Generowanie sieci losowych przy użyciu modeli Albert Barabsi, Erdős-Rényi i Watts-Strogatz

2020

Model Albert-Barabasi

Sieć bezskalowa

Sieć, w której rozkład stopni wierzchołków jest równy

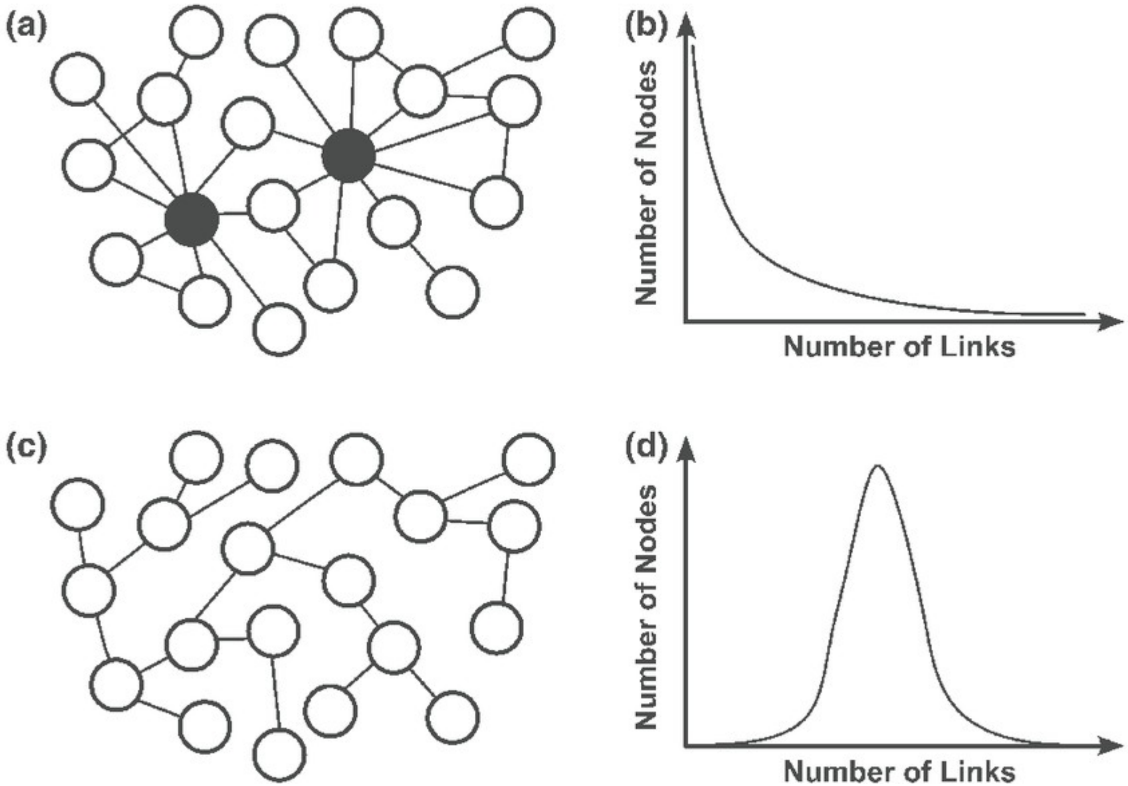
$$P(k) \sim k^{-\gamma}$$

gdzie γ jest paramtrem właściwym dla danej sieci oraz

$$2 < \gamma < 3$$

Sieć bezskalowa:

Graf losowy:



Sieć bezskalowa

Własności sieci bezskalowych:

