Preferential attachment graphs and branching random walks

Peter Mörters

University of Bath

joint work with Steffen Dereich (Münster)

and Maren Eckhoff (Bath)

Preferential attachment graphs

Preferential attachment was made popular by Barabási and Albert (1999). It is a principle that aims to explain the emergent features of large complex networks.

Preferential attachment graphs

Preferential attachment was made popular by Barabási and Albert (1999). It is a principle that aims to explain the emergent features of large complex networks.

The building principle of preferential attachment graphs is that

Preferential attachment graphs

Preferential attachment was made popular by Barabási and Albert (1999). It is a principle that aims to explain the emergent features of large complex networks.

The building principle of preferential attachment graphs is that

- An increasing family of graphs is constructed by adding vertices one-by-one.

Preferential attachment graphs

Preferential attachment was made popular by Barabási and Albert (1999). It is a principle that aims to explain the emergent features of large complex networks.

The building principle of preferential attachment graphs is that

- An increasing family of graphs is constructed by adding vertices one-by-one.
- When a new vertex is introduced, it is linked by edges to a fixed or random number of existing vertices with a probability proportional to an increasing function f of their degree.

Preferential attachment was made popular by Barabási and Albert (1999). It is a principle that aims to explain the emergent features of large complex networks.

The building principle of preferential attachment graphs is that

- An increasing family of graphs is constructed by adding vertices one-by-one.
- When a new vertex is introduced, it is linked by edges to a fixed or random number of existing vertices with a probability proportional to an increasing function f of their degree.

Preferential attachment graphs

Preferential attachment was made popular by Barabási and Albert (1999). It is a principle that aims to explain the emergent features of large complex networks.

The building principle of preferential attachment graphs is that

- An increasing family of graphs is constructed by adding vertices one-by-one.
- When a new vertex is introduced, it is linked by edges to a fixed or random number of existing vertices with a probability proportional to an increasing function f of their degree.
- In short: the higher the degree of a vertex, the more likely it is to attract further links.

Preferential attachment graphs

Preferential attachment was made popular by Barabási and Albert (1999). It is a principle that aims to explain the emergent features of large complex networks.

The building principle of preferential attachment graphs is that

- An increasing family of graphs is constructed by adding vertices one-by-one.
- When a new vertex is introduced, it is linked by edges to a fixed or random number of existing vertices with a probability proportional to an increasing function f of their degree.
- In short: the higher the degree of a vertex, the more likely it is to attract further links.

We have analysed a mathematically convenient variant of such graphs. Crucially, the essential features of preferential attachment graphs should not depend on the choice of variant but only on the strength of the preferential attachment and the edge density.

Definition of the model

Take parameters $0 \le \gamma < 1$ and $0 < \beta < 1$ and choose the affine function

$$f : \mathbb{N} \cup \{0\} \to (0, \infty) \text{ given by } f(k) = \gamma k + \beta.$$

Definition of the model

Take parameters $0 \leq \gamma < 1$ and $0 < \beta < 1$ and choose the affine function

$$f : \mathbb{N} \cup \{0\} \to (0, \infty) \text{ given by } f(k) = \gamma k + \beta.$$

Model evolution: At time $n = 1$, we have a single vertex (labeled 1). In each time step $n \to n+1$ we

Definition of the model

Take parameters $0 \leq \gamma < 1$ and $0 < \beta < 1$ and choose the affine function

$$f : \mathbb{N} \cup \{0\} \to (0, \infty) \text{ given by } f(k) = \gamma k + \beta.$$

Model evolution: At time $n = 1$, we have a single vertex (labeled 1). In each time step $n \to n + 1$ we

- add a new vertex labeled $n + 1$, and

Definition of the model

Take parameters $0 \le \gamma < 1$ and $0 < \beta < 1$ and choose the affine function

$$f : \mathbb{N} \cup \{0\} \to (0, \infty) \text{ given by } f(k) = \gamma k + \beta.$$

Model evolution: At time $n = 1$, we have a single vertex (labeled 1). In each time step $n \to n+1$ we

- add a new vertex labeled $n+1$, and
- for each $m \le n$ independently introduce an oriented edge from the new vertex $n+1$ to the old vertex m with probability

$$\frac{f(\text{indegree of } m \text{ at time } n)}{n}.$$

Definition of the model

Take parameters $0 \leq \gamma < 1$ and $0 < \beta < 1$ and choose the affine function

$$f: \mathbb{N} \cup \{0\} \to (0, \infty) \text{ given by } f(k) = \gamma k + \beta.$$

Model evolution: At time $n = 1$, we have a single vertex (labeled 1). In each time step $n \to n+1$ we

- add a new vertex labeled $n+1$, and
- for each $m \leq n$ independently introduce an oriented edge from the new vertex $n+1$ to the old vertex m with probability

$$\frac{f(\text{indegree of } m \text{ at time } n)}{n}.$$

Example:

Definition of the model

Take parameters $0 \leq \gamma < 1$ and $0 < \beta < 1$ and choose the affine function

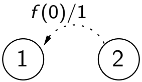
$$f : \mathbb{N} \cup \{0\} \to (0, \infty) \text{ given by } f(k) = \gamma k + \beta.$$

Model evolution: At time $n = 1$, we have a single vertex (labeled 1). In each time step $n \to n+1$ we

- add a new vertex labeled $n+1$, and
- for each $m \leq n$ independently introduce an oriented edge from the new vertex $n+1$ to the old vertex m with probability

$$\frac{f(\text{indegree of } m \text{ at time } n)}{n}.$$

Example:



