

GIS - sprawozdanie 3

Autorzy:

Marcin Dzieżyc
Mateusz Statkiewicz

Prowadzący:

prof. dr hab. Jacek Wojciechowski

Temat:

Opracowanie generatora modeli grafowych różnego typu sieci.

Cel projektu:

Celem projektu jest opracowanie generatora różnego typu sieci, potrafiącego wyświetlić wygenerowane sieci. Program będzie generował sieci losowe, euklidesowe, bezskalowe i sieci małego świata. Dodatkowo będzie sprawdzał spójność sieci.

Założenia początkowe:

Sieci generowane przez program będą grafami nieskierowanymi, z określonymi wagami krawędzi.

Wykorzystane biblioteki:

- **graphviz (libgvc):**
<http://www.graphviz.org/> - API definiowania i plotowania grafów

Algorytmy zastosowane w projekcie:

- **generowanie losowej sieci (algorytm Erdosa-Renyiego)**

Algorytm losujący graf na podstawie niezależnego prawdopodobieństwa wyboru danej krawędzi ze zbioru krawędzi grafu pełnego o n wierzchołkach (w odróżnieniu od generatora losującego z prawdopodobieństwem równomiernym graf o n wierzchołkach i m krawędziach spośród wszystkich możliwych o tych parametrach). Generator pobiera na wejściu liczbę wierzchołków oraz prawdopodobieństwo wyboru poszczególnych krawędzi. Każdy możliwy do wygenerowania graf o tych parametrach będzie mógł pojawić się na wyjściu z prawdopodobieństwem zgodnym z rozkładem Bernoulliego. Wagi poszczególnych krawędzi losowane będą z rozkładem jednostajnym z przedziału $[0, 1]$.

- **generowanie sieci euklidesowej**

Na wejściu generatora tej sieci znajdują się 3 argumenty:

- liczba wierzchołków sieci;
- bok kwadratu, na którym rozmieszczone będą wierzchołki;
- promień, który będzie określał maksymalną odległość wierzchołków połączonych krawędzią;

Algorytm polega na wylosowaniu odpowiedniej liczby punktów wewnątrz kwadratu o danym boku. Następnie krawędzie tworzone są między wierzchołkami, których odległość jest mniejsza od zadanego promienia. Aby zapewnić spełnienie nierówności trójkąta krawędzie mają wagi odpowiadające odległości między wierzchołkami, które łączą.

- **generowanie sieci małego świata (algorytm Watts-Strogatz)**

Na wejściu generatora tej sieci znajdują się 3 argumenty:

- liczba wierzchołków sieci;;
- początkowa liczba sąsiadów każdego wierzchołka;
- prawdopodobieństwo przestawienia krawędzi;

Algorytm polega na utworzeniu pierścienia, w którym każdy z wierzchołków jest połączony krawędziami z daną liczbą sąsiadów. Kolejny krok to wymiana wierzchołków krawędzi z zadanym prawdopodobieństwem.

