

Sieci społecznościowe w postaci grafów

Patryk Michalak

10 lutego 2025

Spis treści

1	Wprowadzenie	3
2	Erdos-Renyi	3
2.1	Zasada działania algorytmu	3
2.2	Erdos 1 - 256 wierzchołków, Szansa połączenia - 0.005	4
2.3	Erdos 2 - 512 wierzchołków, Szansa połączenia - 0.05	4
2.4	Erdos 3 - 1024 wierzchołków, Szansa połączenia - 0.1	4
2.5	Wniosek	5
3	Watts-Strogatz	5
3.1	Zasada działania algorytmu	5
3.2	Watts 1 - 256 wierzchołków, odległość maksymalna sąsiada - 2, Szansa przekierowania - 0.05	6
3.3	Watts 2 - 512 wierzchołków, odległość maksymalna sąsiada - 4, Szansa przekierowania - 0.1	6
3.4	Watts 2 - 1024 wierzchołków, odległość maksymalna sąsiada - 8, Szansa przekierowania - 0.15	6
3.5	Wniosek	7
4	Barabasi-Albert	7
4.1	Zasada działania algorytmu	7
4.2	Barabasi 1 - Rozmiar początkowego grafu: 4, Rozmiar końcowy: 256, Liczba krawędzi do przyłączenia: 8	8
4.3	Barabasi 2 - Rozmiar początkowego grafu: 4, Rozmiar końcowy: 512, Liczba krawędzi do przyłączenia: 12	8
4.4	Barabasi 3 - Rozmiar początkowego grafu: 8, Rozmiar końcowy: 1024, Liczba krawędzi do przyłączenia: 16	8
4.5	Wniosek	9
5	Podsumowanie	9

1 Wprowadzenie

W poniższym badaniu będziemy badać modele sieci społecznościowych w postaci grafów i różne sposoby na ich stworzenie. Zastosujemy trzy różne algorytmy na tworzenie grafów: Erdos-Renyi, Watts-Strogatz i Barabasi-Albert. Będziemy badać jak wejściowe parametry dla każdego algorytmu wpływa na końcowe wyniki grafów. By ocenić zmiany struktury grafu, poddamy obserwacji pod kątem:

- Rozkładu liczby krawędzi
- Promienia grafu
- Średnicy grafu

Także wyniki będą porównywane z dwoma implementacjami: własno-ręczną i z biblioteki `networkx`.

2 Erdos-Renyi

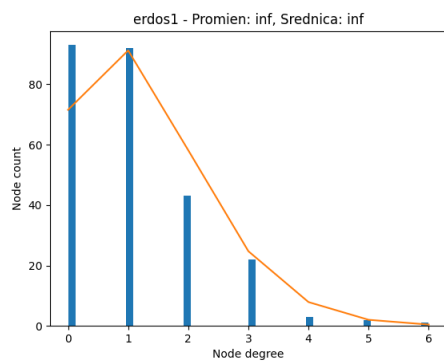
2.1 Zasada działania algorytmu

Algorytm przyjmuje dwa parametry w obu implementacjach - liczbę wierzchołków w grafie (n) i szansa na połączenie pomiędzy parą wierzchołków (p). Algorytm na początku generuje graf o zadanej liczbie wierzchołków, po czym dla każdej pary wierzchołków losuje z prawdopodobieństwem p czy wierzchołki są ze sobą połączone. Rozkład liczby krawędzi w teorii powinien być zbliżony do rozkładu Poissona.

