

Uniwersytet Łódzki
Wydział Fizyki i Informatyki Stosowanej

SPRAWOZDANIE Z LABORATORIUM
Zaawansowane metody obliczeniowe

**Generowanie i analiza sieci losowych
ER, WS i BA**

Implementacja własna oraz biblioteczna

Autor: Katarzyna Stańczyk
Kierunek studiów: Informatyka
Data wykonania: 8 luty 2026
Rok akademicki: 2025/2026

Łódź, 2026

1 Wstęp

Celem ćwiczenia było zbudowanie i porównanie trzech losowych modeli sieci, często używanych jako uproszczone modele sieci społecznościowych: **Erdős-Rényiego (ER)**, **Watts-Strogatza (WS)** oraz **Barabásiego-Alberta (BA)**.

Dla jednego z modeli wykonano własną implementację algorytmu generowania, a następnie zestawiono charakterystyki uzyskanych grafów.

Dla każdej sieci wyznaczono: rozkład liczby sąsiadów (rozkład stopni), współczynnik klasteryzacji, średnicę i promień grafu. Dodatkowo zbadano zależność tych wielkości od liczby wierzchołków N oraz dla wybranych modeli - od parametrów generowania.

Ponieważ rozważane modele są losowe, pojedyncze uruchomienie może dawać wyniki odbiegające od wartości typowych. W związku z tym metryki obliczano dla wielu realizacji i raportowano średnie oraz odchylenia standardowe, co pozwala porównywać modele w sposób bardziej stabilny statystycznie [7].

1.1 Graf

1.1.1 Definicja grafu

Grafem nazywamy parę

$$G = (V, E), \quad (1)$$

gdzie V jest zbiorem wierzchołków (węzłów), a E jest zbiorem krawędzi opisujących połączenia pomiędzy wierzchołkami.

W kontekście sieci społecznościowych wierzchołki mogą reprezentować np. osoby, a krawędzie – relacje (znajomość, obserwowanie).

W dalszej części pracy analizowane są wyłącznie własności topologiczne grafu (struktura połączeń), niezależnie od interpretacji dziedzinowej [7].

1.1.2 Graf prosty, nieskierowany

W eksperymentach rozważany jest graf **nieskierowany**, tzn. krawędź $\{u, v\} \in E$ nie ma kierunku i oznacza połączenie obustronne.

Dodatkowo przyjmujemy graf **prosty**: bez pętli własnych (brak krawędzi $\{v, v\}$) oraz bez wielokrawędzi (pomiędzy dwiema parami wierzchołków występuje co najwyżej jedna krawędź). Takie założenie upraszcza interpretację stopni wierzchołków oraz obliczanie metryk, takich jak klasteryzacja czy odległości najkrótszych ścieżek.

1.1.3 Reprezentacja

Graf może być przechowywany na różne sposoby. Dwie najczęściej spotykane reprezentacje to:

- **Macierz sąsiedztwa** $A \in \{0, 1\}^{|V| \times |V|}$, gdzie

$$A_{uv} = \begin{cases} 1 & \text{gdy } \{u, v\} \in E, \\ 0 & \text{w przeciwnym razie.} \end{cases} \quad (2)$$

Reprezentacja ta jest wygodna przy operacjach macierzowych, lecz dla grafów rzadkich zajmuje $O(|V|^2)$ pamięci.

- **Lista sąsiedztwa**, czyli dla każdego wierzchołka v przechowujemy zbiór (lub listę) jego sąsiadów $N(v)$. Dla grafów rzadkich jest to zwykle efektywniejsze pamięciowo: $O(|V| + |E|)$.

W badanych sieciach, gdzie liczba krawędzi rośnie w przybliżeniu liniowo z liczbą wierzchołków, naturalnym wyborem jest reprezentacja oparta o listy sąsiedztwa, ułatwiająca m.in. wyznaczanie stopni i przegląd grafu (BFS/DFS) podczas obliczeń odległości.

1.2 Sieć

1.2.1 Sieć jako graf z interpretacją

W dalszej części pracy pojęcia *graf* i *sieć* są blisko powiązane, jednak nie mają takiego samego znaczenia.

Graf (1) jest formalnym modelem matematycznym opisującym strukturę połączeń, natomiast **sieć** to graf wraz z interpretacją znaczenia wierzchołków i krawędzi w danym problemie.

Tym samym ta sama struktura grafu może opisywać różne zjawiska, zależnie od tego, co reprezentują elementy V i E .

1.2.2 Sieć społecznościowa

Sieć społecznościowa to szczególny rodzaj sieci, w której wierzchołki reprezentują uczestników życia społecznego (osoby, konta użytkowników lub organizacje), a krawędzie opisują relacje między nimi (znajomość, współpracę, kontakt lub obserwowanie).

W analizie takich sieci często bada się cechy struktury połączeń, m.in. rozkład stopni, klasteryzację oraz odległości w grafie (średnicę i promień)[7, 1].

1.3 Stopień wierzchołka i rozkład stopni

1.3.1 Stopień wierzchołka i średni stopień

Stopień wierzchołka v (k_v) definiujemy jako liczbę jego sąsiadów, czyli liczbę krawędzi incydentnych z v . W grafie nieskierowanym można go zapisać jako

$$k_v = |N(v)|, \quad (3)$$

gdzie $N(v)$ oznacza zbiór sąsiadów wierzchołka v .

