



Agata Fronczak

Wykładnicze grafy przypadkowe: teoria i przykłady zastosowań do analizy rzeczywistych sieci złożonych

CZĘŚĆ 1.

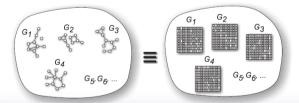
Wykładnicze grafy przypadkowe - trochę teorii



Zespoły statystyczne sieci przypadkowych

 $\textbf{Sie\acute{c}\ przypadkowa} \rightarrow \text{pojedyncza\ sie\acute{c}\ taka,\ } \\ \dot{\textbf{z}} \\ \textbf{e} \\ \textbf{w} \\ \textbf{jej} \\ \textbf{procedurze\ konstrukcyjnej\ jest\ pewien\ element\ losowości;} \\ \textbf{r} \\ \textbf{v} \\ \textbf{v} \\ \textbf{v} \\ \textbf{jej} \\ \textbf{procedurze\ konstrukcyjnej\ jest\ pewien\ element\ losowości;} \\ \textbf{v} \\ \textbf{$

 $\textbf{Sie\acute{c}\ przypadkowa} \rightarrow \text{element}\ \underline{\text{zespołu}\ \text{statystycznego\ sieci}}; \textbf{Zesp\acute{o}l\ statystyczny}\ \text{to\ zbi\acute{o}r\ mo\acute{z}liwych\ realizacji\ sieci\ } \{G\},$ którym to realizacjom (tj. mikrostanom) są przypisane odpowiednie prawdopodobieństwa P(G);



Wykładnicze grafy przypadkowe: sieci o zadanym hamiltonianie strukturalnym

$$P(G) = \frac{e^{\overline{H}(G)}}{Z},$$

$$Z = \sum_{(G)} \overline{e}^{H(G)}$$

$$H(G) = \sum_{r} \theta_r m_r(G)$$

Besag (1974), Holland & Leinhardt (1981), Frank & Strauss (1986), Park & Newman (2004), .

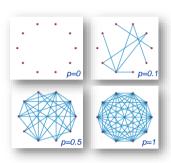


Przykład: Zespół klasycznych grafów przypadkowych

Procedura konstrukcyjna:

- Ustalona liczba węzłów N
- Każdą parę węzłów łączymy z prawdopodobieństwem p.





prawdopodobieństwo grafu w zespole:

E(G) – liczba krawędzi w grafie

$$\begin{split} P(G) &= p^{E(G)}(1-p)^{\binom{N}{2}-E(G)} \\ &= \left(\frac{p}{1-p}\right)^{E(G)}(1-p)^{\binom{N}{2}} \\ &= \underbrace{e^{\int \mathcal{U}(\underbrace{\mathcal{V}_{\ell-p}})\cdot E(G)}}_{(\ell-p)^{-\binom{N}{2}}} = \frac{\overline{e}^{H(G)}}{Z} \end{split}$$

hamiltonian zespołu klasycznych grafów przypadkowych:

$$H(G) = \ln \left[\frac{1 - p}{\log m} \right] E(G) = \theta E(G)$$

suma statystyczna:

$$Z = \frac{1}{(1-p)^{\binom{N}{2}}} = (1 + \overline{e}^{\theta})^{\binom{N}{2}}$$

Park & Newman (2004)

5

Zespół kanoniczny



Układ, który może wymieniać energię z otoczeniem o stałej temperaturze, T=const

Przestrzeń stanów układu: $\{\varOmega\}$

Prawdopodobieństwo, że układ będzie przebywał w mikrostanie \varOmega

$$Ω$$

$$P(Ω) = \frac{e^{-βE(Ω)}}{Z}$$

Średnia energia układu

$$\langle E \rangle = -\frac{\partial \ln Z}{\partial \beta}$$

Wykładnicze grafy przypadkowe: na przykładzie klasycznego grafu przypadkowego



Graf, który wymienia krawędzie z rezerwuarem krawędzi, tj. grafem o stałej gęstości połączeń p=const

Przestrzeń stanów, tj. zbiór wszystkich grafów prostych o ustalonej liczbie węzłów: $\{G\}$

Prawdopodobieństwo, że układ będzie w stanie ${\cal G}$

$$P(G) = \frac{\vec{e}^{\theta E(G)}}{Z}, \quad Z = \binom{N}{2} \ln \left[1 + \vec{e}^{\theta} \right], \quad \theta = 0$$

Średnia liczba krawędzi w grafie

$$\langle E \rangle = \frac{\partial \ln Z}{\partial \theta} = \binom{N}{2} p$$

Gdy potrafimy napisać hamiltonian układu, to potem własności tego układu (np. średnią energię układu – $\langle E \rangle$) można uzyskać wykonując symulacje MC (patrz wykład z *KMS, MN*: algorytm Metropolis).