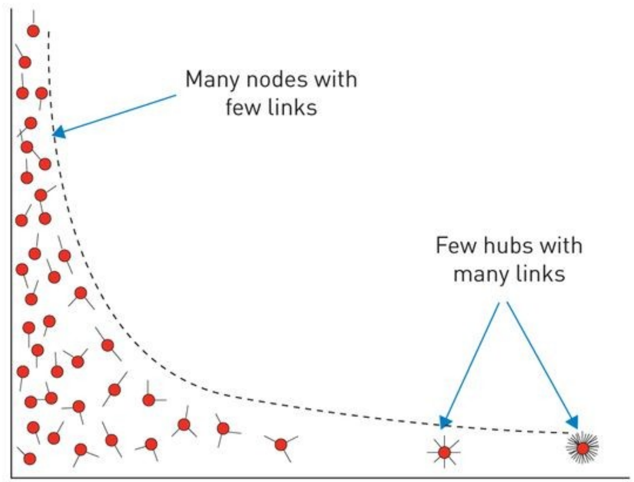
- **Wierzchołki o dużym stopniu (huby)**
- ***Rich-gets-richer***
- **Huby z małym współczynnikiem klasteryzacji, pojedyncze wierzcholki z większym**

Sieć bezskalowa - przykłady

Nazwa sieci	Wierzchołki	Krawędzie
sieci społeczne	ludzie	znajomości
mózg	neurony	synapsy
www	strony	hiperłącza
hollywood	aktorzy	filmy
środowisko naukowe	artykuły naukowe	odwołania, cytaty

Air traffic network

Scale-free



Number of links

Model Albert-Barabasi

- Najpopularniejszy model sieci bezskalowych (rok 1999)

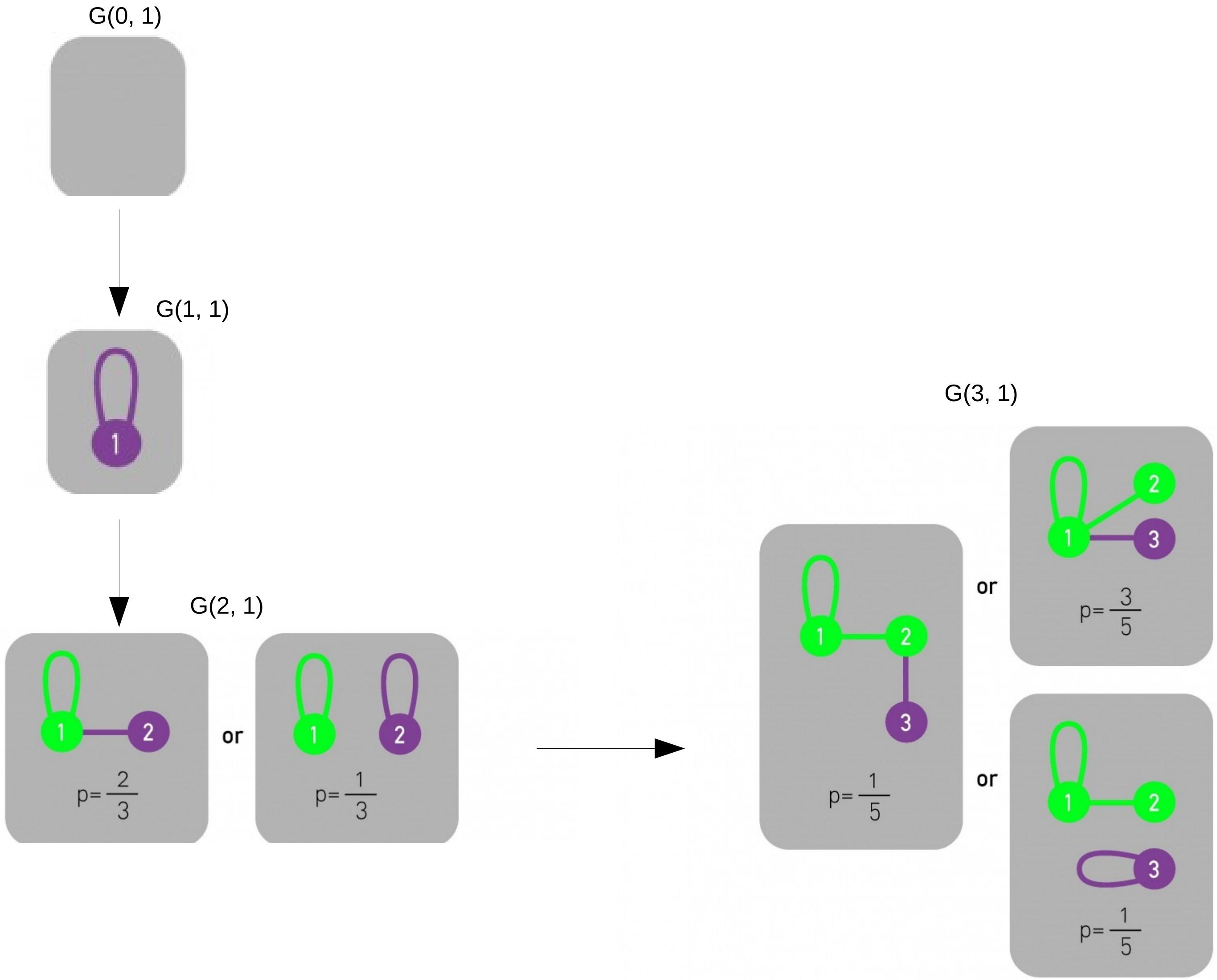
- Im więcej wierzchołek ma sąsiadów tym większa szansa że nowe wierzchołki będą z nim połączone krawędzią