Definition of the model

Take parameters $0 \le \gamma < 1$ and $0 < \beta < 1$ and choose the affine function

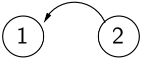
$$f : \mathbb{N} \cup \{0\} \to (0, \infty) \text{ given by } f(k) = \gamma k + \beta.$$

Model evolution: At time $n = 1$, we have a single vertex (labeled 1). In each time step $n \to n + 1$ we

- add a new vertex labeled $n + 1$, and
- for each $m \le n$ independently introduce an oriented edge from the new vertex $n + 1$ to the old vertex m with probability

$$\frac{f(\text{indegree of } m \text{ at time } n)}{n}.$$

Example:



Take parameters $0 \leq \gamma < 1$ and $0 < \beta < 1$ and choose the affine function

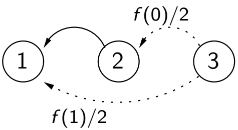
$$f : \mathbb{N} \cup \{0\} \to (0, \infty) \text{ given by } f(k) = \gamma k + \beta.$$

Model evolution: At time $n = 1$, we have a single vertex (labeled 1). In each time step $n \to n + 1$ we

- add a new vertex labeled $n + 1$, and
- for each $m \leq n$ independently introduce an oriented edge from the new vertex $n + 1$ to the old vertex m with probability

$$\frac{f(\text{indegree of } m \text{ at time } n)}{n}.$$

Example:



Definition of the model

Take parameters $0 \le \gamma < 1$ and $0 < \beta < 1$ and choose the affine function

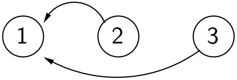
$$f : \mathbb{N} \cup \{0\} \to (0, \infty) \text{ given by } f(k) = \gamma k + \beta.$$

Model evolution: At time $n = 1$, we have a single vertex (labeled 1). In each time step $n \to n + 1$ we

- add a new vertex labeled $n + 1$, and
- for each $m \le n$ independently introduce an oriented edge from the new vertex $n + 1$ to the old vertex m with probability

$$\frac{f(\text{indegree of } m \text{ at time } n)}{n}.$$

Example:



Definition of the model

Take parameters $0 \leq \gamma < 1$ and $0 < \beta < 1$ and choose the affine function

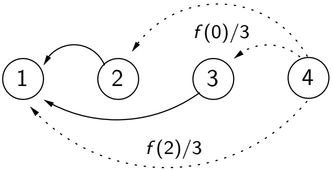
$$f : \mathbb{N} \cup \{0\} \to (0, \infty) \text{ given by } f(k) = \gamma k + \beta.$$

Model evolution: At time $n = 1$, we have a single vertex (labeled 1). In each time step $n \to n+1$ we

- add a new vertex labeled $n+1$, and
- for each $m \leq n$ independently introduce an oriented edge from the new vertex $n+1$ to the old vertex m with probability

$$\frac{f(\text{indegree of } m \text{ at time } n)}{n}.$$

Example:



Definition of the model

Take parameters $0 \leq \gamma < 1$ and $0 < \beta < 1$ and choose the affine function

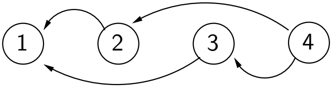
$$f: \mathbb{N} \cup \{0\} \to (0, \infty) \text{ given by } f(k) = \gamma k + \beta.$$

Model evolution: At time $n = 1$, we have a single vertex (labeled 1). In each time step $n \to n + 1$ we

- add a new vertex labeled $n + 1$, and
- for each $m \leq n$ independently introduce an oriented edge from the new vertex $n + 1$ to the old vertex m with probability

$$\frac{f(\text{indegree of } m \text{ at time } n)}{n}.$$

Example:



All edges are ordered from the younger to the older vertex. For the questions of interest, edges may be considered as unordered. We denote the resulting increasing sequence of graphs by $(\mathcal{G}_n)$.

Preferential attachment graphs are scale-free

Our first result shows that the preferential attachment graphs with $\gamma > 0$ are scale-free and identify the power law exponent.

Preferential attachment graphs are scale-free

Our first result shows that the preferential attachment graphs with $\gamma > 0$ are scale-free and identify the power law exponent.

Theorem 1 (M, Dereich 2008)

(a) Suppose X_n is the empirical indegree distribution of $\mathcal{G}_n$. Then

$$\lim_{n\uparrow\infty} X_n(k) = \mu(k) \quad \text{for all } k, \text{ almost surely,}$$

where

$$\mu(k) = \frac{1}{1+f(k)} \prod_{l=0}^{k-1} \frac{f(l)}{1+f(l)} \sim c\, k^{-(1+\frac{1}{\gamma})}.$$

(b) The conditional distribution of the outdegree of the $(n+1)$st vertex, given the graph at time n, converges almost surely in the variational topology to the Poisson distribution with parameter $\sum_{k=0}^{\infty} \mu(k) f(k)$.

Preferential attachment graphs are scale-free

Our first result shows that the preferential attachment graphs with $\gamma > 0$ are scale-free and identify the power law exponent.