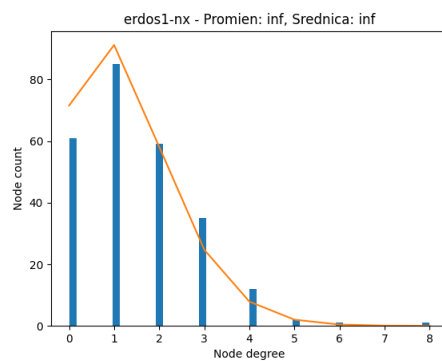
$$P(k) = \frac{(np)^k e^{-np}}{k!}$$

- k - liczba krawędzi wychodząca z wierzchołka

2.2 Erdos 1 - 256 wierzchołków, Szansa połączenia - 0.005

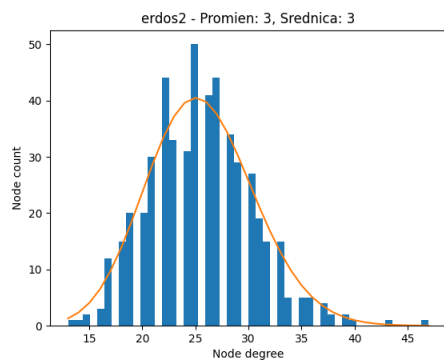


Rysunek 1: Implementacja własnoręczna

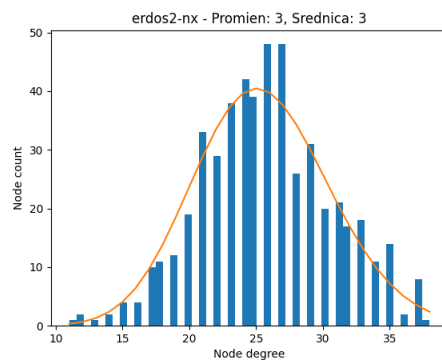


Rysunek 2: Implementacja biblioteczna

2.3 Erdos 2 - 512 wierzchołków, Szansa połączenia - 0.05

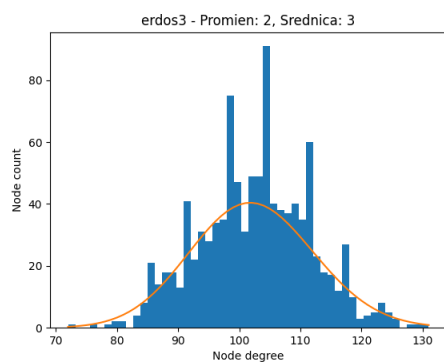


Rysunek 3: Implementacja własnoręczna

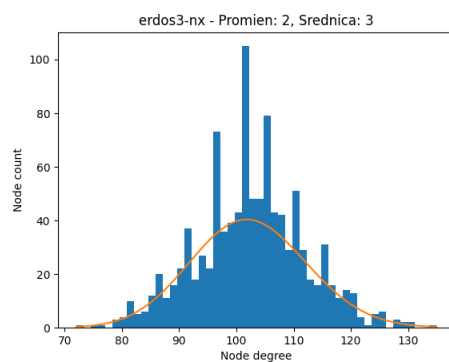


Rysunek 4: Implementacja biblioteczna

2.4 Erdos 3 - 1024 wierzchołków, Szansa połączenia - 0.1



Rysunek 5: Implementacja własnoręczna



Rysunek 6: Implementacja biblioteczna

2.5 Wniosek

Powyższy wykresy przedstawiają że rozkład liczby krawędzi w implementacji bibliotecznej jest zbliżony do teoretycznych wartości rozkładu Poissona. Własnoręczna implementacja jest w mniejszym stopniu zgodna z teoretyczną wartościami. Jednakże wraz z większą liczbą wierzchołków i szansą na połączenia dwóch wierzchołków, rozkłady zbliżają się do oczekiwanych wartości.

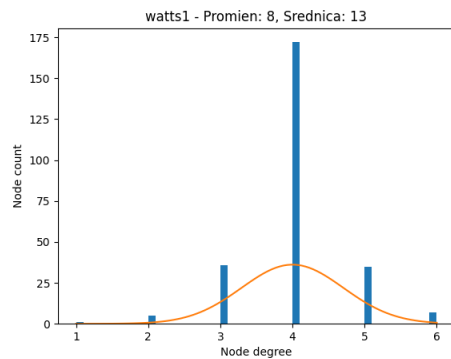
Wraz z wzrostem parametrów, rośnie spójność grafu. Przy niskich wartościach, podgrafy mogą być od siebie odizolowane jednakże w przypadku większych wartości promień i średnica są małe.