Wcale nie musimy rozwiązywać układu w sposób ścisły, nie musimy znać sumy statystycznej ...



Symulacje numeryczne metodą Monte Carlo

Symulacje MC umożliwiają numeryczne badanie układów fizycznych w stanie równowagi termodynamicznej.

Podstawową rolę w symulacjach MC odgrywa generowanie (losowanie) ciągu mikrostanów (realizacji) badanego układu (np. grafu)

$$\{G_i\} = G_1, G_2, G_3, \dots$$

W tym ciągu poszczególne mikrostany układu, G_i , muszą mieć określone prawdopodobieństwo realizacji,

$$P(G_i) = \frac{\overline{e}^{H(G_i)}}{Z}$$

Symulacje numeryczne polegające na bezpośrednim losowaniu ciągu $\{G_i\}$ i akceptacji poszczególnych realizacji G_i układu z prawdopodobieństwem $P(G_i)$ jest bardzo nieefektywne, ponieważ z reguły $P(G_i) \ll 1$.

W symulacjach metodą MC wspomniany ciąg mikrostanów G_1,G_2,\dots generuje się wykorzystując ergodyczny proces Markowa, w którym prawdopodobieństwo $p(i\to j)$ wygenerowania stanu G_j ze stanu G_i jest dobrane w taki sposób, by $P(G_j)$ miało określoną postać.

Newman & Barkema Monte Carlomethods in statistical physics Oxford University Press (1999)

Ponieważ badany układ jest w stanie równowagi musi być

$$\sum_{j} P(G_i) \; p(i \rightarrow j) = \sum_{j} P(G_j) \; p(j \rightarrow i)$$

 \mathbf{Wg} tej zależności: \mathbf{W} stanie równowagi częstość przejść ze stanu G_i do dowolnego innego stanu, np. G_j , musi być równa częstości przejść dowolnego stanu G_j do G_i .

Warunek równowagi szczegółowej:

$$P(G_i) p(i \rightarrow j) = P(G_j) p(j \rightarrow i)$$

$$\frac{p(i \to j)}{p(j \to i)} = \frac{P(G_j)}{P(G_i)} = \overline{e}^{\Delta H_j} = \mathcal{C}^{\Delta H_j}$$

gdzie
$$\Delta H_{ij} = H(G_i) - H(G_i) = - \Delta H_{ji}$$

jest zmianą hamiltonianu układu, jaka nastąpiła w wyniku zmiany stanu $i \to j$ (tj. $G_i \to G_j$)

A. Fronczak, Wykładnicze grafy przypadkowe: teoria. przykłady, symulacje numeryczne (str. 16-19) http://www.if.pw.edu.pl/~agatka/pub/2014rozdzERG.pdf

7



Algorytm Metropolis

Warunek równowagi szczegółowej:

$$\frac{p(i \to j)}{p(i \to i)} = e^{\Delta H_{ij}} = e^{\Delta H_{ji}}$$

Algorytm Metropolis realizuje jedną z najprostszych postaci prawdopodobieństwa zmiany stanów $p(i \rightarrow j)$:

gdy
$$\Delta H_i < 0$$
 wheely $p(i
ightarrow j) = 1$



Algorytm Metropolis

Warunek równowagi szczegółowej:

$$\frac{p(i \to j)}{p(j \to i)} = \bar{e}^{\Delta H} \dot{j} = e^{\Delta H} \dot{j} \iota \quad ()$$

Algorytm Metropolis realizuje jedną z najprostszych postaci prawdopodobieństwa zmiany stanów $p(i \to j)$:

gdy
$$\Delta H_{i} < 0$$
 wheely $p(i \rightarrow j) = 1$

gdy
$$\Delta H>0$$
 wholy $p(i o j)=\overline{e}^{\Delta h}$

- 2) 200 Dany, re H(G) < H(Gi) => AHij <0
- 3) Hg. algoryamu Hetropolis powinuishy takie pnejsie raakoeptovać z prontologo. p(i-j)=1.