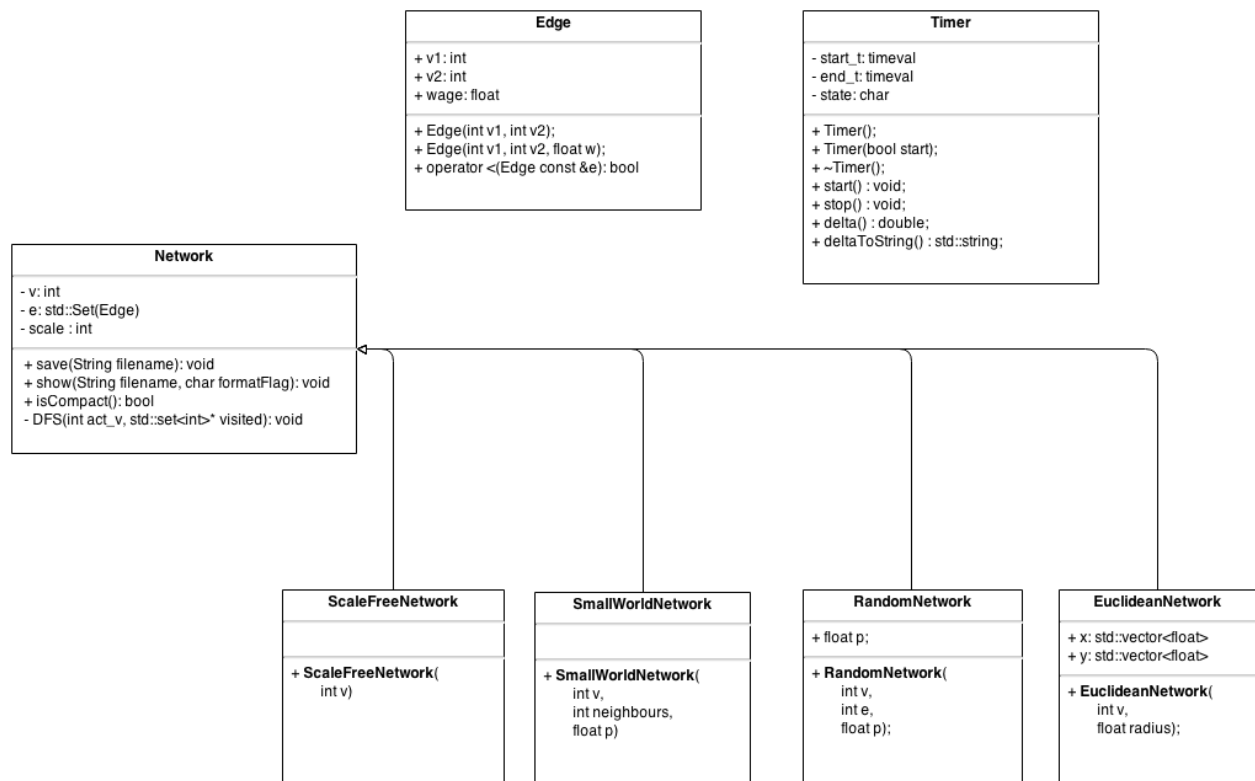
- **generowanie sieci bezskalowej (algorytm Barabasi-Albert)**

Algorytm budujący sieć bezskalową iteracyjnie dodając nowe wierzchołki oraz łącząc je z resztą sieci w oparciu o model Barabasi-Alberta. Prawdopodobieństwo wyboru krawędzi zależy wprost proporcjonalnie od stopnia przeciwległego wierzchołka krawędzi, co skutkuje kumulacją krawędzi w pojedynczych wierzchołkach w miarę rozrastania się grafu. Sieć taka charakteryzować się będzie rozkładem stopni wierzchołków postaci $P(k) \sim k^{-\gamma}$, gdzie $P(k)$ to ułamek wierzchołków stopnia k w sieci, zaś γ jest parametrem charakterystycznym sieci, zwykle z przedziału $[2, 3]$.

- **sprawdzanie spójności grafu (algorytm DFS)**

Rekurencyjne sprawdzenie czy wszystkie wierzchołki należą do spójnego grafu. Rozpoczynając od dowolnego wierzchołka, oznacz go jako odwiedzony i przejdź do wszystkich sąsiednich, które nie zostały wcześniej oznaczone.

Struktury danych



Złożoność obliczeniowa

Dla każdego generatora złożoność obliczeniowa wynosi $O(n^2)$, gdzie n - liczba wierzchołków lub, w przypadku sieci małego świata, liczba generowanych krawędzi.

