Modelowanie i analiza sieci złożonych

VII. Probabilistyczne aspekty sieci złożonych.

Grzegorz Siudem

Politechnika Warszawska

Zadanie 10 pn.
„Przygotowanie i uruchomienie nowego kierunku studiów na studiach II stopnia - Inżynieria i Analiza Danych (IAD)”

realizowane jest w ramach projektu
„NERW PW. Nauka – Edukacja – Rozwój – Współpraca”
współfinansowanego ze środków Unii Europejskiej
w ramach Europejskiego Funduszu Społecznego

Projekt

Równanie master dla sieci BA - studium przypadku

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

$$N_{m-1}(t) = 0.$$

Ćwiczenie 1. (dla bardzo ambitnych)

Rozwiąż (asymptotycznie) równanie master.

Burza mózgów:

Jakie mamy narzędzia i pomysły?

Równanie master dla sieci BA - *taśma dla amatorów*

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

Spoiler alert!

Równanie master dla sieci BA - *taśma dla amatorów*

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

$$N_{m-1}(t) = 0.$$

Ćwiczenie 2. – wskazówka 1

Zacznijmy od N_m

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

gdzie (w naszym przypadku) $c = 1$, $b(t) = m/2$.

Lemat 1.

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

gdzie $b(t) \to b$.

Równanie master dla sieci BA - *taśma dla amatorów*

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

Ćwiczenie 2. – wskazówka 2

Rozważmy teraz N_k dla $k > m$

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

gdzie $g(t) \to g$, $b(t) \to b$, a w naszym przypadku
$g(t) = (k-1)N_{k-1}(t)/2t$ (dlaczego to zbiega?), $b(t) = k/2$.

Lemat 2.

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

Równanie master dla sieci BA - *taśma dla amatorów*

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

$$N_{m-1}(t) = 0.$$

Ćwiczenie 2. – wskazówka 3

Poszukiwana wielkość to

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

zgodnie z Lematem 1 (dlaczego?) mamy:

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

a zgodnie z Lematem 2 (dlaczego?):

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

Równanie master dla sieci BA

P7.1 Do wykresów z poprzednich zajęć dodaj rozwiązanie równania master. [10%]

P7.2 Przetestuj, które rozwiązanie lepiej dopasowuje sie do rozkładu empirycznego. [20%]

Perkolacja w grafach ER

Wprowadźmy rozkład $\mathcal{Q}(k)$

$\mathcal{Q}(k)$ to rozklad stopni na końcach losowo wybranej krawędzi.

Perkolacja w grafach ER

Wprowadźmy rozkład $\mathcal{Q}(k)$

$\mathcal{Q}(k)$ to rozklad stopni na końcach losowo wybranej krawędzi.

Ćwiczenie 3.

Uzasadnij, że w grafach nieskorelowanych $\mathcal{Q}(k) \propto k\mathcal{P}(k)$, a zatem

$$\mathcal{Q}(k) = \frac{k}{\langle k \rangle}\mathcal{P}(k).$$

Wprowadźmy rozkład $\mathcal{Q}(k)$

$\mathcal{Q}(k)$ to rozklad stopni na końcach losowo wybranej krawędzi.

Ćwiczenie 3.

Uzasadnij, że w grafach nieskorelowanych $\mathcal{Q}(k) \propto k\mathcal{P}(k)$, a zatem

$$\mathcal{Q}(k) = \frac{k}{\langle k \rangle}\mathcal{P}(k).$$

Ćwiczenie 4.

Uzasadnij, że próg perkolacji można zdefiniować jako

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

co jest równoważne

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

Perkolacja w grafach ER

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

P7.3 Udowodnij, ze powyższy warunek dla grafów ER sprowadza się do $\langle k \rangle = 1$, co prowadzi do $p_c = \frac{1}{N}$. [20%]

P7.3 Sprawdź symulacyjnie powyższy wynik rysując wykres rozmiaru największego klastra w funkcji $\langle k \rangle = pN$. [20%]

Formalizm funkcji generujących

Aby wyznaczyć rozmiar największego klastra analitycznie posłużymy się formalizmem funkcji generujących:

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

Ćwiczenie 5.

Udowodnij, że

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

Funkcje generujące dla grafów ER

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

P7.4 Wyznacz funkcję G_0 dla grafów ER, przyjmując, że $\mathcal{P}(k)$ ma rozkład Poissona. [20%]

Funkcje generujące dla grafów ER

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

P7.4 Wyznacz funkcję G_0 dla grafów ER, przyjmując, że $\mathcal{P}(k)$ ma rozkład Poissona. [20%]

Rozwiązanie:

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

Funkcje generujące dla grafów ER

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

P7.5 Wyznacz funkcję G_1 dla grafów ER, przyjmując, że $\mathcal{P}(k)$ ma rozkład Poissona. [20%]

Funkcje generujące dla grafów ER

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

P7.5 Wyznacz funkcję G_1 dla grafów ER, przyjmując, że $\mathcal{P}(k)$ ma rozkład Poissona. [20%]

Rozwiązanie:

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

Funkcje generujące

Zdefiniujmy funkcję H_1

$H_1(x)$ jest funkcją generującą rozkładu prawdopodobieństwa $h_1(s)$, że idąc w dowolnym kierunku losowo wybranej krawędzi rozważanego grafu (dowolnego, nie koniecznie ER!) w skończonej liczbie kroków możemy dojść do s wierzchołków.