Model Albert-Barabasi

Krok algorytmu:

- **Wejście: graf bezskalowy (np. graf pusty)**
- **Dodajemy wierzchołek, po czym łączymy go *m* krawędziami z obecnymi wierzcholkami**
- **Prawdopodobieństwo połączenia nowego wierzchołka z wierzchołkiem *i* wynosi:** $p_i = \frac{k_i}{\sum_j k_j}$

Model Albert-Barabasi, algorytm



Model Albert-Barabasi

Własności algorytmu:

- **Rozkład stopni:**

$$P(k) \sim k^{-3}$$

- **Średnia długość ścieżki:**

$$l = \frac{\ln(n)}{\ln\ln(n)}$$

Model Erdős-Rényi

Tą nazwą określane są dwa podobne modele służące do generowania losowych grafów:

- **G(n, M)** – graf jest wybierany losowo ze zbioru wszystkich grafów, które mają **n** wierzchołków i **M** krawędzi
  - Na przykład w modelu G(3, 2) każdy z trzech możliwych grafów z trzema wierzchołkami i dwiema krawędziami jest wybierany z prawdopodobieństwem ⅓
- **G(n, p)** – graf jest konstruowany przez losowe dodawanie krawędzi. Każda krawędź jest uwzględniana w grafie z prawdopodobieństwem **p**, niezależnie od innych krawędzi
  - Odpowiednio, wszystkie grafy z n węzłami i M krawędziami mają to samo prawdopodobieństwo wylosowania równe $p^M(1-p)^{\binom{n}{2}-M}$

Model Erdős-Rényi

Własności:

- **G(n, p) odpowiada grafowi G(n, M) z $M = \binom{n}{2}p$ przy $pn^2 \to \infty$**
- **Jeśli $np < 1$ to graf G(n, p) prawie napewno nie będzie posiadał spójnych klik rozmiaru większego niż $O(\log(n))$**
- **Jeśli $np = 1$ to graf G(n, p) prawie napewno będzie miał duże kliki rozmiaru rzędu $n^{2/3}$**
- **Jeśli $p < \frac{(1-\varepsilon) ln\, n}{n}$ to graf G(n, p) prawie napewno nie jest spójny**
- **Jeśli $p > \frac{(1+\varepsilon) ln\, n}{n}$ to graf G(n, p) prawie napewno jest spójny**

Model Erdős-Rényi

Ograniczenia:

- **Ze względu na to że prawdopodobieństwo połączenia dwóch wierzchołków jest stałe, sieć ma niski współczynnik klasteryzacji (*clustering coefficient*)**