Theorem 1 (M, Dereich 2008)

(a) Suppose X_n is the empirical indegree distribution of $\mathcal{G}_n$. Then

$$\lim_{n\uparrow\infty} X_n(k) = \mu(k) \quad \text{for all } k, \text{ almost surely,}$$

where

$$\mu(k) = \frac{1}{1+f(k)} \prod_{l=0}^{k-1} \frac{f(l)}{1+f(l)} \sim c\, k^{-(1+\frac{1}{\gamma})}.$$

(b) The conditional distribution of the outdegree of the $(n+1)$st vertex, given the graph at time n, converges almost surely in the variational topology to the Poisson distribution with parameter $\sum_{k=0}^{\infty} \mu(k) f(k)$.

$\Rightarrow$ The empirical degree distribution converges to a power law with exponent

$$\tau := 1 + \frac{1}{\gamma}.$$

Comparison with other models

In the preferential attachment model we have, for $m \leq n$, that

$$\mathbb{E}\big[\text{indegree of } m \text{ at time } n\big] \approx \Big(\frac{n}{m}\Big)^{\gamma},$$

and hence

$$\mathbb{P}\{m \leftrightarrow n\} \approx \frac{1}{(m \wedge n)^{\gamma}(m \vee n)^{1-\gamma}}.$$

Comparison with other models

In the preferential attachment model we have, for $m \leq n$, that

$$\mathbb{E}\big[\text{indegree of } m \text{ at time } n\big] \approx \Big(\frac{n}{m}\Big)^{\gamma},$$

and hence

$$\mathbb{P}\{m \leftrightarrow n\} \approx \frac{1}{(m \wedge n)^{\gamma}(m \vee n)^{1-\gamma}}.$$

If $\gamma \neq \frac{1}{2}$ this is not the product of a function of n with a function of m and hence preferential attachment is not a rank one model.

Comparison with other models

In the preferential attachment model we have, for $m \leq n$, that

$$\mathbb{E}\big[\text{indegree of } m \text{ at time } n\big] \approx \Big(\frac{n}{m}\Big)^{\gamma},$$

and hence

$$\mathbb{P}\{m \leftrightarrow n\} \approx \frac{1}{(m \wedge n)^{\gamma}(m \vee n)^{1-\gamma}}.$$

If $\gamma \neq \frac{1}{2}$ this is not the product of a function of n with a function of m and hence preferential attachment is not a rank one model.

This makes our model much more difficult to study than rank one models such as

- the LCD-model of Bollobas and Riordan, which is a preferential attachment model that necessarily has $\gamma = \frac{1}{2}$, or, equivalently, $\tau = 3$.
- the configuration model which, loosely speaking, is the uniform distribution on the collection of graphs with a given power law degree distribution.

I. Robustness of networks

Questions:

(1) For which parameters β, γ is there a giant component?

(2) When is the network robust?

(3) What is the size of the giant component near criticality?

I. Robustness of networks

Questions:

(1) For which parameters β, γ is there a giant component?

(2) When is the network robust?

(3) What is the size of the giant component near criticality?

Earlier work:

- In the case that $\gamma = 0$ so that there is no preferential attachment the model was first studied by Dubins (1984). The questions were answered by Shepp (1989) and Riordan (2005).
- For the LCD model the questions were answered by Bollobas, Riordan (2003) and Riordan (2005).

Coupling the graph to a branching process

The answers to our questions are based on a coupling of the neighbourhood of a randomly chosen vertex to the genealogy of a killed branching random walk.

Coupling the graph to a branching process

The answers to our questions are based on a coupling of the neighbourhood of a randomly chosen vertex to the genealogy of a killed branching random walk.

Particle positions are on the real line and types are given by the relative position of their father. Define the pure birth process $(Z_t \colon t \geq 0)$ by its generator

$$Lg(k) = f(k)\big(g(k+1) - g(k)\big).$$

Coupling the graph to a branching process

The answers to our questions are based on a coupling of the neighbourhood of a randomly chosen vertex to the genealogy of a killed branching random walk.

Particle positions are on the real line and types are given by the relative position of their father. Define the pure birth process $(Z_t : t \geq 0)$ by its generator

$$Lg(k) = f(k)\big(g(k+1) - g(k)\big).$$

A particle which has its parent to its left generates offspring

- to its right with relative positions at the jumps of the process $(Z_t : t \geq 0)$;
- to its left with relative positions distributed according to the Poisson process Π on $(-\infty, 0]$ with intensity measure $e^t\, \mathbb{E}[f(Z_{-t})]\, dt$.

Coupling the graph to a branching process

For $\tau > 0$ we let $(Z_t^{(\tau)} : t \geq 0)$ be the pure birth process $(Z_t : t \geq 0)$ conditioned to have a birth at time τ with that particle removed.

Coupling the graph to a branching process

For $\tau > 0$ we let $(Z_t^{(\tau)} : t \geq 0)$ be the pure birth process $(Z_t : t \geq 0)$ conditioned to have a birth at time τ with that particle removed.

Miracle: The law of $(Z_t^{(\tau)} : t \geq 0)$ does not depend on τ!

Coupling the graph to a branching process

For $\tau > 0$ we let $(Z_t^{(\tau)} : t \geq 0)$ be the pure birth process $(Z_t : t \geq 0)$ conditioned to have a birth at time τ with that particle removed.

Miracle: The law of $(Z_t^{(\tau)} : t \geq 0)$ does not depend on τ!

A particle which has its parent to its right generates offspring

- to its right with relative positions at the jumps of $(Z_t^{(\tau)}(t) : t \geq 0)$.
- to its left in the same manner as before.

We start the branching random walk with one initial particle in location $-X$, where X is standard exponential and kill particles and their offspring if their position is to the right of the origin.

Application of the branching process

- There exists a giant component iff
  the killed branching process survives with positive probability.

