *time scale parameter* $\beta \in [0, 1]$. For reasons that will become evident below, the regime $\beta = 1$ will be referred to as the **macroscopic** delay regime

b) A *delay distribution* μ on $(\mathbb{R}_+, \mathcal{B}(\mathbb{R}_+))$. In the $\beta = 1$ regime, we can assume that μ is supported on $[0, 1]$ without loss of generality.

c) An *attachment function* $f : \{0, 1, \ldots\} \to \mathbb{R}_+$, measuring the attractiveness of individuals based on their local information, assumed to be strictly positive.

↪ Motivation from areas like blockchain. Ask Partha Dey!

Definition 1.1 (Network evolution with delay). *We grow a sequence of random trees $\{\mathcal{T}(n) : n \geqslant 1\}$ as follows:*

i) *The initial tree $\mathcal{T}(1)$ consists of a single vertex v_1. At time 2, the tree $\mathcal{T}(2)$ consists of two vertices, labeled as $\{v_1, v_2\}$ and attached by a single edge directed from v_2 to v_1. Call the vertex v_1 as the "root" of the tree.*

ii) *Suppose the network has been constructed till time n for $n \geqslant 2$. Let the vertices in $\mathcal{T}(n)$ be labeled as $\{v_1, \dots, v_n\}$ with edges directed from children to parent. For vertex $v_i \in \mathcal{T}(n)$ (which entered the system at time i) and $j \geqslant i$, let $\deg(v_i, j)$ denote the degree of v_i at time j (which for all vertices other than the root is equal to the in-degree $+1$) interpreted in various applications as a measure of trust accumulated till time i. Initialize always with $\deg(v_i, i) = 1$.*

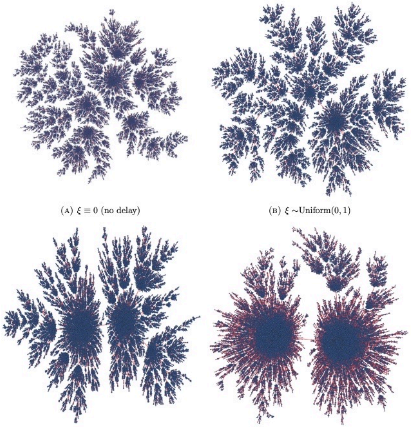
iii) *At time $n+1$, a new vertex v_{n+1} enters the system. A* normalized *time delay $\xi_{n+1} \sim \mu$ independent of $\mathcal{T}(n)$ is sampled. The information available to v_{n+1} is the graph $\mathcal{T}(\lfloor n - n^{\beta}\xi_{n+1} \rfloor)$*

iv) *Conditional on $\mathcal{T}(\lfloor n - n^{\beta}\xi_{n+1} \rfloor)$, this new vertex attaches to a vertex $v \in \mathcal{T}(\lfloor n - n^{\beta}\xi_{n+1} \rfloor)$ with probability proportional to*

$$f(\deg(v, \lfloor n - n^{\beta}\xi_{n+1} \rfloor)).$$

Write $\{\mathcal{T}(n) : n \geqslant 1\}$ for the corresponding sequence of growing random trees and let $\mathcal{L}(\beta, \mu, f)$ denote the corresponding probability distribution of the sequence of growing random trees.

Mesoscopic delay: Simulations



(A) $\xi \equiv 0$ (no delay) (B) $\xi \sim$Uniform$(0,1)$

(C) $\xi \sim 1/$Uniform$(0,1)$ (D) $\xi \sim 1/[\text{Uniform}(0,1)]^2$

FIGURE 1. Simulated Linear Preferential Attachment trees in the mesoscopic regime with $\beta = 1/2$, attachment function taken to be the pure linear preferential attachment function $f(k) = k$ with $\theta = 1/2$ and different delay distributions arranged with increasing gradation of mass towards the right tail. Each simulation is on a network size of $n = 50000$ nodes with different delay distributions. All four settings satisfy the assumptions of Theorem 2.7 (i.e., (6)) and thus have the same local limit as the no delay regime in a large network limit. Delay regime (D) does not satisfy the conditions of Theorem 2.12(i), and thus, for example, the root degree has different scaling than the setting without delay.