Funkcje generujące

Zdefiniujmy funkcję H_1

$H_1(x)$ jest funkcją generującą rozkładu prawdopodobieństwa $h_1(s)$, że idąc w dowolnym kierunku losowo wybranej krawędzi rozważanego grafu (dowolnego, nie koniecznie ER!) w skończonej liczbie kroków możemy dojść do s wierzchołków.

Ćwiczenie 6.

Uzasadnij dlaczego

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

gdzie $h_k(s-1)$ to prawdopodobieństwo, że do $s-1$ węzłów można dotrzeć idąc wzdłuż k różnych krawędzi.

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

P7.6 Udowodnij, że gdy

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

oraz

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

to wówczas [30%]

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

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

Ćwiczenie 7.

Oblicz średni rozmiar klastra dany (dlaczego?) wzorem

$$\langle s \rangle = H_0'(1) = 1 + G_0'(1)H_1'(1)$$

.

Średni rozmiar klastra przed perkolacją

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

Ćwiczenie 7.

Oblicz średni rozmiar klastra dany (dlaczego?) wzorem

$$\langle s \rangle = H_0'(1) = 1 + G_0'(1)H_1'(1)$$

.

Rozwiązanie:

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

.

Prawdopodobieństwo $\mathcal{P}_\infty$

Zdefiniujmy prawdopodobieństwo $\mathcal{P}_\infty$, że losowo wybrany węzeł należy do klastra perkolacyjnego

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

Cwiczenie 8.

Wyznacz wartość $\mathcal{P}_\infty$ korzystając z zależności

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

gdzie $v = H_1(1)$ jest rozwiązaniem równania $v = G_1(v)$.

Rozmiar klastra perkolacyjnego

Prawdopodobieństwo $\mathcal{P}_\infty$

Zdefiniujmy prawdopodobieństwo $\mathcal{P}_\infty$, że losowo wybrany węzeł należy do klastra perkolacyjnego

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

Cwiczenie 8.

Wyznacz wartość $\mathcal{P}_\infty$ korzystając z zależności

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

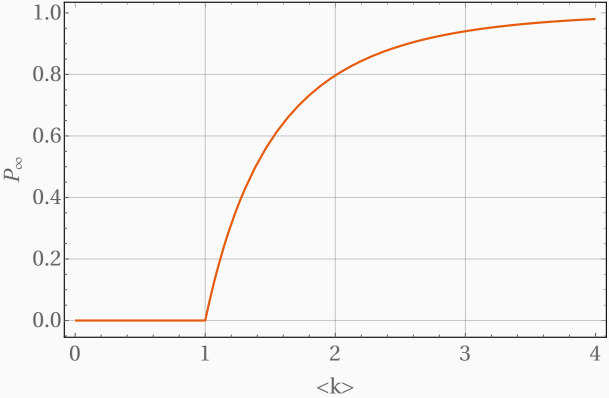
gdzie $v = H_1(1)$ jest rozwiązaniem równania $v = G_1(v)$.

Rozwiązanie:

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

Rozmiar klastra perkolacyjnego – podsumowanie

P7.7 Porównaj uzyskany wynik z symulacjami z zadania P7.3. [10%]



Perkolacje w sieciach bezskalowych

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

P7.8 Oblicz próg perkolacji dla sieci o potęgowym rozkładzie stopni wierzchołków: [20%]

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

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

P7.8 Oblicz próg perkolacji dla sieci o potęgowym rozkładzie stopni wierzchołków: [20%]

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

Rozwiązanie:

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

Perkolacje w sieciach bezskalowych

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

Ćwiczenie 9.

Jakie wnioski wyciągamy z powyższego o progu perkolacji?

Uwaga!

Ćwiczenia 1-9 warte są 40% punktów za projekt.

P7.9 Wyznacz funkcje G_0 i G_1 dla sieci bezskalowych. [20%]

Perkolacje w sieciach bezskalowych

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

Ćwiczenie 9.

Jakie wnioski wyciągamy z powyższego o progu perkolacji?

Uwaga!

Ćwiczenia 1-9 warte są 40% punktów za projekt.

P7.9 Wyznacz funkcje G_0 i G_1 dla sieci bezskalowych. [20%]

Rozwiązanie:

$$G_0(x) = (\alpha - 1)\mathrm{Li}_\alpha(x),$$

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

Projekt

P7.10 Zbadaj symulacyjnie [50%]

- odporność grafów ER i sieci beskalowych na przypadkowe awarie,
- odporność grafów ER i sieci beskalowych na intencjonalne ataki.

P7.11 Który z typów sieci jest bardziej odporny na które z zagrożeń? Jak to wytłumaczyć? [10%]

P7.12 **Zadanie dla ambitnych:** wykonaj rachunki analityczne dla opisu odporności na ataki (por. 5.3.2 w książce Fronczaków). [60%]

Dziękuję za uwagę!

Fundusze Europejskie
Wiedza Edukacja Rozwój

Politechnika Warszawska

Unia Europejska
Europejski Fundusz Społeczny

Zadanie 10 pn.
„Przygotowanie i uruchomienie nowego kierunku studiów na studiach II stopnia
- Inżynieria i Analiza Danych (IAD)”

realizowane jest w ramach projektu
„NERW PW. Nauka – Edukacja – Rozwój – Współpraca”
współfinansowanego ze środków Unii Europejskiej
w ramach Europejskiego Funduszu Społecznego