Application of the branching process

- There exists a giant component iff
  the killed branching process survives with positive probability.
- The size of the giant component equals
  the survival probability of the killed branching process.

Application of the branching process

- There exists a giant component iff
  the killed branching process survives with positive probability.
- The size of the giant component equals
  the survival probability of the killed branching process.
- The network is robust iff
  the branching process has infinite mean growth on survival.

Application of the branching process

- There exists a giant component iff
  the killed branching process survives with positive probability.
- The size of the giant component equals
  the survival probability of the killed branching process.
- The network is robust iff
  the branching process has infinite mean growth on survival.

For $0 < \alpha < 1$ and the unkilled branching process we define the matrix $A^{(\alpha)}$ by

$$A^{(\alpha)}_{i,j} := \mathbb{E}_i\Big[\sum_{|x|=1} e^{-\alpha \mathrm{Position}(x)} \mathbf{1}\{\mathrm{Type}(x) = j\}\Big]$$

where $i, j \in \{\text{left}, \text{right}\}$. Denote by $\rho(A^{(\alpha)})$ its spectral radius.

Application of the branching process

- There exists a giant component iff
  the killed branching process survives with positive probability.
- The size of the giant component equals
  the survival probability of the killed branching process.
- The network is robust iff
  the branching process has infinite mean growth on survival.

For $0 < \alpha < 1$ and the unkilled branching process we define the matrix $A^{(\alpha)}$ by

$$A^{(\alpha)} = \begin{pmatrix} \frac{\beta}{\alpha-\gamma} & \frac{\beta}{1-\alpha-\gamma} \\ \frac{\beta+\gamma}{\alpha-\gamma} & \frac{\beta}{1-\alpha-\gamma} \end{pmatrix}$$

where $i, j \in \{\text{left}, \text{right}\}$. Denote by $\rho(A^{(\alpha)})$ its spectral radius.

- The killed branching process dies iff
  there exists $0 < \alpha < 1$ with $\rho(A^{(\alpha)}) \leq 1$.

Application of the branching process

- There exists a giant component iff
  the killed branching process survives with positive probability.
- The size of the giant component equals
  the survival probability of the killed branching process.
- The network is robust iff
  the branching process has infinite mean growth on survival.

For $0 < \alpha < 1$ and the unkilled branching process we define the matrix $A^{(\alpha)}$ by

$$A_{i,j}^{(\alpha)} := \mathbb{E}_i\Big[\sum_{|x|=1} e^{-\alpha \mathrm{Position}(x)} \mathbf{1}\{\mathrm{Type}(x) = j\}\Big]$$

where $i, j \in \{\mathrm{left}, \mathrm{right}\}$. Denote by $\rho(A^{(\alpha)})$ its spectral radius.

- The killed branching process dies iff
  there exists $0 < \alpha < 1$ with $\rho(A^{(\alpha)}) \leq 1$.
- The killed branching process has infinite mean growth iff
  the matrix $A^{(\alpha)}$ is ill-defined for any $0 < \alpha < 1$.

Existence of a giant component

Theorem 2 Dereich, M (2010)

(a) A giant component exists if and only if

$$\gamma \geq \frac{1}{2} \text{ or } \beta > \frac{(\frac{1}{2}-\gamma)^2}{1-\gamma}.$$

Existence of a giant component

Theorem 2 Dereich, M (2010)

(a) A giant component exists if and only if

$$\gamma \geq \frac{1}{2} \text{ or } \beta > \frac{(\frac{1}{2} - \gamma)^2}{1 - \gamma}.$$

(b) The network is robust if and only if $\gamma \geq \frac{1}{2}$.

Existence of a giant component

Theorem 2 Dereich, M (2010)

(a) A giant component exists if and only if

$$\gamma \geq \frac{1}{2} \text{ or } \beta > \frac{(\frac{1}{2}-\gamma)^2}{1-\gamma}.$$

(b) The network is robust if and only if $\gamma \geq \dfrac{1}{2}$.

(c) If $\gamma < \frac{1}{2}$ the critical percolation parameter in the network is

$$p_c = \left(\tfrac{1}{2\gamma} - 1\right)\left(\sqrt{1+\tfrac{\gamma}{\beta}} - 1\right).$$

Existence of a giant component

Theorem 2 Dereich, M (2010)

(a) A giant component exists if and only if

$$\gamma \geq \frac{1}{2} \text{ or } \beta > \frac{(\frac{1}{2} - \gamma)^2}{1 - \gamma}.$$

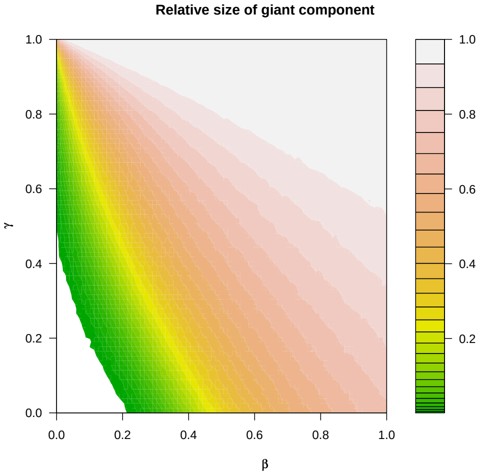
(b) The network is robust if and only if $\gamma \geq \dfrac{1}{2}$.