Średni stopień grafu wynosi

$$\langle k \rangle = \frac{1}{|V|} \sum_{v \in V} k_v, \quad (4)$$

a dla grafu nieskierowanego zachodzi również zależność wynikająca z faktu, że każda krawędź zwiększa sumę stopni o 2:

$$\sum_{v \in V} k_v = 2|E|. \quad (5)$$

1.3.2 Rozkład stopni $P(k)$

Rozkład stopni $P(k)$ opisuje, jakie stopnie przyjmują wierzchołki w sieci. Definiujemy go jako odsetek (prawdopodobieństwo), że losowo wybrany wierzchołek ma stopień równy k :

$$P(k) = \frac{|\{v \in V : k_v = k\}|}{|V|}. \quad (6)$$

W praktyce rozkład stopni wyznacza się przez zliczenie wartości k_v dla wszystkich wierzchołków, a następnie przedstawia w postaci histogramu lub wykresu $P(k)$.

Rozkład stopni jest jedną z kluczowych cech odróżniających modele sieci. W sieciach jednorodnych (np. o małej losowości) większość wierzchołków ma stopnie bliskie $\langle k \rangle$, natomiast w sieciach z “hubami” pojawiają się wierzchołki o stopniach znacznie większych od średniej [7, 1].

1.4 Współczynnik klasteryzacji

Klasteryzacja opisuje, na ile „zwarte” są lokalne sąsiedztwa w sieci: sprawdza, jak często sąsiedzi danego wierzchołka są połączeni między sobą [7, 8]. Dla wierzchołka v o stopniu $k_v \geq 2$ lokalny współczynnik klasteryzacji definiujemy jako

$$C_v = \frac{2T_v}{k_v(k_v - 1)}, \quad (7)$$

gdzie T_v oznacza liczbę trójkątów (cykli długości 3) zawierających wierzchołek v .

Wartość $C_v \in [0, 1]$: $C_v = 0$ oznacza brak połączeń pomiędzy sąsiadami v , a $C_v = 1$ oznacza, że wszyscy sąsiedzi v tworzą klikę.

W analizie porównawczej wygodnie jest raportować **średni współczynnik klasteryzacji**, zdefiniowany jako średnia z wartości lokalnych po wszystkich wierzchołkach:

$$\langle C \rangle = \frac{1}{|V|} \sum_{v \in V} C_v. \quad (8)$$

W dalszej części pracy współczynnik klasteryzacji wykorzystano do porównania modeli ER, WS i BA oraz analizy zależności od N i parametrów generowania sieci.

1.5 Odległości, promień i średnica

1.5.1 Odległość i najkrótsze ścieżki

Dla dwóch wierzchołków $u, v \in V$ **odległość** $d(u, v)$ definiujemy jako długość (liczbę krawędzi) najkrótszej ścieżki łączącej u z v .

$$d(u, v) = \min_{\pi: u \leadsto v} |\pi|, \quad (9)$$

gdzie π oznacza ścieżkę od u do v , a $|\pi|$ jej długość.

W grafach nieskierowanych bez wag odległości można efektywnie wyznaczać algorytmem BFS uruchomionym z wybranego wierzchołka.

1.5.2 Ekscentryczność

Ekscentryczność wierzchołka v jest miarą tego, jak daleko v znajduje się od najbardziej odległego (osiągalnego) wierzchołka w grafie:

$$e(v) = \max_{u \in V} d(u, v). \quad (10)$$

1.5.3 Promień i średnica

Na podstawie ekscentryczności definiuje się dwie globalne charakterystyki grafu: **promień** (najmniejsza ekscentryczność) oraz **średnica** (największa ekscentryczność):

$$r = \min_{v \in V} e(v), \quad (11)$$

$$D = \max_{v \in V} e(v). \quad (12)$$

Promień wskazuje istnienie „centralnych” wierzchołków (o małej maksymalnej odległości do innych), a średnica opisuje największą odległość pomiędzy parami wierzchołków w sieci [7].

1.5.4 Spójność i największa spójna składowa (LCC)

Jeśli graf nie jest spójny, to dla części par (u, v) nie istnieje ścieżka, a odległość $d(u, v)$ nie jest skończona, więc promień i średnica całego G nie są jednoznacznie określone. Dlatego w analizie obie miary liczono dla **największej spójnej składowej** (Largest Connected Component, LCC). Oznaczając przez G_{LCC} fragment grafu złożony z wierzchołków należących do tej składowej, otrzymujemy:

$$r_{\text{LCC}} = r(G_{\text{LCC}}), \quad (13)$$

$$D_{\text{LCC}} = D(G_{\text{LCC}}). \quad (14)$$

Podejście to jest szczególnie istotne w modelu ER przy małych wartościach parametru p , gdy graf może dzielić się na wiele małych składowych [3, 5].

1.6 Modele sieci

1.6.1 Model Erdős–Rényiego (ER)

W modelu Erdős–Rényiego rozważa się graf losowy $G(N, p)$, w którym dla zadanej liczby wierzchołków N każdą parę różnych wierzchołków łączy się krawędzią niezależnie z prawdopodobieństwem p [4, 5, 3]. Oczekiwana liczba krawędzi wynosi

$$\mathbb{E}[|E|] = p \binom{N}{2}, \quad (15)$$

a oczekiwany stopień wierzchołka jest równy

$$\mathbb{E}[k] = p(N - 1). \quad (16)$$

Model ER jest często traktowany jako punkt odniesienia dla sieci “bez struktury” poza losowością połączeń.

1.6.2 Model Wattsa–Strogatza (WS)

Model Wattsa–Strogatza konstruuje się w dwóch krokach. Najpierw buduje się regularną sieć na pierścieniu (*ring-lattice*): każdy wierzchołek łączy się z k najbliższymi sąsiadami (zwykle k jest parzyste), co daje graf o dużej klasteryzacji.