Model Watts-Strogatz

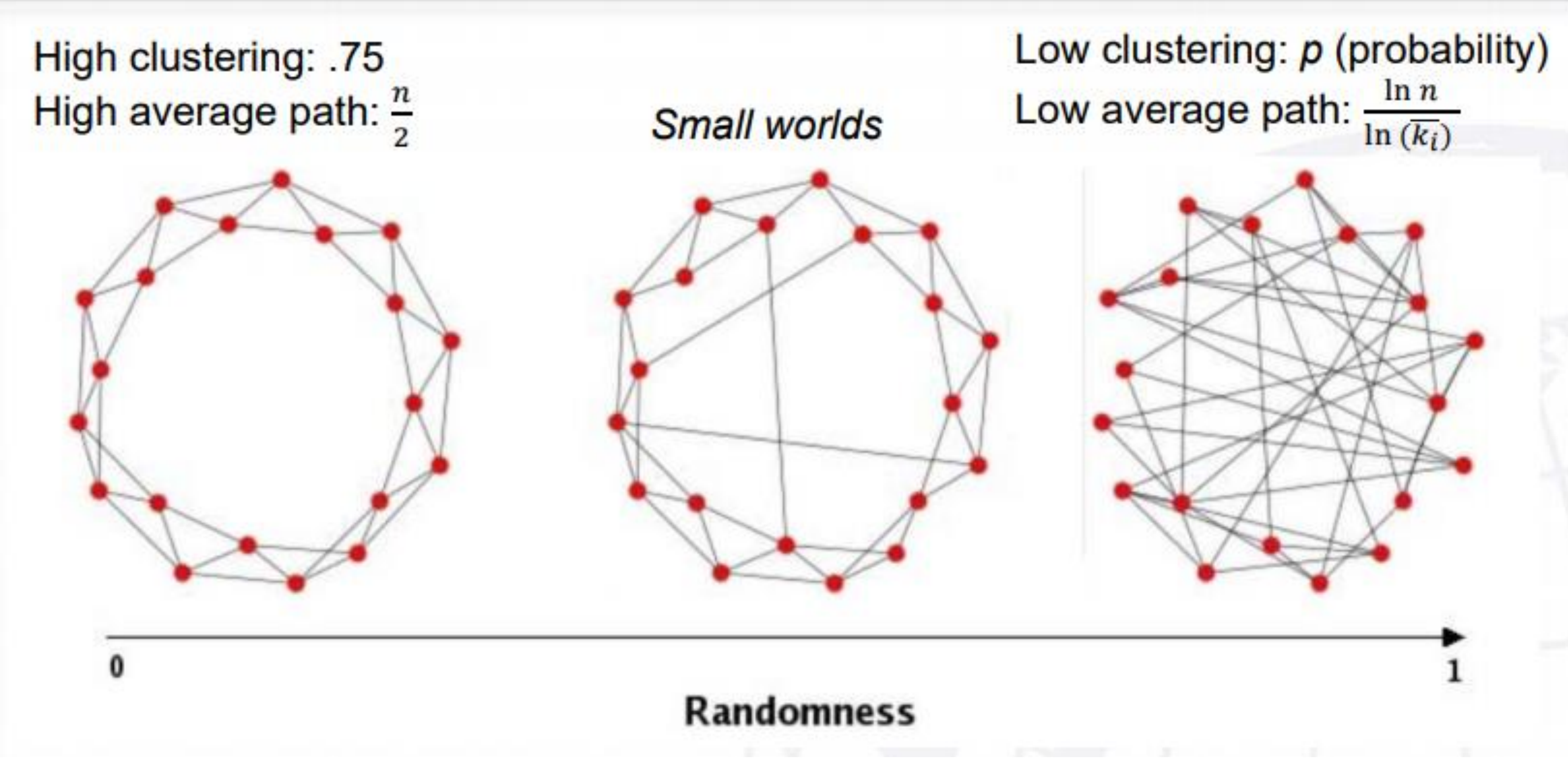
Tą nazwą określany jest model służący do generowania losowych grafów z właściwościami Small World, czyli z krótką średnią długością ścieżki i wysokim współczynnikiem klasteryzacji