(c) If $\gamma < \frac{1}{2}$ the critical percolation parameter in the network is

$$p_c = \left(\frac{1}{2\gamma} - 1\right)\left(\sqrt{1 + \frac{\gamma}{\beta}} - 1\right).$$

Remark: We also have (slightly less explicit) results in the case that the attachment rule f is concave rather than affine.

Phase diagram



Simulation of the survival probability of the killed branching process.

Size of the giant component near criticality

Denote by $\zeta(\gamma,\beta)$ the size of the giant component which is the asymptotic proportion of vertices in the largest graph component.

Theorem 3 Eckhoff, M (2013)

(a) Let $0 \le \gamma < \frac{1}{2}$ and denote $\beta_c(\gamma) = \frac{(\frac{1}{2}-\gamma)^2}{1-\gamma}$. Then

$$\limsup_{\beta \downarrow \beta_c(\gamma)} \sqrt{\beta - \beta_c(\gamma)} \log \zeta(\gamma,\beta) \le -\frac{\pi}{2} \frac{1}{\sqrt{1-\gamma}}.$$

(b) Let $0 < \beta < \frac{1}{4}$ and denote $\gamma_c(\beta) = \frac{1}{2}(1 - \beta - \sqrt{\beta^2 + 2\beta})$. Then

$$\limsup_{\gamma \downarrow \gamma_c(\beta)} \sqrt{\gamma - \gamma_c(\beta)} \log \zeta(\gamma,\beta) \le -\frac{\pi}{2} \left(\frac{1}{\beta^2 + 2\beta}\right)^{\frac{1}{4}}.$$

Size of the giant component near criticality

Denote by $\zeta(\gamma,\beta)$ the size of the giant component which is the asymptotic proportion of vertices in the largest graph component.

Theorem 3 Eckhoff, M (2013)

(a) Let $0 \leq \gamma < \frac{1}{2}$ and denote $\beta_c(\gamma) = \frac{(\frac{1}{2}-\gamma)^2}{1-\gamma}$. Then

$$\limsup_{\beta \downarrow \beta_c(\gamma)} \sqrt{\beta - \beta_c(\gamma)} \log \zeta(\gamma,\beta) \leq -\frac{\pi}{2} \frac{1}{\sqrt{1-\gamma}}.$$

(b) Let $0 < \beta < \frac{1}{4}$ and denote $\gamma_c(\beta) = \frac{1}{2}(1 - \beta - \sqrt{\beta^2 + 2\beta})$. Then

$$\limsup_{\gamma \downarrow \gamma_c(\beta)} \sqrt{\gamma - \gamma_c(\beta)} \log \zeta(\gamma,\beta) \leq -\frac{\pi}{2} \left(\frac{1}{\beta^2 + 2\beta}\right)^{\frac{1}{4}}.$$

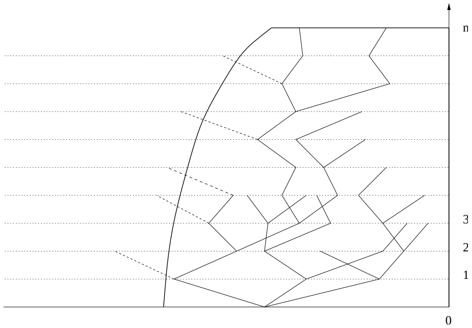
Remark: This is in stark contrast to the behaviour of the configuration model!

Size of the giant component: Idea of proof

We need to give an upper bound on the survival probability of the killed branching process. Our arguments adapt a technique from a recent paper by Gantert, Hu and Shi (2011) to our setup.

Size of the giant component: Idea of proof

We need to give an upper bound on the survival probability of the killed branching process. Our arguments adapt a technique from a recent paper by Gantert, Hu and Shi (2011) to our setup.



Main ingredients of the proof

- the basis is a truncated first moment method,
- use a many-to-one lemma,
- adapt and use a large deviation theorem for Markov chains.

II. Vulnerability of the network

Besides their robustness under random attack one of the characteristic features of many real networks is their vulnerability to targeted attack. Removal of a small number of key hubs typically changes their behaviour dramatically. We are going to explore this feature in our model of preferential attachment graphs.

II. Vulnerability of the network

Besides their robustness under random attack one of the characteristic features of many real networks is their vulnerability to targeted attack. Removal of a small number of key hubs typically changes their behaviour dramatically. We are going to explore this feature in our model of preferential attachment graphs.

From now on our focus is on the case $\gamma \geq \frac{1}{2}$ when the network is robust. To simulate a targeted attack on the graph we remove a small number of the oldest vertices and investigate how this changes the behaviour of the network.

II. Vulnerability of the network

Besides their robustness under random attack one of the characteristic features of many real networks is their vulnerability to targeted attack. Removal of a small number of key hubs typically changes their behaviour dramatically. We are going to explore this feature in our model of preferential attachment graphs.

From now on our focus is on the case $\gamma \geq \frac{1}{2}$ when the network is robust. To simulate a targeted attack on the graph we remove a small number of the oldest vertices and investigate how this changes the behaviour of the network.

Questions:

- How many old vertices have to be removed to destroy the robustness?

II. Vulnerability of the network

Besides their robustness under random attack one of the characteristic features of many real networks is their vulnerability to targeted attack. Removal of a small number of key hubs typically changes their behaviour dramatically. We are going to explore this feature in our model of preferential attachment graphs.

From now on our focus is on the case $\gamma \geq \frac{1}{2}$ when the network is robust. To simulate a targeted attack on the graph we remove a small number of the oldest vertices and investigate how this changes the behaviour of the network.

