

Social Networks & Recommendation Systems

VII. Probabilistic aspects of the complex networks.

Grzegorz Siudem

Warsaw University of Technology



**European
Funds**
Knowledge Education Development

**Warsaw University
of Technology**

European Union
European Social Fund



MSc program in Data Science has been developed
as a part of task 10 of the project
„NERW PW. Science - Education - Development - Cooperation”
co-funded by European Union from European Social Fund.

Project

Master equation for BA networks - case study

$$N_k(t+1) - N_k(t) = \frac{k-1}{2t} N_{k-1}(t) - \frac{k}{2t} N_k(t) + \delta_{km}$$
$$N_{m-1}(t) = 0.$$

P7.1 Solve (asymptotically) master equation i.e. find the asymptotic $t \rightarrow \infty$ formula for the distribution of the degrees. [20% for following the idea from the next slides or 50% for different approach]

Brainstorming:

What tools and ideas do we have?

$$N_k(t+1) - N_k(t) = \frac{k-1}{2t} N_{k-1}(t) - \frac{k}{2t} N_k(t) + \delta_{km}$$
$$N_{m-1}(t) = 0.$$

Spoiler alert!

Master equation for BA networks - *for amateurs*

$$N_k(t+1) - N_k(t) = \frac{k-1}{2t} N_{k-1}(t) - \frac{k}{2t} N_k(t) + \delta_{km}$$
$$N_{m-1}(t) = 0.$$

1. hint

Let us start with N_m

$$N_m(t+1) = c + \left(1 - \frac{b(t)}{t}\right) N_m(t),$$

where in our case $c = 1$, $b(t) = m/2$.

Lemma 1.

$$N_m(t)/t \rightarrow \frac{c}{1+b},$$

where $b(t) \rightarrow b$.

Master equation for BA networks - *for amateurs*

$$N_k(t+1) - N_k(t) = \frac{k-1}{2t} N_{k-1}(t) - \frac{k}{2t} N_k(t) + \delta_{km}$$
$$N_{m-1}(t) = 0.$$

2. hint

Let us consider now N_k for $k > m$

$$N_k(t+1) = g(t) + \left(1 - \frac{b(t)}{t}\right) N_k(t),$$

where $g(t) \rightarrow g$, $b(t) \rightarrow b$, in our case $g(t) = (k-1)N_{k-1}(t)/2t$ (Why does it converge?), $b(t) = k/2$.

Lemma 2.

$$N_k(t)/t \rightarrow \frac{g}{1+b},$$

Master equation for BA networks - *for amateurs*

$$N_k(t+1) - N_k(t) = \frac{k-1}{2t} N_{k-1}(t) - \frac{k}{2t} N_k(t) + \delta_{km}$$
$$N_{m-1}(t) = 0.$$

3. hint

We are looking for

$$\mathcal{P}(k) = \lim_{t \rightarrow \infty} \frac{N_k t}{t},$$

according to Lemma 1 (why?) we have:

$$\mathcal{P}(m) = \frac{2}{m+2},$$

according to Lemma 2 (why?)

$$\mathcal{P}(k) = \mathcal{P}(k-1) \frac{k-1}{k+2}.$$

- P7.2 Add the above solutions for the plots from the previous project. [10%]
- P7.3 Test which solution fits better to the simulation's data. [10%]

Percolation in ER graphs

Let us introduce distribution $\mathcal{Q}(k)$

$\mathcal{Q}(k)$ is the distribution of the degrees on the ends on the randomly selected edge.

P7.4 Justify, that for non-correlated graphs $\mathcal{Q}(k) \propto k\mathcal{P}(k)$, and thus

$$\mathcal{Q}(k) = \frac{k}{\langle k \rangle} \mathcal{P}(k). \quad [20\%]$$

P7.5 Justify that percolation threshold may be defined as follows

$$\sum_k k \mathcal{Q}(k) \geq 2,$$

which is equivalent to

$$\langle k \rangle_{nn} = \frac{\langle k^2 \rangle}{\langle k \rangle} = 2. \quad [30\%]$$

$$\langle k \rangle_{nn} = \frac{\langle k^2 \rangle}{\langle k \rangle} = 2.$$

P7.6 Prove that the above condition for ER graphs leads to $\langle k \rangle = 1$, which implies $p_c = \frac{1}{N}$. [20%]

P7.7 Check with simulation the above result. Plot the size of the largest cluster as a function of $\langle k \rangle = pN$. [20%]

Generating functions formalism

To determine the size of the largest cluster, we will use the formalism of generating functions:

- $G_0(x) = \sum_{k=0}^{\infty} \mathcal{P}(k)x^k$,
- $\mathcal{F}(k) = \mathcal{Q}(k+1) = \frac{k+1}{\langle k \rangle} \mathcal{P}(k+1)$ (why such definition is more convenient?),
- $G_1(x) = \sum_{k=0}^{\infty} \mathcal{F}(k)x^k$.

