Analysis of Complex Networks Lecture 3: Galton-Watson Branching Process

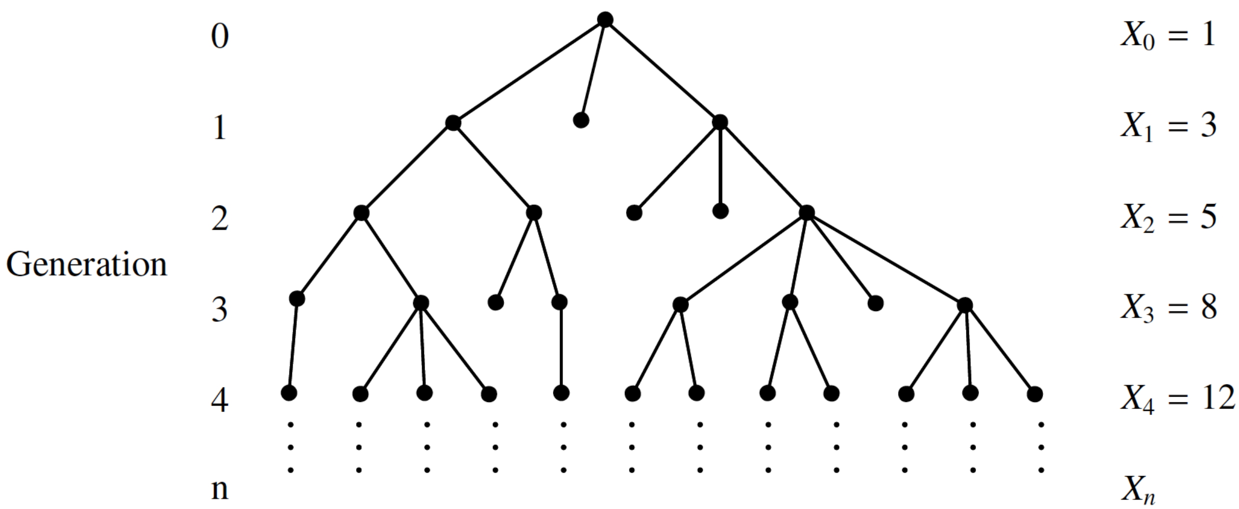
Prof. Yung Yi, yiyung@kaist.edu,
http://lanada.kaist.ac.kr
EE, KAIST

September 7, 2016

Contents

- What is Galton-Watson Branching process (or simply branching process)?
- Phase transition in the process
- Later, connection to the analysis of when a giant component of ER graph emerges

GW Branching Process



- X_n: the number of individuals at generation n. Then,

$$X_{n+1} = \sum_{i=1}^{X_n} \xi_{n,i},$$

  where $\xi_{n,i}$ is a random variable which represents the number of children of i-th individual of the n-th generation.

- Assume that $\xi_{n,i}$ is i.i.d. over n and i, thus we don't care about n and i and use the notation $\xi = \xi_{n,i}$.

- Denote by p_k the distribution of ξ, i.e,

$$p_k := \mathbb{P}[\xi = k], \quad \text{for all non-negative integer } k.$$

- Then, a branching process $\mathcal{T}$ is characterized by $\{p_k\}$.

- $\{p_k\}$ is often called the offspring distribution.

- Graph theoretically, it is an oriented tree spanning the descendants of the ancestor and rooted at the ancestor.

Questions

- Q1. Under what conditions of $\{p_k\}$, $\mathcal{T}$ experiences extinction?
- Q2. If $\mathcal{T}$ becomes extinct, what is the total population size?
- Let p_{ext} be the extinction probability, i.e., the probability that $X_n = 0$ for some finite n.

Depth-first Exploration

- For any child of the root, the subtree rooted at this child has the same statistical properties:

- Denote by $|\mathcal{T}|$ the number of nodes of $\mathcal{T}$. Then, we have:

$$\begin{aligned} p_{\text{ext}} &= \mathbb{P}[|\mathcal{T}| < \infty] \\ &= \\ &= \quad . \end{aligned} \tag{1}$$

- Consider the following generating function of ξ: $\phi_\xi(s) := \sum_{k \in N} p_k s^k$. Then, from (1), p_{ext} is the solution of:

$$x = \quad . \tag{2}$$

Depth-first Exploration: Total population $\mathcal{X}$

- Total population $\mathcal{X} = \sum_{k=0}^{\infty} X_k$, and let its generation function $\phi_{\mathcal{X}}$.

- Theorem. $\phi_{\mathcal{X}}(s) = s\phi_{\xi}(\phi_{\mathcal{X}}(s))$.

- Meaning of Theorem
  - One-to-one correspondence between a random variable and its generating function
  - Once ξ's generation function is known, we can know the characterization of $\mathcal{X}$, hopefully, the distribution of the total population.

- Proof.

Breadth-first Exploration

- Let's investigate (2) more rigorously. From this, we can provide more explicit characterization of p_{ext}.