Questions:

- How many old vertices have to be removed to destroy the robustness?
- How does the critical percolation probability behave when the number of removed vertices approaches criticality?

II. Vulnerability of the network

Besides their robustness under random attack one of the characteristic features of many real networks is their vulnerability to targeted attack. Removal of a small number of key hubs typically changes their behaviour dramatically. We are going to explore this feature in our model of preferential attachment graphs.

From now on our focus is on the case $\gamma \geq \frac{1}{2}$ when the network is robust. To simulate a targeted attack on the graph we remove a small number of the oldest vertices and investigate how this changes the behaviour of the network.

Questions:

- How many old vertices have to be removed to destroy the robustness?
- How does the critical percolation probability behave when the number of removed vertices approaches criticality?
- How are other features like
    - the power law property of the network,
    - the largest degree in the network,
    - the typical distance of vertices in the network,

  affected by the attack?

Loss of connectivity

It turns out that in order for an attack to lead to a qualitative change of the network we have to remove an arbitrarily small proportion ε of the oldest vertices.

Loss of connectivity

It turns out that in order for an attack to lead to a qualitative change of the network we have to remove an arbitrarily small proportion ε of the oldest vertices.

We denote by $\mathcal{G}_n^{(\varepsilon)}$ the graph $\mathcal{G}_n$ with vertices $1, 2, \ldots, \varepsilon n$ and adjacent edges removed.

Loss of connectivity

It turns out that in order for an attack to lead to a qualitative change of the network we have to remove an arbitrarily small proportion ε of the oldest vertices.

We denote by $\mathcal{G}_n^{(\varepsilon)}$ the graph $\mathcal{G}_n$ with vertices $1, 2, \ldots, \varepsilon n$ and adjacent edges removed.

Theorem 5 Eckhoff, M (2012)

Let $p_c(\varepsilon)$ be the critical percolation probability of $\mathcal{G}_n^{(\varepsilon)}$.

(a) If $\gamma = \frac{1}{2}$, then as $\varepsilon \downarrow 0$ we have

$$p_c(\varepsilon) \asymp \log(1/\varepsilon)^{-1}.$$

(b) If $\gamma > \frac{1}{2}$, then as $\varepsilon \downarrow 0$ we have

$$p_c(\varepsilon) = \frac{2\gamma-1}{\sqrt{\beta(\gamma+\beta)}}\, \varepsilon^{\gamma-\frac{1}{2}}\, \big(1 + o(1)\big).$$

Loss of connectivity

It turns out that in order for an attack to lead to a qualitative change of the network we have to remove an arbitrarily small proportion ε of the oldest vertices.

We denote by $\mathcal{G}_n^{(\varepsilon)}$ the graph $\mathcal{G}_n$ with vertices $1, 2, \ldots, \varepsilon n$ and adjacent edges removed.

Theorem 5 Eckhoff, M (2012)

Let $p_c(\varepsilon)$ be the critical percolation probability of $\mathcal{G}_n^{(\varepsilon)}$.

(a) If $\gamma = \frac{1}{2}$, then as $\varepsilon \downarrow 0$ we have

$$p_c(\varepsilon) \asymp \log(1/\varepsilon)^{-1}.$$

(b) If $\gamma > \frac{1}{2}$, then as $\varepsilon \downarrow 0$ we have

$$p_c(\varepsilon) = \frac{2\gamma - 1}{\sqrt{\beta(\gamma+\beta)}}\, \varepsilon^{\gamma - \frac{1}{2}} \big(1 + o(1)\big).$$

Remark: In (a) we conjecture $p_c(\varepsilon) = \frac{1}{\sqrt{\beta(\gamma+\beta)}} \log(1/\varepsilon)^{-1} \big(1 + o(1)\big)$.

Loss of connectivity

It turns out that in order for an attack to lead to a qualitative change of the network we have to remove an arbitrarily small proportion ε of the oldest vertices.

We denote by $\mathcal{G}_n^{(\varepsilon)}$ the graph $\mathcal{G}_n$ with vertices $1, 2, \dots, \varepsilon n$ and adjacent edges removed.

Theorem 5 Eckhoff, M (2012)

Let $p_c(\varepsilon)$ be the critical percolation probability of $\mathcal{G}_n^{(\varepsilon)}$.

(a) If $\gamma = \frac{1}{2}$, then as $\varepsilon \downarrow 0$ we have

$$p_c(\varepsilon) \asymp \log(1/\varepsilon)^{-1}.$$

(b) If $\gamma > \frac{1}{2}$, then as $\varepsilon \downarrow 0$ we have

$$p_c(\varepsilon) = \frac{2\gamma-1}{\sqrt{\beta(\gamma+\beta)}}\,\varepsilon^{\gamma-\frac{1}{2}}\,\big(1+o(1)\big).$$

Remark: A similar result holds for the configuration model but the exponent in (b) is doubled from $\gamma - \frac{1}{2}$ to $2\gamma - 1$.

Loss of connectivity: Idea of proof

Loss of connectivity: Idea of proof

Removing a proportion ε of the oldest vertices means that for the branching process a second killing boundary at $\log \varepsilon$ is introduced.

Loss of connectivity: Idea of proof

Removing a proportion ε of the oldest vertices means that for the branching process a second killing boundary at $\log \varepsilon$ is introduced.

Recall: In the case of one killing boundary at zero, on survival of the branching process the leftmost particle converges to $-\infty$ and therefore away from the killing boundary. The analysis of survival could therefore be based on the operators $A^{(\alpha)}$ relating to the unkilled process.