Info at time

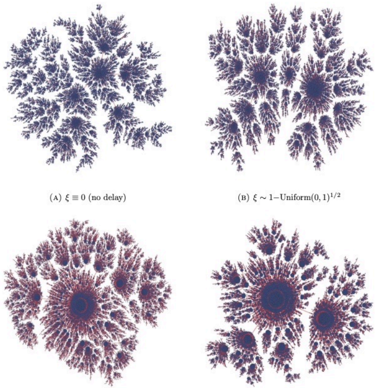
$n =$

$\tau_{[n - \sqrt{n}\,\xi_{n+1}]}$

Theorem [Sayan Banerjee, SB, Partha Dey, Akshay Sakanavelli]

In the mesoscopic regime, under general technical assumptions on the attachment function and minor conditions on the delay distribution (e.g $E[\log_+(\xi)] < \infty$) the local structure asymptotics IS THE SAME as without delay

MESOSCOPIC REGIME : $\beta = 1$



(A) $\xi \equiv 0$ (no delay)

(B) $\xi \sim 1-\text{Uniform}(0,1)^{1/2}$

(C) $\xi \sim 1-\text{Uniform}(0,1)$

(D) $\xi \sim 1-\text{Uniform}(0,1)^2$

FIGURE 1. Simulated Linear Preferential Attachment trees on 50000 nodes with different macroscopic delay distributions.

Info at time n is

$$\tau_{(n - n\xi_{n+1})}$$

$$= \tau_{(n(1-\xi_{n+1}))}$$

$f(k) = k+\alpha$

ξ = Delay r.v

Delay: $n(1-\xi)$

Define random variable η by,

$$\eta := -\log(1-\xi) \text{ with } \mathbb{P}(\eta \in dx) := dF_\eta(x), \text{ for } x \in [0,\infty]. \tag{8}$$

Note that $\xi = 1$ corresponds to the event $\eta = \infty$. Let $\sigma_i = \sum_{j=1}^{i} \mathcal{E}_{j-1\rightsquigarrow j}$, with $\sigma_0 \equiv 0$. The construction proceeds as follows:

a) **Base case $\mathcal{E}_{0\rightsquigarrow 1}$:** Define the hazard function $h_{0\rightsquigarrow 1}(\cdot)$ via,

$$h_{0\rightsquigarrow 1}(x) := \frac{1+\alpha}{2+\alpha}\int_0^x e^u dF_\eta(u), \qquad x \in [0,\infty]. \tag{9}$$

Let $\mathcal{E}_{0\rightsquigarrow 1}$ be the random variable on $[0,\infty]$ with the above hazard rate so that for any x, $\mathbb{P}(\mathcal{E}_{0\rightsquigarrow 1} > s) = \exp(-\int_0^s h_{0\rightsquigarrow 1}(x)dx)$.

b) **General case (general i):** Having constructed $\{\mathcal{E}_{j-1\rightsquigarrow j} : 1 \leqslant j \leqslant i\}$, conditional on the above sequence, consider the hazard function for $x > 0$,

$$h_{i\rightsquigarrow i+1}(x) := \frac{1}{2+\alpha}\left[\sum_{j=1}^{i}\int_{x+\sigma_i-\sigma_j}^{x+\sigma_i-\sigma_{j-1}}(j+\alpha)e^u dF_\eta(u) + \int_0^x (i+1+\alpha)e^u dF_\eta(u)\right]. \tag{10}$$

Let $\mathcal{E}_{i\rightsquigarrow i+1} > 0$ a.s. be the random variable with the above hazard rate.

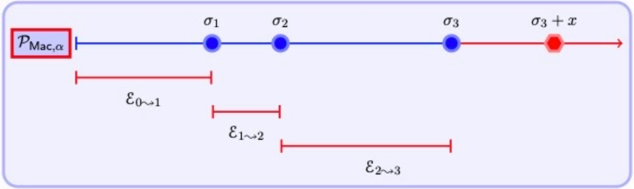


FIGURE 5. Showing the construction of $\mathcal{E}_{3\rightsquigarrow 4}$ conditional on the preceding inter-arrival times and the corresponding point process $\mathcal{P}_{\mathsf{Mac},\alpha}$ in the Macroscopic regime.

Attachment function $f(k) = k+\alpha$

$\{\xi_{i\to i+1} : i \geq 0\}$

Theorem [Sayan Banerjee, SB, Partha Dey, Akshay Sakanaveeti]

In the macroscopic regime, under techinical conditions on the delay (finite MGF of η-transform) Network converges in the local weak limit sense to an infinite random tree whose fringe distribution

= BP with offspring distr^n on previous page run for exp(1) amount of time

- Results on the tail exponent of the limit degree
- In many settings, network has heavier tail than without delay
- Conditions for condensation of the network

Thank you for your attention!

ANY QUESTIONS ?