- Denote by $\mu := \mathbb{E}[\xi]$, i.e., the average number of children per one individual.

- Theorem. First, p_{ext} is the smallest solution of (2). Second, the following holds:

  (i) Subcritical regime: If $\mu < 1$, then $p_{\text{ext}} = 1$.
  (ii) Critical regime: If $\mu = 1$ and $p_1 < 1$, then $p_{\text{ext}} = 1$.
  (iii) Supercritical regime: If $\mu > 1$, then $p_{\text{ext}} < 1$.

- Meaning

  - If the offspring distribution is Poisson with parameter λ, how can we find p_{ext}?

  - (ii)?

- Proof.

.

One-by-one Exploration

- Throughout the exploration, one keeps track of a number of so-called active nodes, i.e., they have been unveiled by our exploration but their children have not yet unveiled.

- Starting with the ancestor as the only active node, each step of the exploration consists of: One (i) picks (in an arbitrary fashion) some active node, say u, (ii) unveils its children (who thus become active) and (iii) deactivates the originally selected active node u.

- Later, this exploration method will be used to compute the bound of the total population's distribution, i.e., $\mathbb{P}[X \geq k]$.

- Let
  - A_n: the number of active nodes at step n
  - ξ_n: the number of children of the node chosen at step n (which follows the offspring distribution $\xi = \{p_k\}$ due to i.i.d. assumption).

- Then,

$$A_0 = 1,$$
$$A_n = A_{n-1} - 1 + \xi_n, \quad n \geq 1.$$

- This exploration stops at step $\mathcal{Y}$, where

$$\mathcal{Y} = \inf\{n > 0 : A_n = 0\},$$

i.e., at the step when no active nodes to pick exists.

- $\mathcal{Y}$ is actually the total population size, i.e., $\mathcal{Y} = \mathcal{X}$. why?

Two Definitions

- Definition. (History) The history of a branching process (due to one-by-one exploration) is given by the sequence $H = \{\xi_1, \ldots, \xi_{\mathcal{X}}\}$, satisfying the constraints:

$$A_n > 0, \quad n = 0, \ldots, \mathcal{X} - 1,$$
$$A_{\mathcal{X}} = 0,$$

- The distribution of the history H follows:

$$\mathbb{P}[H = (x_1, \ldots, x_k)] = \qquad\qquad\qquad\qquad .$$

- Definition. (Exponentially tilted distribution) For a given reference offspring distribution $\{q_k\}$ with $q_0 > 0$ $\sum_{k \geq 0} k q_k = 1$. Using a parameter λ, we define the corresponding exponentially tilted distribution, denoted by $\{p_k(\lambda)\}$, by:

$$p_k(\lambda) = q_k \frac{\lambda^k}{\phi(\lambda)},$$

where $\phi(\lambda)$ is a normalization term to make it a probability distribution, i.e.,

$$\phi(\lambda) := \sum_{k \geq 0} q_k \lambda^k.$$

- Example. When $q_k = \frac{e^{-1}}{k!}$, what is the distribution of $p_k(\lambda)$?

- If $\lambda > 1$ (resp. $\lambda < 1$) the offstring distribution $p_k(\lambda)$ leads to a supercritical (resp. subscritical) process. Why?

Dual Parameter

- Definition. Given a reference distribution $\{q_k\}$, the parameter μ is said to be dual (or conjugate) to λ, if

$$\frac{\phi(\lambda)}{\lambda} = \frac{\phi(\mu)}{\mu}.$$

  - How to drive the conjugate of a given parameter? Next proposition.

- Proposition. Given a parameter $\lambda > 1$, there is a unique conjugate parameter $\mu \neq \lambda$. It satisfies $\mu < 1$, given by:

$$\mu = \lambda p_{\text{ext}}(\lambda),$$

  where $p_{\text{ext}}(\lambda)$ is the extinction probability associated with the offspring distribution $\{p_k(\lambda)\}$.

- Proof.

Duality of Branching Process

- Theorem. Fix $\lambda > 1$. The distribution of the history $\{\xi_1, \dots, \xi_{\mathcal{X}}\}$ under the offspring distribution $\{p_k(\lambda)\}$, conditioned on extinction, coincides with the distribution of the history under offspring distribution $\{p_k(\mu)\}$, where μ is the dual parameter of λ, i.e., $\mu = \lambda p_{\text{ext}}(\lambda)$.

- Meaning
  - Supercritical branching process is highly complex to see, especially, conditioned on extinction. Can we look at this complex supercritical branching process from a simple structural viewpoint?
  - Yes. It can be characterized by its dual subcritical branching process.

- Proof.

.

Total population size

- Next question: I want to know the statistical property of the total population size.

- Distribution of $\mathcal{X}$?

- In other words, $\mathbb{P}[\mathcal{X} > x]$

- We need a new mathematical tool called Chernoff bound.