Następnie każdą krawędź przepina się (*rewiring*) z prawdopodobieństwem p_{rew} do losowo wybranego wierzchołka (z zachowaniem braku pętli własnych i wielokrawędzi).

W praktyce parametr p_{rew} kontroluje poziom losowości: dla $p_{\text{rew}} = 0$ otrzymujemy sieć regularną, a dla $p_{\text{rew}} \approx 1$ sieć zbliżoną do losowej [8, 7].

1.6.3 Model Barabásiego–Alberta (BA)

Model Barabásiego–Alberta opisuje sieci rosnące z mechanizmem **preferencyjnego przyłączania**. Zaczynamy od małego grafu początkowego, a następnie w każdym kroku dodajemy nowy wierzchołek i łączymy go z m istniejącymi wierzchołkami. Prawdopodobieństwo wyboru wierzchołka v jest proporcjonalne do jego stopnia:

$$\mathbb{P}(v) = \frac{k_v}{\sum_{u \in V} k_u}. \quad (17)$$

Mechanizm ten sprzyja powstawaniu wierzchołków o bardzo dużym stopniu (tzw. *hubów*), co jest typowe dla wielu sieci rzeczywistych [2, 1].

1.6.4 Parametry modeli i spodziewane efekty

Każdy z modeli posiada parametry, które w największym stopniu wpływają na obserwowane charakterystyki sieci:

- **ER:** parametr p (lub równoważnie $\langle k \rangle$ z (16)) kontroluje zagęszczenie grafu. Zwiększenie p podnosi średni stopień, zwykle zmniejsza odległości oraz zwiększa prawdopodobieństwo spójności.
- **WS:** parametr k ustala stopień sieci startowej, a p_{rew} kontroluje liczbę “skrótów” tworzonych przez przepinanie krawędzi. Dla małego p_{rew} oczekuje się dużej klasteryzacji, natomiast już umiarkowane p_{rew} może istotnie skrócić średnice/odległości (efekt *small-world*).
- **BA:** parametr m kontroluje, ile krawędzi dodaje każdy nowy wierzchołek, a więc również typowy poziom zagęszczenia sieci. Preferencyjne przyłączanie z (17) prowadzi do powstawania hubów, co zwykle obniża odległości w grafie w porównaniu do sieci jednorodnych o podobnym $\langle k \rangle$.

W części eksperymentalnej badano, jak metryki opisane w sekcjach (6)–(14) zmieniają się wraz z N oraz wraz z powyższymi parametrami modeli.

1.7 Losowość modeli i uśrednianie wyników

Wiele realizacji, średnia i odchylenie standardowe

Modele ER, WS i BA zawierają element losowy (odpowiednio: losowanie krawędzi, losowe przepinanie krawędzi oraz losowanie wierzchołków przy preferencyjnym przyłączaniu), dlatego metryki obliczone dla pojedynczej realizacji mogą istotnie różnić się pomiędzy uruchomieniami. Aby uzyskać wyniki bardziej reprezentatywne, dla każdej kombinacji parametrów i liczby wierzchołków N generowano R niezależnych realizacji grafu i obliczano interesujące wielkości (np. $\langle C \rangle$, r_{LCC} , D_{LCC}).

Dla dowolnej metryki X (np. średniej klasteryzacji) oraz wartości X_1, \dots, X_R z R realizacji wyznaczano średnią:

$$\bar{X} = \frac{1}{R} \sum_{i=1}^R X_i, \quad (18)$$

oraz odchylenie standardowe:

$$s_X = \sqrt{\frac{1}{R-1} \sum_{i=1}^R (X_i - \bar{X})^2}. \quad (19)$$

Wyniki przedstawiano w postaci wykresów zależności $\bar{X}(N)$, a rozrzut pomiędzy realizacjami ilustrowano za pomocą s_X (np. jako słupki błędów). Podobnie rozkład stopni $P(k)$ wyznaczano na podstawie stopni z wielu realizacji, co ogranicza wpływ losowych wahań typowych dla pojedynczego grafu [7].

2 Rozwinięcie

2.1 Środowisko i konfiguracja programu

Eksperymenty wykonano w języku Python w środowisku IDE **PyCharm 2024.3.4**, korzystając z interpretera **Python 3.11**.

Wykorzystano biblioteki: `NetworkX` [6], `NumPy`, `Matplotlib`, a także moduły standardowe: `os`, `csv`, `random` i `collections` (`Counter`).

Program uruchamiano jako skrypt `ZM0_StanczykK_z4.py`, który automatycznie tworzy katalogi wyjściowe i zapisuje wyniki w postaci plików graficznych oraz CSV.

W celu poprawy powtarzalności wyników ustawiono ziarna generatorów liczb losowych dla modułów `random` oraz `numpy.random`. Należy jednak podkreślić, że losowy charakter modeli oznacza, iż wyniki pojedynczej realizacji mogą różnić się między uruchomieniami, dlatego w dalszej analizie stosowano uśrednianie po wielu realizacjach zgodnie z wzorami (18), (19).

2.1.1 Struktura plików wynikowych

Wyniki zapisywane są w dwóch głównych katalogach:

- `plots/` – wykresy i wizualizacje:
 - `ER/`, `WS/`, `BA/` – wyniki dla poszczególnych modeli (graf, histogram stopni, $P(k)$ w skali log-log),
 - `compare/` – wykresy porównawcze modeli (m.in. $P(k)$, $\langle C \rangle(N)$, $D_{LCC}(N)$, $r_{LCC}(N)$),
 - dodatkowo dla WS: wykres $\langle C \rangle(p_{rew})$.
- `data/` – dane wyjściowe:
 - `ER_metrics.csv`, `WS_metrics.csv`, `BA_metrics.csv` – średnie i odchylenia standardowe metryk dla kolejnych N .