3 Watts-Strogatz

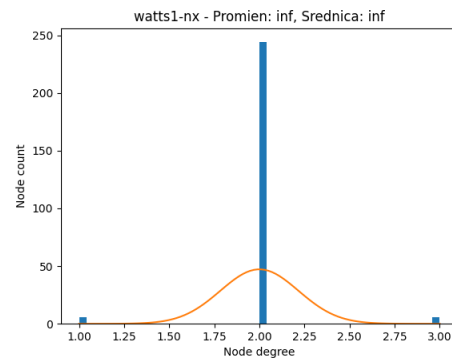
3.1 Zasada działania algorytmu

Algorytm Watts-Strogatz'a przyjmuje jako parametry liczbę wierzchołków n , odległość najdalszego sąsiada k i prawdopodobieństwo przekierowania p . Tworzymy graf, którego wierzchołki są połączone z k poprzednimi sąsiadami. Potem, dla każdej kolejnej krawędzi losujemy z prawdopodobieństwem p czy krawędź połączy się z innym wierzchołkiem końcowym.

3.2 Watts 1 - 256 wierzchołków, odległość maksymalna sąsiada - 2, Szansa przekierowania - 0.05

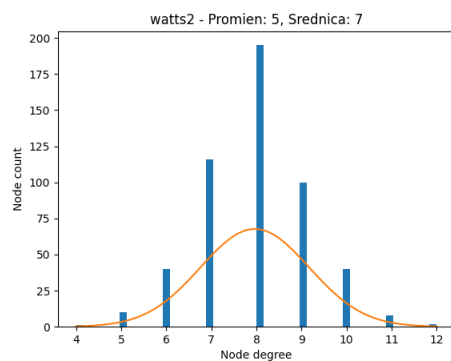


Rysunek 7: Implementacja własnoręczna

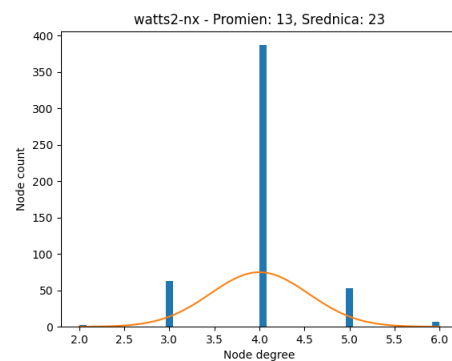


Rysunek 8: Implementacja biblioteczna

3.3 Watts 2 - 512 wierzchołków, odległość maksymalna sąsiada - 4, Szansa przekierowania - 0.1

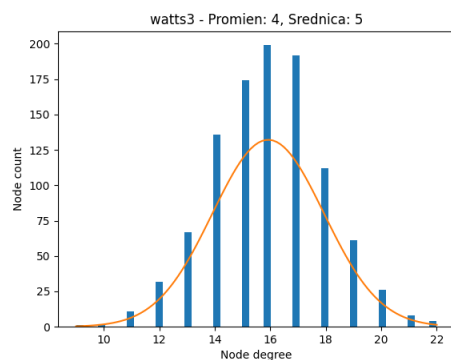


Rysunek 9: Implementacja własnoręczna

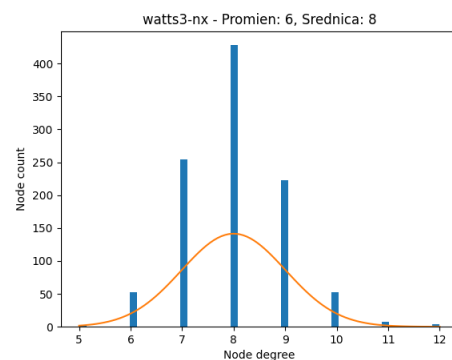


Rysunek 10: Implementacja biblioteczna

3.4 Watts 2 - 1024 wierzchołków, odległość maksymalna sąsiada - 8, Szansa przekierowania - 0.15



Rysunek 11: Implementacja własnoręczna



Rysunek 12: Implementacja biblioteczna

3.5 Wniosek