Loss of connectivity: Idea of proof

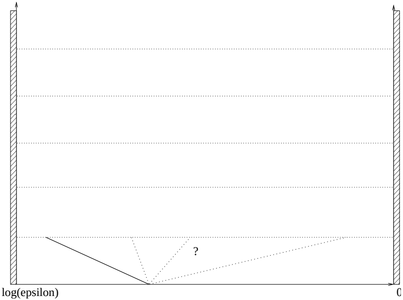
Removing a proportion ε of the oldest vertices means that for the branching process a second killing boundary at $\log \varepsilon$ is introduced.

Recall: In the case of one killing boundary at zero, on survival of the branching process the leftmost particle converges to $-\infty$ and therefore away from the killing boundary. The analysis of survival could therefore be based on the operators $A^{(\alpha)}$ relating to the unkilled process.

Now: In the case of two killing boundary at $\log \varepsilon$ and zero, the ancestral line of a surviving particle is confined to the strip $[\log \varepsilon, 0] \times \mathbb{N}$. What is the optimal strategy for a particle to survive in such a strip?

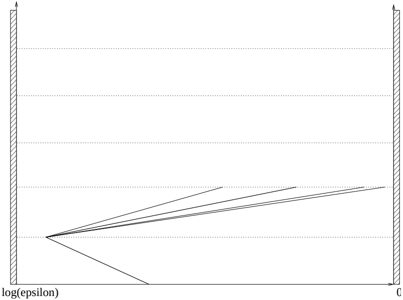
Loss of connectivity: Idea of proof

Removing a proportion ε of the oldest vertices means that for the branching process a second killing boundary at $\log \varepsilon$ is introduced.



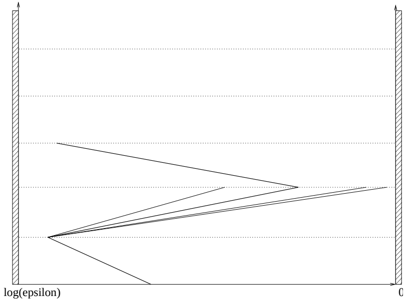
Loss of connectivity: Idea of proof

Removing a proportion ε of the oldest vertices means that for the branching process a second killing boundary at $\log \varepsilon$ is introduced.



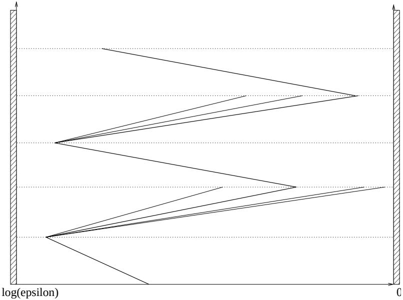
Loss of connectivity: Idea of proof

Removing a proportion ε of the oldest vertices means that for the branching process a second killing boundary at $\log \varepsilon$ is introduced.



Loss of connectivity: Idea of proof

Removing a proportion ε of the oldest vertices means that for the branching process a second killing boundary at $\log \varepsilon$ is introduced.



Loss of connectivity: Idea of proof

Removing a proportion ε of the oldest vertices means that for the branching process a second killing boundary at $\log \varepsilon$ is introduced.

Steps of the proof:

Loss of connectivity: Idea of proof

Removing a proportion ε of the oldest vertices means that for the branching process a second killing boundary at $\log \varepsilon$ is introduced.

Steps of the proof:

- the process survives if the associated expectation operator A has $\rho(A) > 1$,

Loss of connectivity: Idea of proof

Removing a proportion ε of the oldest vertices means that for the branching process a second killing boundary at $\log \varepsilon$ is introduced.

Steps of the proof:

- the process survives if the associated expectation operator A has $\rho(A) > 1$,
- in this case the critical percolation probability is $1/\rho(A)$,

Loss of connectivity: Idea of proof

Removing a proportion ε of the oldest vertices means that for the branching process a second killing boundary at $\log \varepsilon$ is introduced.

Steps of the proof:

- the process survives if the associated expectation operator A has $\rho(A) > 1$,
- in this case the critical percolation probability is $1/\rho(A)$,
- for upper bounds guess the eigenfunction ϕ,

Loss of connectivity: Idea of proof

Removing a proportion ε of the oldest vertices means that for the branching process a second killing boundary at $\log \varepsilon$ is introduced.

Steps of the proof:

- the process survives if the associated expectation operator A has $\rho(A) > 1$,
- in this case the critical percolation probability is $1/\rho(A)$,
- for upper bounds guess the eigenfunction ϕ,
- for lower bounds write $A = A_1 + A_2$ for a left offspring operator A_1 and a right offspring operator A_2, and show that main contribution to A^n comes from alternating products of A_1 and A_2.

Loss of the power law

Theorem 6 Eckhoff, M (2012)

Suppose $X_n^{(\varepsilon)}$ is the empirical indegree distribution of $\mathcal{G}_n^{(\varepsilon)}$. Then

$$\lim_{n\uparrow\infty} X_n(k) = \mu^{(\varepsilon)}(k) \quad \text{for all } k, \text{ almost surely,}$$

where the limiting measure $\mu^{(\varepsilon)}$ is deterministic and satisfies

$$\lim_{k\to\infty} \frac{1}{k} \log \mu^{(\varepsilon)}(k,\infty) = \log(1-\varepsilon^{\gamma}).$$

Loss of the power law

Theorem 6 Eckhoff, M (2012)

Suppose $X_n^{(\varepsilon)}$ is the empirical indegree distribution of $\mathcal{G}_n^{(\varepsilon)}$. Then

$$\lim_{n\uparrow\infty} X_n(k) = \mu^{(\varepsilon)}(k) \quad \text{for all } k, \text{ almost surely,}$$

where the limiting measure $\mu^{(\varepsilon)}$ is deterministic and satisfies

$$\lim_{k\to\infty} \frac{1}{k} \log \mu^{(\varepsilon)}(k,\infty) = \log(1-\varepsilon^{\gamma}).$$

Remark: After the attack the network is no longer scale-free. Instead its asymptotic degree distribution has exponential tails.