2.1.2 Parametry eksperymentu

Badano zależność metryk od liczby wierzchołków N dla wartości $\{50, 100, 200, 500, 1000\}$. Dodatkowo, w celu uzyskania czytelnych rozkładów stopni $P(k)$ oraz histogramów, użyto mniejszych rozmiarów grafu $\{10, 20, 50, 100, 200\}$, dla których wizualizacje są mniej obciążające obliczeniowo.

Dla każdej konfiguracji generowano 30 niezależnych realizacji grafu, a następnie obliczano średnią i odchylenie standardowe (wzory (18)–(19)). W przypadku średnicy i promienia stosowano podejście oparte na największej spójnej składowej (LCC), zgodnie z definicjami (13)–(14).

Parametry modeli ustawiono następująco:

- **ER:** prawdopodobieństwo krawędzi $p = 0.05$.
- **WS:** stopień sieci startowej $k = 4$ oraz prawdopodobieństwo przepinania $p_{\text{rew}} = 0.1$. Dodatkowo zbadano wpływ parametru p_{rew} dla kilku wartości $p_{\text{rew}} \in \{0.0, 0.01, 0.05, 0.1, 0.5\}$.
- **BA:** liczba dołączanych krawędzi $m = 2$.

Dobór parametrów pozwala porównać modele przy zbliżonym poziomie zagęszczenia sieci (średni stopień), a jednocześnie uwidoczyć różnice strukturalne charakterystyczne dla danego mechanizmu generowania (losowanie krawędzi, przepinanie, preferencyjne przyłączanie).

2.2 Implementacja modeli sieci

2.2.1 ER i BA (bibliotecznie)

Modele Erdős–Rényiego (ER) oraz Barabásiego–Alberta (BA) generowano z wykorzystaniem gotowych funkcji biblioteki NetworkX. Dla ER użyto wariantu $G(N, p)$, a dla BA modelu rosnącego z parametrem m (liczbą krawędzi dołączanych przez nowy wierzchołek).

2.2.2 WS (implementacja własna)

Model Wattsa–Strogatza (WS) zaimplementowano własnoręcznie w dwóch etapach:

- (1) budowa regularnej sieci na pierścieniu, w której każdy wierzchołek łączy się z k najbliższymi sąsiadami,
- (2) przepinanie krawędzi z prawdopodobieństwem p_{rew} do losowo wybranego wierzchołka.

W implementacji przyjęto założenia grafu prostego i nieskierowanego (bez pętli własnych i wielokrawędzi), a w trakcie przepinania kontrolowano, aby nowy koniec krawędzi nie był równy wierzchołkowi źródłowemu i nie tworzył duplikatu istniejącej krawędzi.

Poprawność implementacji weryfikowano przez sprawdzenie podstawowych własności: zachowanie liczby wierzchołków, brak pętli i wielokrawędzi, oraz zgodność typowego rozkładu stopni z oczekiwaniemi (Dla $p_{\text{rew}} = 0$ stopnie są skoncentrowane wokół wartości k , natomiast wraz ze wzrostem p_{rew} rozkład ulega “rozmyciu”). Dodatkowo kontrolowano, czy dla małych p_{rew} klasteryzacja pozostaje wysoka, a dla większych wartości obserwuje się skracanie odległości (efekt small-world).

2.3 Przygotowanie danych do obliczeń

2.3.1 Zakres badanych rozmiarów grafu N

W celu zbadania zależności metryk od liczby wierzchołków analizowano kilka wartości N , tak aby z jednej strony umożliwić czytelną wizualizację struktur, a z drugiej - zaobserwować trendy przy wzroście rozmiaru sieci. Dla rozkładów stopni $P(k)$ wykorzystywano dodatkowo mniejsze wartości N , ponieważ dla dużych grafów wizualizacje stają się mniej czytelne, a samo generowanie wykresów bardziej kosztowne obliczeniowo.

2.3.2 Uśrednianie po wielu realizacjach

Ponieważ każdy z modeli jest losowy, dla każdej konfiguracji parametrów i wartości N generowano wiele niezależnych realizacji grafu. Dla każdej realizacji obliczano zestaw metryk, a następnie agregowano je za pomocą średniej i odchylenia standardowego zgodnie z definicjami (18)–(19). Takie podejście zmniejsza wpływ fluktuacji losowych i pozwala porównywać modele na podstawie wartości typowych, a nie pojedynczych przypadków.