4) up algorytmu prejsie precime gdy zaistnieje po-nele jago 1856. zahorenie partnieju zeabeptova z prevalgo. $P(j\rightarrow i) = e^{-5H_j i} = e^{5H_j i}$

5) Dieloc obydue wytotenie dostępy

(*): $\frac{p(i \rightarrow i)}{p(j \rightarrow i)} = \frac{1}{e^{OH}i} = e^{-\Delta Hij}$

9



Klasyczne grafy przypadkowe

Przykład wykorzystania algorytm Metropolis

Klasyczne grafy przypadkowe

 $H(G) = \theta E(G),$

E(G) liczba krawędzi w grafie

 $\theta = \ln \left[\frac{\text{N-}p}{\text{Not }p} \right]$

p prawdopodobieństwo krawędzi

Niech zmiana stanów $i\to j$ polega na dodawaniu lub usuwaniu krawędzi w grafie G_i oraz akceptowaniu lub odrzucaniu zmian.

 $\Delta H = H(G_j) - H(G_i) = \pm \theta$

- $+\theta$ oznacza dodanie krawędzi
- $-\theta$ oznacza usunięcie krawędzi

Dalej rozważamy sytuację, gdy heta>0, czyli gdy $p<\frac{1}{2}$

Ponieważ $\theta>0$ wynika stąd, że:

- 1. dodawanie krawędzi akceptujemy z prawdopodobieństwem: $p(i \rightarrow j) = e^{-\theta}$
- 2. usuniecie krawedzi akceptujemy zawsze:

 $p(i \rightarrow j) = 1$

- Symulację MC rozpoczynamy od dowolnego grafu prostego, np. od grafu pustego (tj. od pustej macierzy sąsiedztwa).
- ${\bf 2.}\;\;$ Losujemy dowolny element macierzy sąsiedztwa $A_{ij}.$

Jeśli $A_{ij} = 1$ wtedy zamieniamy: $A_{ij} = 0$

Jeśli $A_{ij} = 0$ wtedy z prawd. $e^{-\theta}$ zamieniamy: $A_{ij} = 1$

3. Sprawdzamy, jak w czasie zmienia się liczba krawędzi $\boldsymbol{E}(t).$

Gdy $\langle E \rangle = const$ oznacza to, że kolejne grafy w ciągu

 $\{\ldots,G_t,G_{t+1},G_{t+2},\ldots\}$

pojawiają się z wykładniczą częstością:

 $P(G) = \frac{\overline{e}^{H(G)}}{Z}$



Sformułowanie problemu ...

Problem: Przypuśćmy, że badamy pewną sieć rzeczywistą i chcemy stworzyć model tej sieci

Krok 1.

Ustalamy te spośród strukturalnych cech badanej sieci, które są dla nas ważne, np.

- liczba krawędzi w sieci, F.
- liczba trójkatów, T.
- sekwencja stopni węzłów: $\{k_i\} = k_1, k_2, \dots, k_N$.
- ogólnie, pewien zbiór parametrów: $\{x_i\}=x_1,x_2,\ldots,x_r$

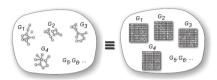
Na podstawie sieci rzeczywistej, określamy średnie wartości tych parametrów:

- · (...)
- $\bullet \quad \textit{og\'olnie:} \ \{x_i^*\} = x_1^*, x_2^*, \dots, x_r^*.$

Krok 3.

Ustalamy zbiór możliwych realizacji badanej sieci $\{G\}$:

- zbiór wszystkich grafów prostych o zadanej liczbie węzłów
- zbiór grafów ważonych o wagach z ustalonego zbioru



Krok 4.

Wybieramy rozkład P(G) taki, by typowe realizacje zespołu sieci miały określone cechy sieci rzeczywistej.

$$\langle x_i \rangle = \sum_G x_i(G)P(G) = x_i^*$$

Pytanie: W jaki sposób wybrać P(G)?