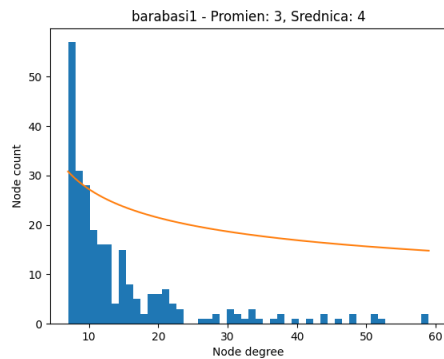
Powyższe wykresy wykazują że rozkład liczby krawędzi podchodzi pod rozkład normalny. W przypadku implementacji własnoręcznej, promienie i średnice grafu osiągają mniejsze wartości niż w przypadku funkcji bibliotecznej. Jednakże, funkcja biblioteczna cechuje się większym dopasowaniem do rozkładu normalnego. W obu przypadkach, promień i średnica maleją, oznaczając lepszą spójność grafu.

4 Barabasi-Albert

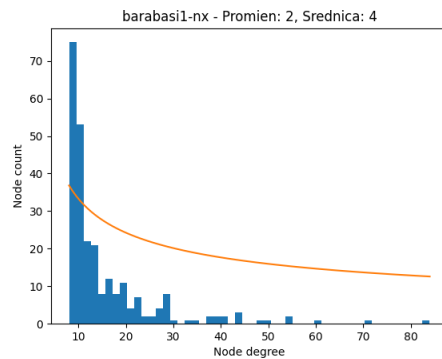
4.1 Zasada działania algorytmu

Algorytm przyjmuje trzy parametry - rozmiar początkowego grafu pełnego, liczbę wierzchołków końcowego grafu i liczbę krawędzi dla każdego nowego wierzchołka. Początkowo, tworzony jest graf pełny o danym rozmiarze. Następnie dodajemy nowe wierzchołki, i przypisujemy jej krawędzie z preferencją do wierzchołkami z większą liczbą krawędzi. Szanse na przyłączenie wynosi $p_i = \frac{k_i}{\sum_j k_j}$, gdzie k_i jest stopień aktualnie wybranego wierzchołka a suma jest sumą wszystkich krawędzi w grafie. Tak musi być połączone do określonej liczby krawędzi. Powtarzamy dopóki rozmiar grafu będzie równa końcowej liczbie określonej jako parametr.