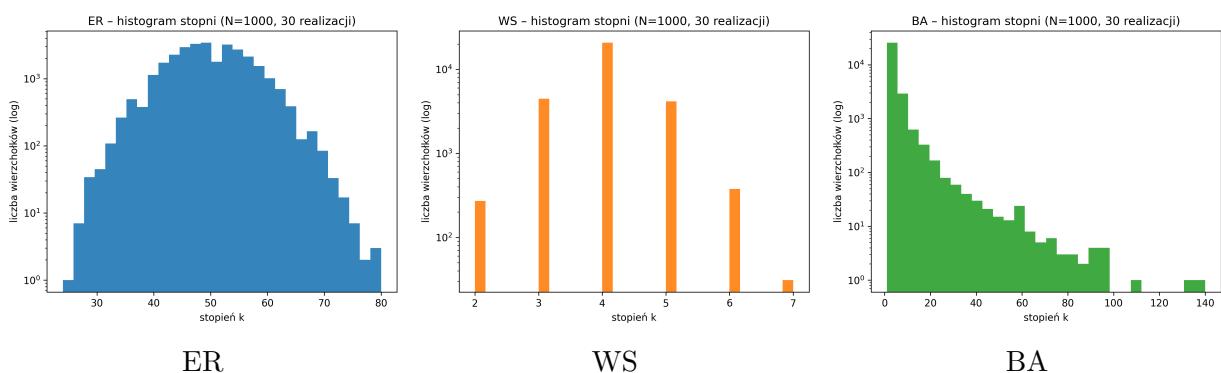
2.3.3 Spójność i liczenie średnicy/promienia

W przypadku grafów niespójnych odległości pomiędzy wierzchołkami z różnych składowych nie są skończone, dlatego promień i średnica nie są wtedy jednoznacznie określone dla całej sieci. Aby zapewnić porównywalność wyników pomiędzy realizacjami, średnicę i promień obliczano dla największej spójnej składowej (LCC), zgodnie z podejściem opisanym w (13)–(14).

2.4 Wyniki

2.4.1 Histogramy stopni

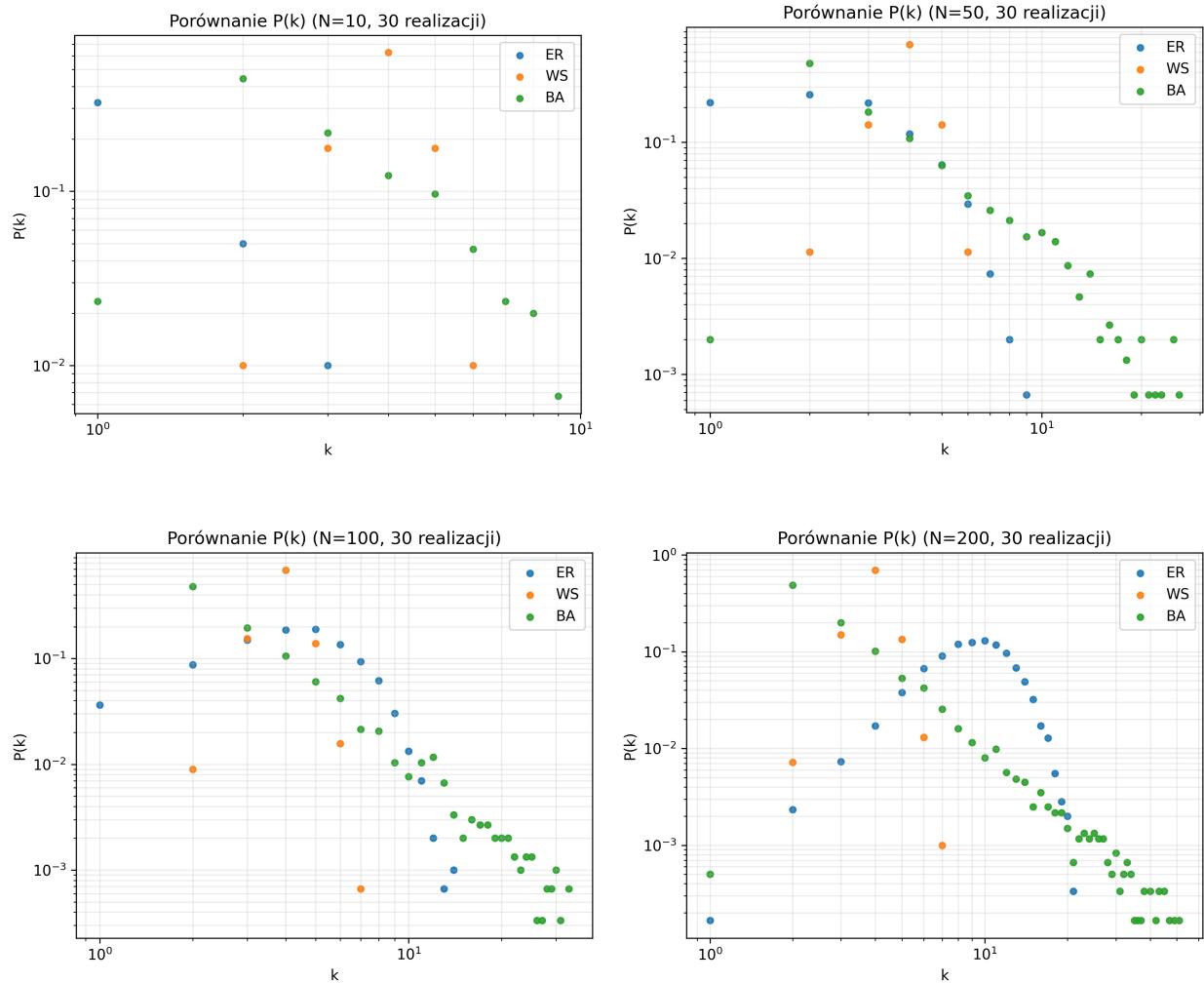
Na rys. 1 przedstawiono histogramy stopni dla większych grafów ($N = 1000$) w modelach ER, WS i BA. Dla ER i WS stopnie są silnie skoncentrowane wokół wartości typowych, natomiast dla BA rozkład jest bardziej rozciągnięty, co wynika z obecności wierzchołków o znacznie większych stopniach (hubów).