11



Zasada maksymalnej entropii

M.E.J. Newman Networks: An Introduction Oxford University Press, 2010

"The best choice of the probability distribution P(G)is the one that maximizes the Shannon/Gibbs entropy. (...) the maximum entropy choice is best in the sense that it makes the minimum assumptions about the distribution other than those imposed upon us by the constraints.'

Entropia Shannona

$$S = -\sum_G P(G) \ln P(G)$$

Więzy narzucone na zespół grafów:

$$\sum_{G} P(G) = 1 \hspace{1cm} \forall_{i} \hspace{0.2cm} \langle x_{i} \rangle = \sum_{G} x_{i}(G) P(G) = x_{i}^{*}$$

Metoda mnożników Lagrange'a:

$$\frac{\partial}{\partial P(G)} \left[S - \alpha \left(1 - \sum_{G \in \mathcal{G}} P(G) \right) - \sum_{i=1}^r \theta_i \left(x_i^* - \sum_{G \in \mathcal{G}} x_i(G) P(G) \right) \right] = 0$$

Dostajemy:

$$P(G) = \exp\left[\alpha - 1 + \sum_{i=1}^{r} \theta_i x_i(G)\right] \qquad P(G) = \frac{e^{H(G)}}{Z},$$

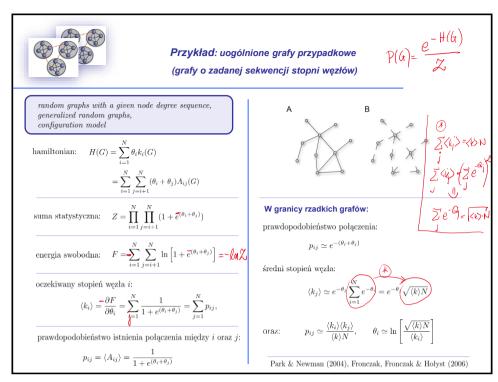
gdzie

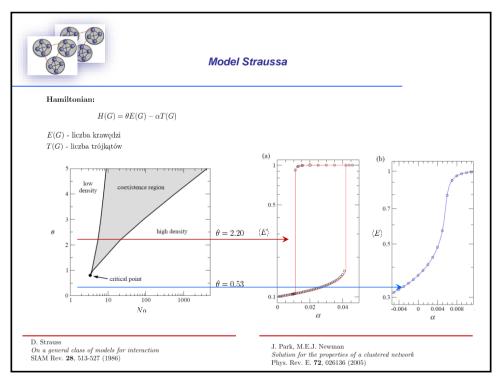
hamiltonian:
$$H(G) = \sum_{i=1}^{r} \theta_i x_i(G)$$

suma statystyczna: $Z = e^{(1-\alpha)}$

energia swobodna: $F = \ln Z$

oczekiwana wartość parametru $x_i \colon \ \langle x_i \rangle = \frac{-1}{Z} \frac{\partial Z}{\partial \theta_i} = \frac{\partial F}{\partial \theta_i}$







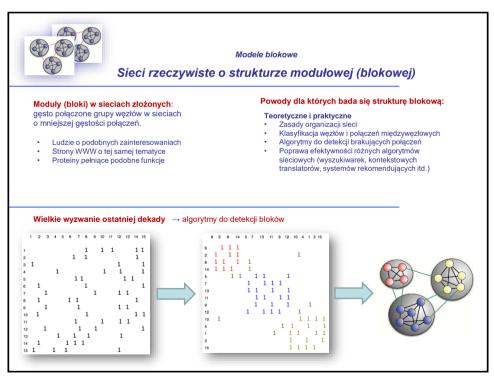
Agata Fronczak

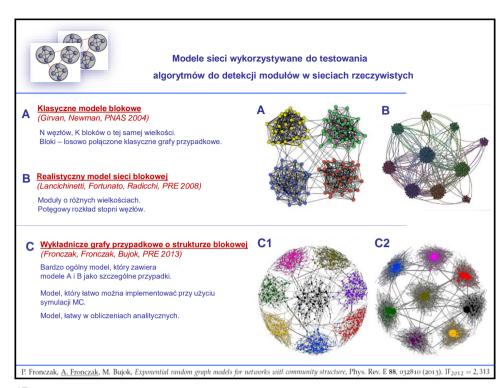
Wykładnicze grafy przypadkowe: teoria i przykłady zastosowań do analizy rzeczywistych sieci złożonych

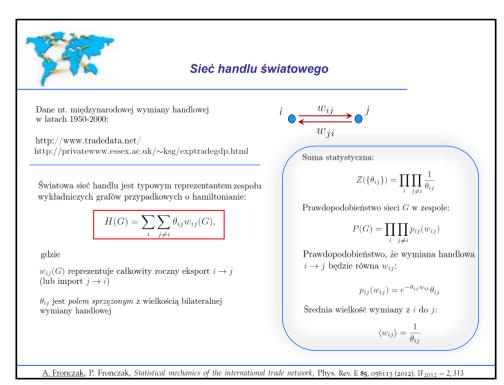
CZĘŚĆ 2.