4.2 Barabasi 1 - Rozmiar początkowego grafu: 4, Rozmiar końcowy: 256, Liczba krawędzi do przyłączenia: 8

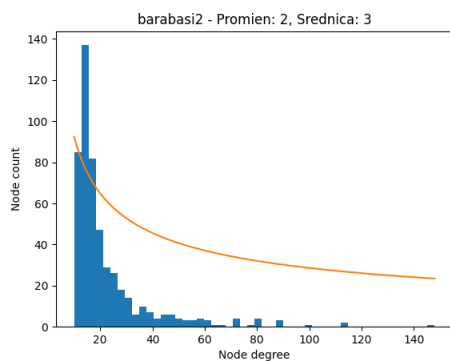


Rysunek 13: Implementacja własnoręczna

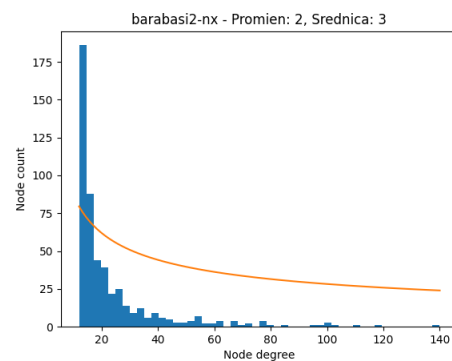


Rysunek 14: Implementacja biblioteczna

4.3 Barabasi 2 - Rozmiar początkowego grafu: 4, Rozmiar końcowy: 512, Liczba krawędzi do przyłączenia: 12

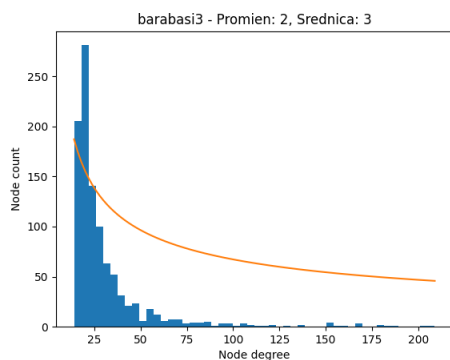


Rysunek 15: Implementacja własnoręczna

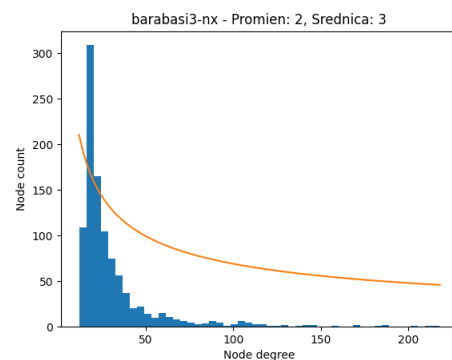


Rysunek 16: Implementacja biblioteczna

4.4 Barabasi 3 - Rozmiar początkowego grafu: 8, Rozmiar końcowy: 1024, Liczba krawędzi do przyłączenia: 16



Rysunek 17: Implementacja własnoręczna



Rysunek 18: Implementacja biblioteczna

4.5 Wniosek

Algorytmy w obu przypadkach zwracają podobne wyniki, Przeważająca liczba wierzchołków ma stosunkową małą liczbę krawędzi a tylko pojedyncze mają proporcjonalnie dużą liczbę krawędzi. Wierzchołki z dużą liczbą krawędzi noszą miano "hub-ów" a grafy z takim rozkładem krawędzi posiadają atrybut "małych światów". Rozkład podchodzi pod rozkład potęgowy $p(k) = k^{-3}$

5 Podsumowanie

Implementacje zarówno własnoręczna jak i biblioteczna są zaimplementowane prawidłowo, porównując do siebie oba rozkłady obu implementacji jak i teoretycznych wartości spodziewanych.

Model Erdos-Renyi ma charakter losowy, jego rozkład podlegającym rozkładowi Poissona, która wraz z większą liczbą węzłów i zwiększaną prawdopodobieństwem połączenia dwóch losowo wybranych wierzchołków rośnie spojność grafu, cechując się niskim promieniem i średnicą grafu.

Model Watts-Strogatz'a posiada wyniki najbardziej różniące się, własnoręczna implementacja ma większą wierzchołków o stopniu innym od średniego stopnia grafu w porównaniu do implementacji z biblioteki. To wpływa na końcowe promienie i średnice grafów, gdzie w przypadku własnoręcznej implementacji promień i średnica jest stosunkowo niższa niż w przypadku bibliotecznej. Jednakże, graf z funkcji bibliotecznej wykazuje lepsze dopasowanie do rozkładu normalnego.

Model Barabasi-Alberta cechuje najmniejsza różnica pomiędzy implementacjami i końcowymi cechami. W wszystkich przypadkach występują pojedyncze wierzchołki z dużą ilością krawędzi, duża liczba wierzchołków z małym stopniem i niskim promieniem i średnicą grafu.