Increase of typical distances

Preferential attachment graphs with $\gamma > \frac{1}{2}$ are ultrasmall. More precisely, for vertices V_n, W_n chosen uniformly from the giant component of $\mathcal{G}_n$ we have

$$d_{\mathcal{G}_n}(V_n, W_n) \sim \frac{4 \log\log n}{\log \frac{\gamma}{1-\gamma}} \qquad \text{in probability,}$$

see Dereich, Mönch, M (2011) and Dommers, van der Hofstad, Hooghiemstra (2010).

Increase of typical distances

Preferential attachment graphs with $\gamma > \frac{1}{2}$ are ultrasmall. More precisely, for vertices V_n, W_n chosen uniformly from the giant component of $\mathcal{G}_n$ we have

$$d_{\mathcal{G}_n}(V_n, W_n) \sim \frac{4 \log\log n}{\log \frac{\gamma}{1-\gamma}} \qquad \text{in probability,}$$

see Dereich, Mönch, M (2011) and Dommers, van der Hofstad, Hooghiemstra (2010).

Theorem 7 Eckhoff, M (2012)

Suppose that $\varepsilon > 0$ is sufficiently small so that $\mathcal{G}_n^{(\varepsilon)}$ has a giant component. and let V_n, W_n be independent, uniformly chosen vertices in $\mathcal{G}_n$. Then, for all $\delta > 0$,

$$d_{\mathcal{G}_n^{(\varepsilon)}}(V_n, W_n) \geq \frac{1-\delta}{\log(1/p_c(\varepsilon))} \log n \qquad \text{with high probability.}$$

Increase of typical distances

Preferential attachment graphs with $\gamma > \frac{1}{2}$ are ultrasmall. More precisely, for vertices V_n, W_n chosen uniformly from the giant component of $\mathcal{G}_n$ we have

$$d_{\mathcal{G}_n}(V_n, W_n) \sim \frac{4 \log \log n}{\log \frac{\gamma}{1-\gamma}} \qquad \text{in probability,}$$

see Dereich, Mönch, M (2011) and Dommers, van der Hofstad, Hooghiemstra (2010).

Theorem 7 Eckhoff, M (2012)

Suppose that $\varepsilon > 0$ is sufficiently small so that $\mathcal{G}_n^{(\varepsilon)}$ has a giant component. and let V_n, W_n be independent, uniformly chosen vertices in $\mathcal{G}_n$. Then, for all $\delta > 0$,

$$d_{\mathcal{G}_n^{(\varepsilon)}}(V_n, W_n) \geq \frac{1-\delta}{\log(1/p_c(\varepsilon))} \log n \qquad \text{with high probability.}$$

Remarks: After the attack the network is no longer ultrasmall, and typical distances are now logarithmic, i.e. much larger. The lower bound is conjectured to be sharp.

Summary

- Preferential attachment graphs are built dynamically using a simple and universal principle. From this principle many properties of complex real world networks emerge.

  In particular preferential attachment graphs are scale free, and if the power-law exponent is in the range $2 < \tau < 3$, they are

  - ultrasmall,
  - robust to blind attack, but
  - vulnerable to targeted attack.

Summary

- Preferential attachment graphs are built dynamically using a simple and universal principle. From this principle many properties of complex real world networks emerge.

  In particular preferential attachment graphs are scale free, and if the power-law exponent is in the range $2 < \tau < 3$, they are

  - ultrasmall,
  - robust to blind attack, but
  - vulnerable to targeted attack.
- A targeted attack drastically changes the topology of the graph, and the network is no longer scale-free, ultrasmall or robust.

Summary

- Preferential attachment graphs are built dynamically using a simple and universal principle. From this principle many properties of complex real world networks emerge.

  In particular preferential attachment graphs are scale free, and if the power-law exponent is in the range $2 < \tau < 3$, they are

  - ultrasmall,
  - robust to blind attack, but
  - vulnerable to targeted attack.
- A targeted attack drastically changes the topology of the graph, and the network is no longer scale-free, ultrasmall or robust.
- Preferential attachment graphs are mathematically tractable. The main tool for such an analysis is an approximation of local neighbourhoods by branching random walks and advanced methods from the theory of branching processes.

Summary

- Preferential attachment graphs are built dynamically using a simple and universal principle. From this principle many properties of complex real world networks emerge.

  In particular preferential attachment graphs are scale free, and if the power-law exponent is in the range $2 < \tau < 3$, they are

  - ultrasmall,
  - robust to blind attack, but
  - vulnerable to targeted attack.
- A targeted attack drastically changes the topology of the graph, and the network is no longer scale-free, ultrasmall or robust.
- Preferential attachment graphs are mathematically tractable. The main tool for such an analysis is an approximation of local neighbourhoods by branching random walks and advanced methods from the theory of branching processes.

Thank you very much for your attention!