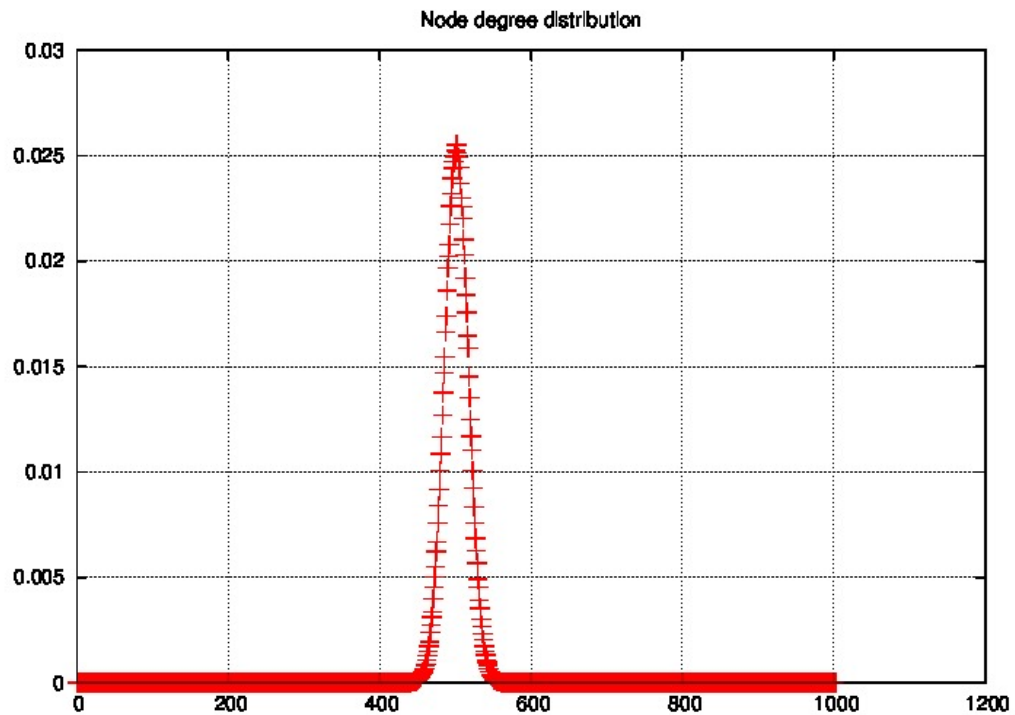
Czas wykonywania

Czas generacji sieci, jak i sprawdzania spójności jest wyświetlany użytkownikowi. Pomiary czasu zostały wykonane w przypadku wygenerowania spójnych sieci. Czasy wywołań przedstawia tabelka (czasy w sekundach):

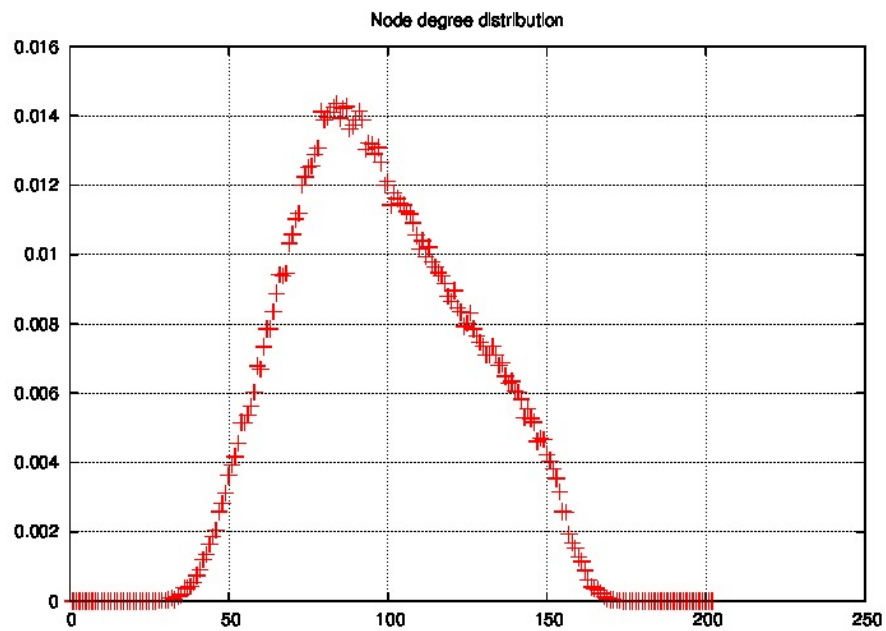
<i>typ sieci</i>	<i>Euklidesowa</i>	<i>Losowa</i>	<i>Małego Świata</i>	<i>Bezskalowa</i>
<i>parametry</i>	<i>$r=0.6$</i>	<i>$p=0.3$</i>	<i>$n=4, p=0.6$</i>	<i>$m0=3, m=2$</i>
<i>$v=10$</i>	0.00011	0.000054	0.000144	0.000083
<i>parametry</i>	<i>$r=0.4$</i>	<i>$p=0.4$</i>	<i>$n=10, p=0.6$</i>	<i>$m0=5, m=4$</i>
<i>$v=100$</i>	0.003673	0.004628	0.003036	0.001222
<i>parametry</i>	<i>$r=0.2$</i>	<i>$p=0.6$</i>	<i>$n=50, p=0.6$</i>	<i>$m0=10, m=6$</i>
<i>$v=1000$</i>	0.199363	1.053605	0.209740	0.035874