Rysunek 1: Histogramy stopni dla $N = 1000$ w modelach ER, WS i BA.

2.4.2 Rozkład stopni $P(k)$ (ER vs WS vs BA)

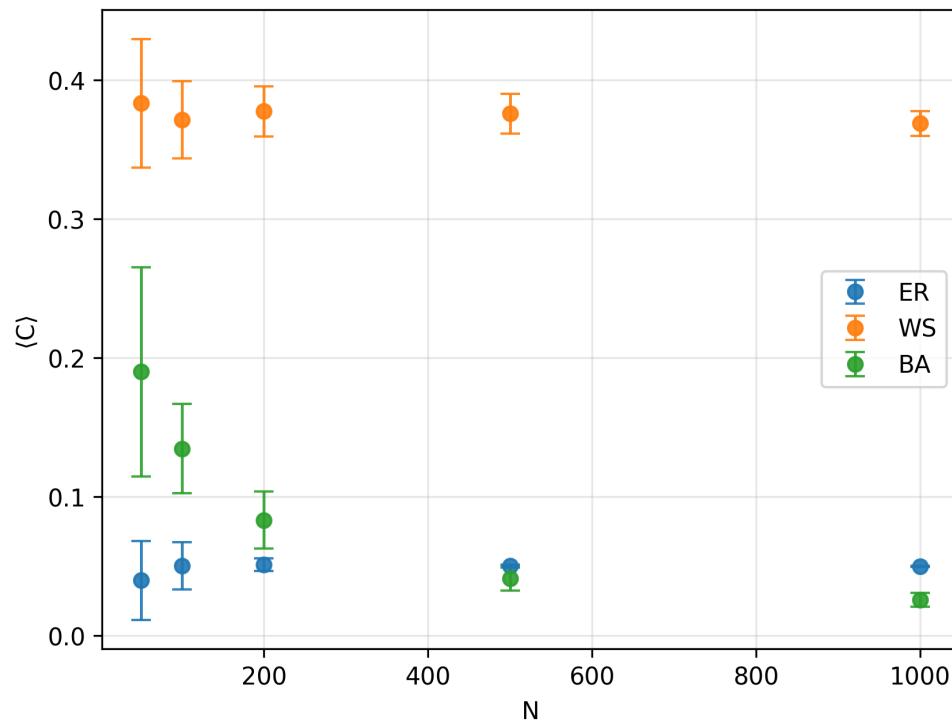
Na rys. 2 pokazano porównanie rozkładów stopni $P(k)$ w skali log–log dla trzech modeli przy tych samych rozmiarach sieci N . W modelach ER i WS rozkład jest skupiony w wąskim zakresie k , natomiast w modelu BA widoczny jest dłuższy ogon, co wskazuje na obecność hubów.



Rysunek 2: Porównanie $P(k)$ (log–log) dla modeli ER, WS i BA dla wybranych wartości N .

2.4.3 Klasteryzacja vs N

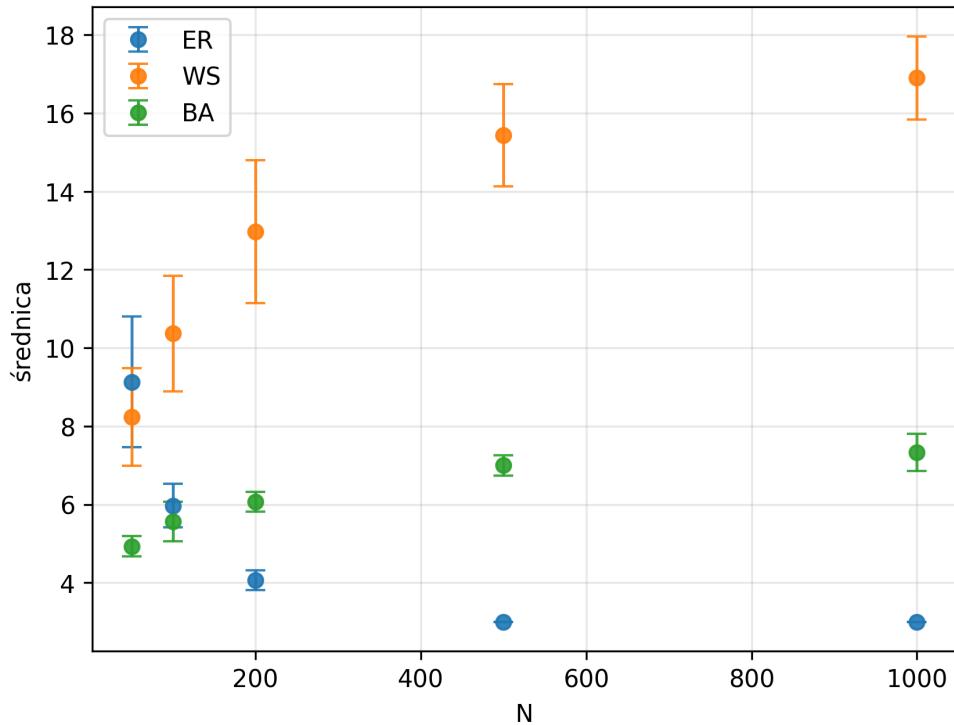
Na rys. 3 przedstawiono średni współczynnik klasteryzacji $\langle C \rangle$ w funkcji liczby wierzchołków N . Model WS charakteryzuje się istotnie większą klasteryzacją niż ER i BA (dla dobranych parametrów), co jest zgodne z intuicją: sieć startowa WS ma lokalną strukturę typu „sąsiedzi sąsiadów”, a losowe przepinanie krawędzi nie usuwa jej całkowicie.