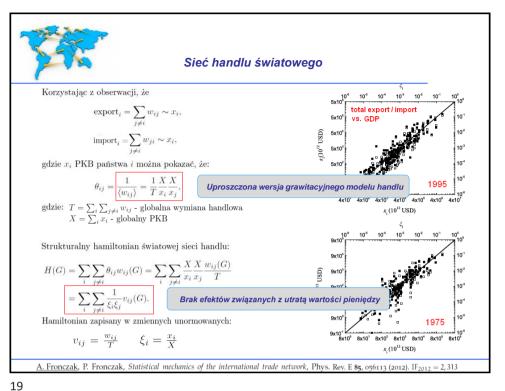
Wykładnicze grafy przypadkowe – przykłady zastosowań

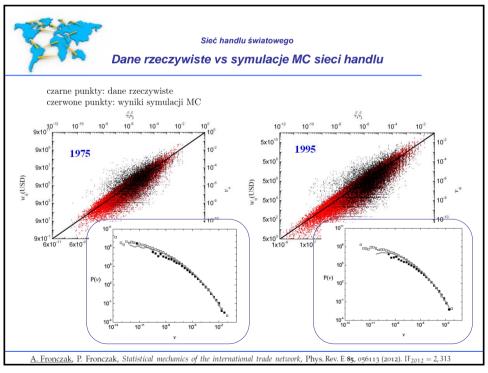
15













Sieć handlu światowego

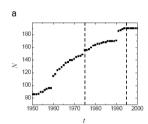
Zmienność sieci w czasie, proces quasistatyczny

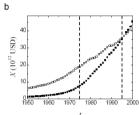
W każdym roku z okresu 1950-2000 badana sieć jest dobrze opisana hamiltonianem strukturalnym:

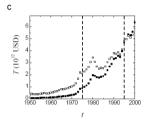
$$H(G) = \sum_i \sum_{j \neq i} \frac{1}{\xi_i \xi_j} v_{ij}(G).$$

Oznacza to, że ewolucję tej sieci w czasie można traktować jako **proces quasistatyczny!**









A. Fronczak, P. Fronczak, Statistical mechanics of the international trade network, Phys. Rev. E 85, 056113 (2012). IF₂₀₁₂ = 2,313

21



Sieć handlu światowego

Relacja fluktuacyjno – dysypacyjna

W każdym roku z okresu 1950-2000 badana sieć jest dobrze opisana hamiltonianem strukturalnym:

$$H(G) = \sum_i \sum_{j \neq i} \frac{1}{\xi_i \xi_j} v_{ij}(G).$$

Oznacza to, że ewolucję tej sieci w czasie można traktować jako ${f proces}$ ${f quasistatyczny!}$





Relacja fluktuacyjno-dysypacyjna (tj. proste twierdzenie fluktuacje-odpowiedź)

$$\langle v_{ij}^2 \rangle - \langle v_{ij} \rangle^2 = -\frac{\partial \langle v_{ij} \rangle}{\partial \theta_{ij}^*},$$





$$\frac{d\langle v_{ij}\rangle}{\langle v_{ij}\rangle} = \frac{d\xi_i}{\xi_i} + \frac{d\xi_j}{\xi_j}$$

 prawa strona: fluktuacje eksportu v_{ij} , lewa strona: podatność eksportu na zmianę zewnętrznego pola θ_{ij}^*

Względne zmiany bilateralnej wymiany handlowej można oszacować badając względne zmiany PKB handlujących ze sobą państw

A. Fronczak, P. Fronczak, Statistical mechanics of the international trade network, Phys. Rev. E 85, 056113 (2012). IF₂₀₁₂ = 2,313