P7.8 Prove that [20% for all]

- $G_0(1) = 1$,
- $G_0^{(n)}(1) = \langle k^n \rangle$,
- $G_1(x) = \frac{G_0'(x)}{G_0'(1)}$,
- $G_1'(1) = \frac{\langle k^2 \rangle}{\langle k \rangle} - 1$.

Generating functions for ER graphs

- $G_0(x) = \sum_{k=0}^{\infty} \mathcal{P}(k)x^k,$
- $\mathcal{F}(k) = \mathcal{Q}(k+1) = \frac{k+1}{\langle k \rangle} \mathcal{P}(k+1),$
- $G_1(x) = \sum_{k=0}^{\infty} \mathcal{F}(x)x^k.$

P7.9 Find function G_0 for ER graphs, assuming that $\mathcal{P}(k)$ is poissonian. [20%]

Generating functions for ER graphs

- $G_0(x) = \sum_{k=0}^{\infty} \mathcal{P}(k)x^k,$
- $\mathcal{F}(k) = \mathcal{Q}(k+1) = \frac{k+1}{\langle k \rangle} \mathcal{P}(k+1),$
- $G_1(x) = \sum_{k=0}^{\infty} \mathcal{F}(k)x^k.$

P7.9 Find function G_0 for ER graphs, assuming that $\mathcal{P}(k)$ is poissonian. [20%]

Solution:

$$G_0(x) = e^{\langle k \rangle (x-1)}.$$

Generating functions for ER graphs

- $G_0(x) = \sum_{k=0}^{\infty} \mathcal{P}(k)x^k,$
- $\mathcal{F}(k) = \mathcal{Q}(k+1) = \frac{k+1}{\langle k \rangle} \mathcal{P}(k+1)$
- $G_1(x) = \sum_{k=0}^{\infty} \mathcal{F}(x)x^k.$

P7.10 Find function G_1 for ER graphs, assuming that $\mathcal{P}(k)$ is poissonian. [20%]

Generating functions for ER graphs

- $G_0(x) = \sum_{k=0}^{\infty} \mathcal{P}(k)x^k,$
- $\mathcal{F}(k) = \mathcal{Q}(k+1) = \frac{k+1}{\langle k \rangle} \mathcal{P}(k+1)$
- $G_1(x) = \sum_{k=0}^{\infty} \mathcal{F}(k)x^k.$

P7.10 Find function G_1 for ER graphs, assuming that $\mathcal{P}(k)$ is poissonian. [20%]

Solution:

$$G_1(x) = e^{\langle k \rangle (x-1)}.$$

Generating functions

Let us define H_1

$H_1(x)$ is a generating function of the distribution $h_1(s)$, that walking in any direction from a randomly chosen edge of the considered graph (not only ER!) in finite number of steps one can reach s vertices.

P7.11 Justify that

$$h_1(s) = \sum_{k=0}^{s-1} \mathcal{F}(k) h_k(s-1),$$

where $h_k(s-1)$ is the probability that going through k different edges one can get to $s-1$ vertices. [20%]

Generating functions

$$h_1(s) = \sum_{k=0}^{s-1} \mathcal{F}(k) h_k(s-1),$$

P7.12 Prove that when

$$H_k(x) = \sum_{s=0}^{\infty} h_k(s) x^s$$

and

$$h_0(s) = \sum_{k=0}^{\infty} \mathcal{P}(k) h_k(s)$$

one gets [30%]

- $H_1(x) = x \sum_{k=0}^{\infty} \mathcal{F}(k) H_k(x),$
- $H_k(x) = [H_1(x)]^k,$
- $H_1(x) = x G_1(H_1(x)),$
- $H_0(x) = x G_0(H_1(x)).$

Average size of the cluster before percolation

$$H_1(x) = xG_1(H_1(x))$$

P7.13 Compute the average size of the cluster given (why?) with the following formula

$$\langle s \rangle = H'_0(1) = 1 + G'_0(1)H'_1(1) \quad [20\%]$$

Average size of the cluster before percolation

$$H_1(x) = xG_1(H_1(x))$$

P7.13 Compute the average size of the cluster given (why?) with the following formula

$$\langle s \rangle = H'_0(1) = 1 + G'_0(1)H'_1(1) \quad [20\%]$$

Solution:

$$\langle s \rangle = 1 + \frac{\langle k \rangle^2}{2\langle k \rangle - \langle k^2 \rangle}.$$

The size of the percolation cluster

Probability \mathcal{P}_∞

Let us define \mathcal{P}_∞ , i.e. probability that randomly chosen vertex belongs to the percolation cluster

$$\sum_{s=0}^{\infty} h_0(s) = 1 - \mathcal{P}_\infty.$$

P7.14 Find \mathcal{P}_∞ using

$$\mathcal{P}_\infty = 1 - H_0(1) = 1 - G_0(v),$$

where $v = H_1(1)$ is the solution of the functional equation $v = G_1(v)$. [20%]