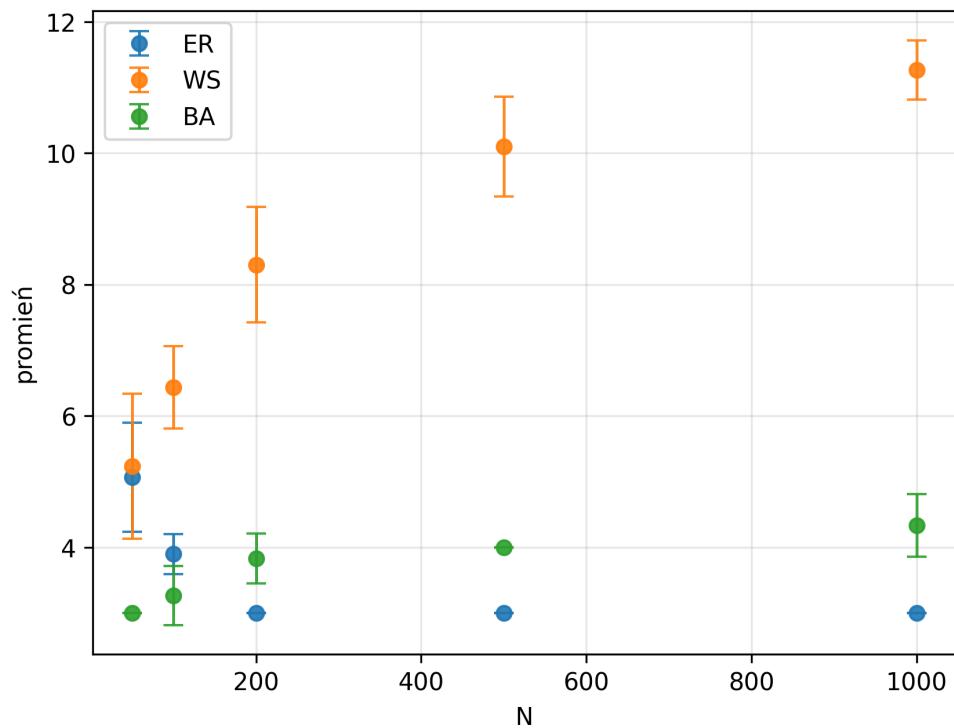
Rysunek 3: Średni współczynnik klasteryzacji $\langle C \rangle$ w funkcji N dla modeli ER, WS i BA (wartości uśrednione po wielu realizacjach).

2.4.4 Średnica i promień vs N

Na rys. 4 oraz 5 pokazano zależność średnicy D_{LCC} i promienia r_{LCC} od liczby wierzchołków. Wartości wyznaczano dla największej spójnej składowej, co zapewnia porównywalność pomiędzy realizacjami również wtedy, gdy cały graf nie jest w pełni spójny.



Rysunek 4: Średnica D_{LCC} w funkcji N dla modeli ER, WS i BA.

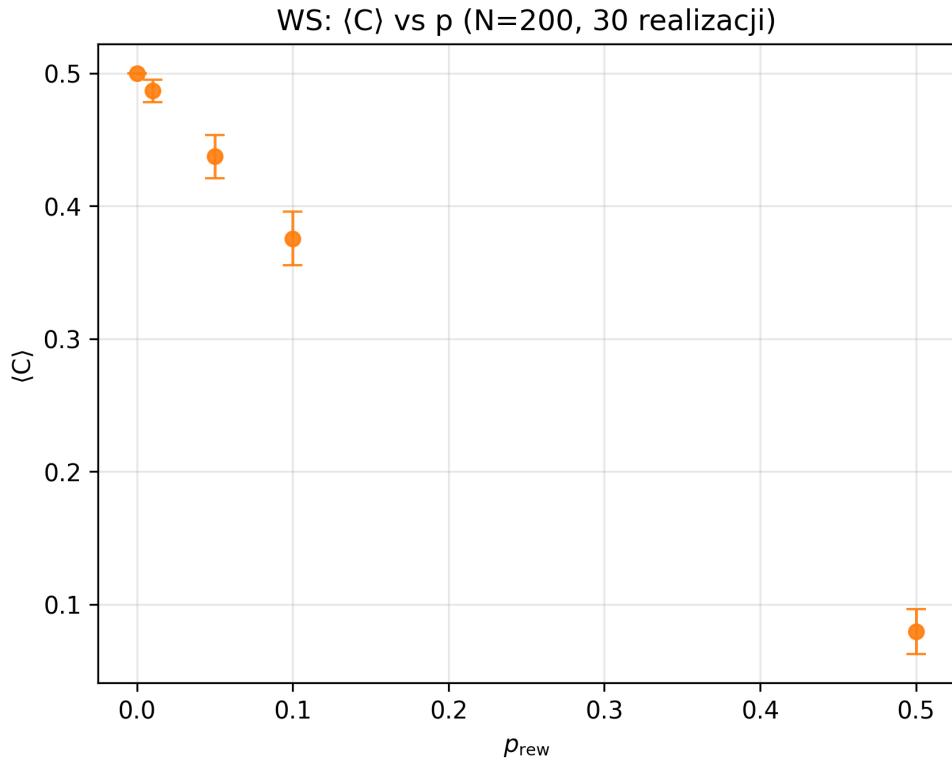


Rysunek 5: Promień r_{LCC} w funkcji N dla modeli ER, WS i BA.