Model Watts-Strogatz

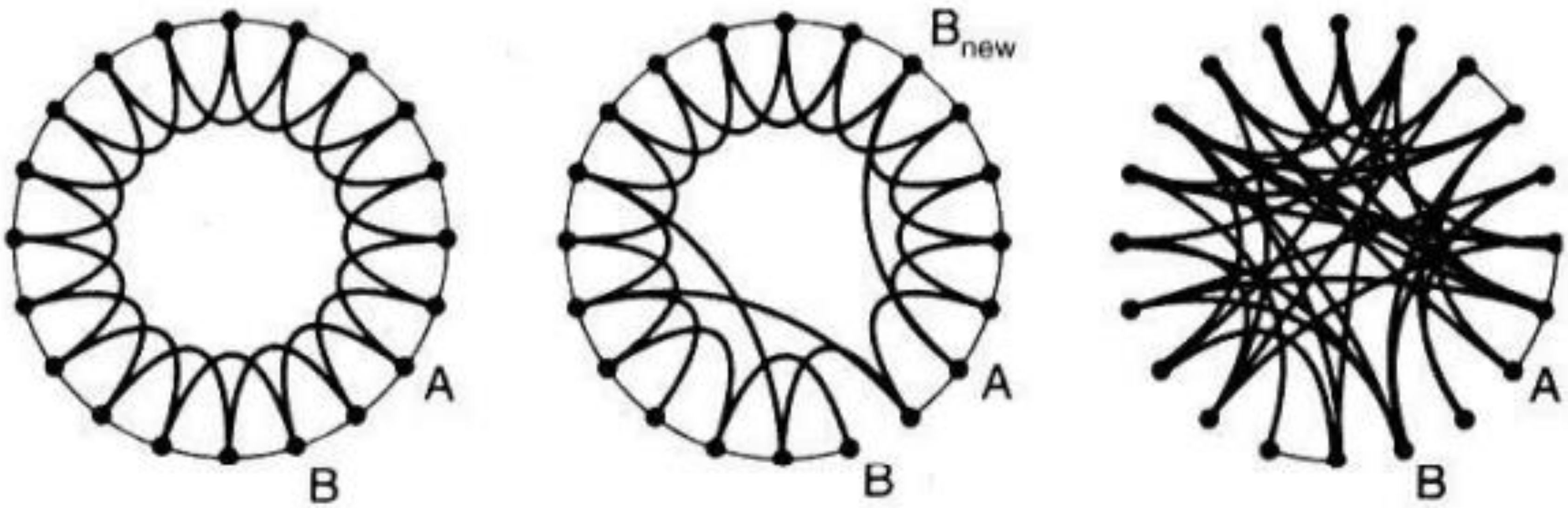
Algorytm tworzenia grafu G(n, k, p):

- **Tworzymy pierścień z n wierzchołków**
- **Każdy wierzchołek łączymy z k najbliższymi sąsiadami**
- **Tworzymy skróty w grafie: dla każdej krawędzi (u, v) w grafie z prawdopodobieństwem p zmień ją na nową krawędź (u, w) z jednostajnym rozkładem wyboru wierzchołka w**

Model Watts-Strogatz



Model Watts-Strogatz



regular lattice:
my friend's friend is always my friend

small world:
mostly structured with a few random connections

random graph:
all connections random

Model Watts-Strogatz

Właściwości:

- ❑ **Średnia długość ścieżki:**
  - ❑ **przy** $p = 0$: $\frac{n}{2k}$
  - ❑ **przy** $p \to 1$: $\frac{\ln n}{\ln k}$
- ❑ **Współczynnik klasteryzacji:**
  - ❑ **przy** $p = 0$: $\frac{3(k-2)}{4(k-1)} \to \frac{3}{4}$
  - ❑ **przy** $p \to 1$: $\frac{k}{n}$
- ❑ **Rozkład stopni wierzchołków**

Model Watts-Strogatz

Ograniczenia:

- **Może tworzyć niemożliwie duży stopień wierzchołka**
- **Ograniczanie ilości węzłów powoduje niemożliwość jej użycia dla modelowania zwiększania/zwiększających sieci**