The size of the percolation cluster

Probability \mathcal{P}_∞

Let us define \mathcal{P}_∞ , i.e. probability that randomly chosen vertex belongs to the percolation cluster

$$\sum_{s=0}^{\infty} h_0(s) = 1 - \mathcal{P}_\infty.$$

P7.14 Find \mathcal{P}_∞ using

$$\mathcal{P}_\infty = 1 - H_0(1) = 1 - G_0(v),$$

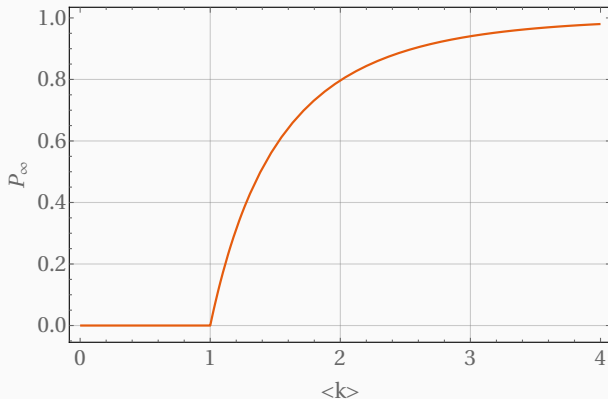
where $v = H_1(1)$ is the solution of the functional equation $v = G_1(v)$. [20%]

Solution:

$$\mathcal{P}_\infty = 1 - e^{-\langle k \rangle \mathcal{P}_\infty}$$

The size of the percolation cluster – summary

P7.15 Compare obtained result with simulation from project P7.3. [10%]



Percolation in scale-free networks

$$\mathcal{P}(k) = \frac{(\alpha - 1)m^{\alpha-1}}{k^\alpha}, \quad k = m, m + 1, \dots, k_{\max} = mN^{1/(\alpha-1)}.$$

P7.16 Compute percolation threshold for scale-free networks with power-law degree distribution: [20%]

$$\frac{\langle k^2 \rangle}{\langle k \rangle} = ?$$

Percolation in scale-free networks

$$\mathcal{P}(k) = \frac{(\alpha - 1)m^{\alpha-1}}{k^\alpha}, \quad k = m, m + 1, \dots, k_{\max} = mN^{1/(\alpha-1)}.$$

P7.16 Compute percolation threshold for scale-free networks with power-law degree distribution: [20%]

$$\frac{\langle k^2 \rangle}{\langle k \rangle} = ?$$

Solution:

$$\frac{\langle k^2 \rangle}{\langle k \rangle} = \begin{cases} k_{\max} & \text{dla } \alpha \in (1, 2) \\ m^{\alpha-2} k_{\max}^{3-\alpha} & \text{dla } \alpha \in (2, 3) \\ m & \text{dla } \alpha \in (3, \infty). \end{cases}$$

$$\frac{\langle k^2 \rangle}{\langle k \rangle} = \left| \frac{2 - \alpha}{3 - \alpha} \right| \begin{cases} k_{\max} & \text{for } \alpha \in (1, 2) \\ m^{\alpha-2} k_{\max}^{3-\alpha} & \text{for } \alpha \in (2, 3) \\ m & \text{for } \alpha \in (3, \infty). \end{cases}$$

P7.17 Compute functions G_0 and G_1 for scale-free networks. [20%]

$$\frac{\langle k^2 \rangle}{\langle k \rangle} = \left| \frac{2 - \alpha}{3 - \alpha} \right| \begin{cases} k_{\max} & \text{for } \alpha \in (1, 2) \\ m^{\alpha-2} k_{\max}^{3-\alpha} & \text{for } \alpha \in (2, 3) \\ m & \text{for } \alpha \in (3, \infty). \end{cases}$$

P7.17 Compute functions G_0 and G_1 for scale-free networks. [20%]

Solution:

$$G_0(x) = (\alpha - 1) \text{Li}_\alpha(x),$$
$$G_1(x) = (\alpha - 2) \frac{\text{Li}_{\alpha-1}(x)}{x}.$$

P7.18 Investigate with simulations [50%]

- robustness of ER graph and scale-free networks to random failures,
- robustness of ER graph and scale-free networks to intentional attacks

P7.19 Which type of graph is more robust on those two threats? How to explain it? [10%]

P7.20 **A task for the ambitious:** make the analytical computation for the robustness to random failures and intentional attacks (see sec. 5.3.2 in Fronczak and Fronczak book). [60%]

Thank you for your attention!



**European
Funds**
Knowledge Education Development

**Warsaw University
of Technology**

European Union
European Social Fund



MSc program in Data Science has been developed
as a part of task 10 of the project
„NERW PW. Science - Education - Development - Cooperation”
co-funded by European Union from European Social Fund.