Testy

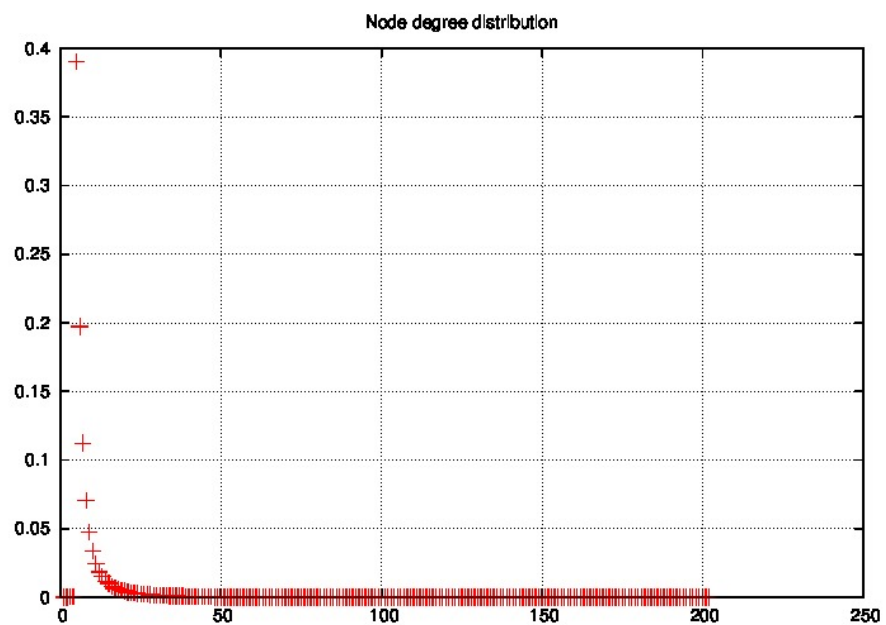
W ramach testów poprawnościowych zaimplementowano procedurę analizującą rozkład stosunku liczby wierzchołków o zadanym stopniu do liczby wszystkich węzłów sieci danego rodzaju. Zamieszczone poniżej zostały przykładowo wygenerowane rozkłady (wraz z parametrami podanymi do generatora).



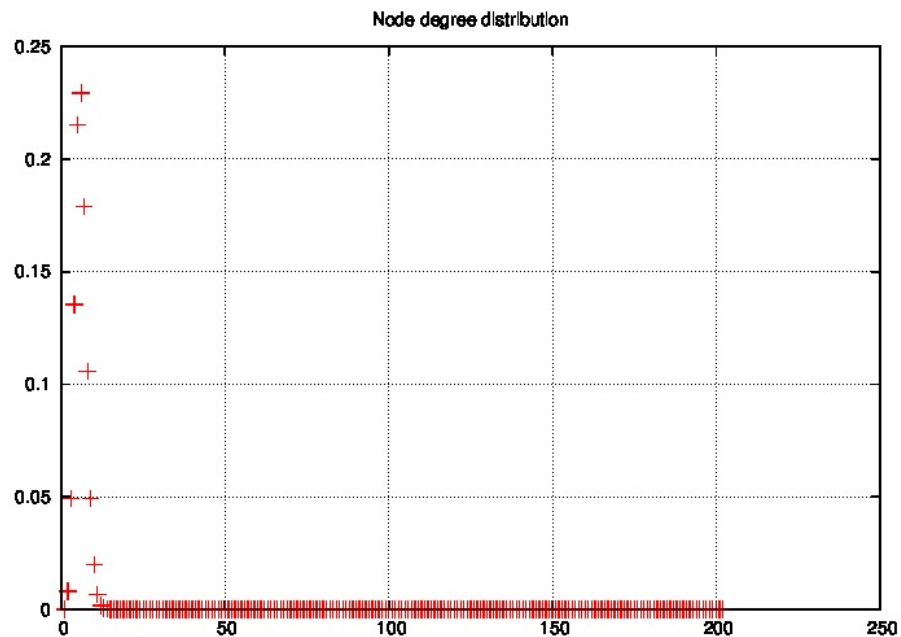
Rozkład sieci losowej (1000 wierzchołków, prawdopodobieństwo wyboru 0.5)



Rozkład sieci euklidesowej (Liczba wierzchołków 200, promień okręgu 0.5)



Rozkład sieci bezskalowej (Liczba wierzchołków 200, początkowa sieć 5 wierzchołków, liczba sąsiadów 3)



Rozkład sieci małego świata (Liczba wierzchołków 200, początkowa liczba sąsiadów 5, prawdopodobieństwo zmiany 0.5)

Wnioski

Osiągane wyniki testów poprawnościowych potwierdziły oczekiwania wywodzące się z naukowych opisów algorytmów oraz modeli wykorzystanych do generacji sieci o zadanym typie i parametrach. Złożoność obliczeniowa nie odbiegała znacząco od przewidywanej złożoności pesymistycznej $O(n^2)$.

Literatura

- A. Reka, A.L. Barabasi, *Statistical mechanics of complex networks. Review of Modern Physics* Vol. 74, January 2002.
- J.Wojciechowski, K.Pieńkosz, *Grafy i sieci*, PWN 2013.