2.4.5 Wpływ parametrów modelu

W ramach analizy wpływu parametrów zbadano dodatkowo model WS dla różnych wartości prawdopodobieństwa przepinania p_{rew} .

Wraz ze wzrostem p_{rew} klasteryzacja maleje, ponieważ losowe przepinanie niszczy lokalne trójkąty charakterystyczne dla sieci regularnej [8, 7].



Rysunek 6: Zależność średniej klasteryzacji $\langle C \rangle$ od parametru przepinania p_{rew} w modelu WS dla $N = 200$.

2.5 Komentarz do wyników

2.5.1 Rozkład stopni $P(k)$: jednorodność vs huby

Na podstawie wykresów rozkładu stopni $P(k)$ (rys. 2) widać wyraźne różnice pomiędzy analizowanymi modelami. W modelach ER i WS większość wierzchołków ma stopnie bliskie wartości typowej (wąski zakres k), co odpowiada sieciom bardziej jednorodnym. W modelu BA pojawia się natomiast dłuższy ogon rozkładu, co oznacza występowanie wierzchołków o znacznie większym stopniu (hubów). Zjawisko to jest naturalną konsekwencją preferencyjnego przyłączania (17) [2, 1].

2.5.2 Klasteryzacja: przewaga modelu WS

Współczynnik klasteryzacji $\langle C \rangle$ (rys. 3) jest najwyższy dla modelu WS, co wynika z konstrukcji sieci startowej typu ring-lattice, gdzie sąsiedzi wierzchołka często są również wzajemnie połączeni. Dla modelu ER klasteryzacja jest relatywnie niska, ponieważ krawędzie powstają

niezależnie i rzadziej tworzą lokalnie zwarte grupy połączeń. Model BA zwykle przyjmuje wartości pośrednie: obecność hubów sprzyja skracaniu ścieżek, ale sam mechanizm dołączania krawędzi nie prowadzi do tak silnego „domykania” połączeń w sąsiedztwach jak w modelu WS.

2.5.3 Średnica i promień: globalne odległości w sieci

Zależności średnicy D_{LCC} i promienia r_{LCC} od liczby wierzchołków N (rys. 4 i 5) pokazują różnice w globalnej strukturze sieci. W praktyce zarówno ER, jak i BA osiągają stosunkowo małe odległości w porównaniu do sieci o silnej regularności, ponieważ losowe połączenia (ER) lub huby (BA) tworzą “skróty” pomiędzy odległymi fragmentami grafu. Dla WS charakterystyczne jest współistnienie wysokiej klasteryzacji i stosunkowo krótkich ścieżek, co odpowiada efektowi *small-world*.

2.5.4 Wpływ parametru p_{rew} w modelu WS

Wpływ parametru przepinania p_{rew} w modelu WS (rys. 6) jest zgodny z oczekiwaniemi: zwiększenie p_{rew} zmniejsza $\langle C \rangle$, ponieważ losowe przepięcia niszczą lokalne trójkąty obecne w sieci regularnej. Jednocześnie nawet niewielkie wartości p_{rew} potrafią istotnie zmieniać strukturę odległości w grafie, dodając “skróty” łączące odległe fragmenty sieci.

2.5.5 Uwagi o losowości i porównywalności

Należy podkreślić, że wyniki zależą od dobranych parametrów (p, k, p_{rew}, m) oraz od losowości pojedynczych realizacji. Z tego względu kluczowe jest uśrednianie metryk po wielu uruchomieniach (wzory (18)-(19)) i analizowanie trendów wraz ze wzrostem N , a nie pojedynczych wartości dla jednej realizacji grafu.

3 Zakończenie

W ćwiczeniu porównano trzy modele sieci losowych: ER, WS i BA. Dla WS wykonano własny generator, a dla wszystkich modeli wyznaczono $P(k)$, $\langle C \rangle$ oraz r_{LCC} i D_{LCC} .

Wyniki wskazują, że ER i WS (dla dobranych parametrów) mają bardziej jednorodne stopnie, natomiast BA tworzy huby (por. (17)). Model WS osiąga najwyższą klasteryzację, a wzrost p_{rew} ją obniża.

Metryki uśredniano po wielu realizacjach i raportowano odchylenia standardowe (18)–(19). W przyszłości analizę można rozszerzyć o inne parametry i dodatkowe miary, np. średnią długość najkrótszej ścieżki.

Literatura

- [1] Réka Albert and Albert-László Barabási. Statistical mechanics of complex networks. *Reviews of Modern Physics*, 74(1):47–97, 2002.
- [2] Albert-László Barabási and Réka Albert. Emergence of scaling in random networks. *Science*, 286(5439):509–512, 1999.
- [3] Béla Bollobás. *Random Graphs*. Cambridge University Press, Cambridge, 2 edition, 2001.
- [4] Paul Erdős and Alfréd Rényi. On random graphs I. *Publicationes Mathematicae*, 6:290–297, 1959.
- [5] Paul Erdős and Alfréd Rényi. On the evolution of random graphs. *Publications of the Mathematical Institute of the Hungarian Academy of Sciences*, 5:17–61, 1960.
- [6] Aric A. Hagberg, Daniel A. Schult, and Pieter J. Swart. Exploring network structure, dynamics, and function using networkx. In *Proceedings of the 7th Python in Science Conference (SciPy2008)*, pages 11–15, 2008.
- [7] Mark E. J. Newman. *Networks: An Introduction*. Oxford University Press, Oxford, 2010.
- [8] Duncan J. Watts and Steven H. Strogatz. Collective dynamics of “small-world” networks. *Nature*, 393:440–442, 1998.