Grafy i sieci:
Generator sieci bezskalowej II (model Barabasiego-Albert)

Eryk Warchulski
Kanstantsin Padmostka
Prowadzący: dr inż. Sebastian Kozłowski

10 marca 2019
wer. 1.0

Spis treści

Streszczenie

Dokument ten zawiera szczegółowy opis zadania projektowego, który ma potwierdzić zrozumienie tematu przez autorów. Tematem zadania jest implementacja szybkiego i przenośnego generatora sieci bezskalowych w wybranym języku programowania. W sekcji (1) znajduje się omówienie tematu zadania, a w sekcjach (2) i (3) ogólny opis grafów losowych oraz opis modelu Barabasiego-Albert.

1 Opis zadania

Postawionym przed nami zadaniem jest zaimplementowanie generatora grafów losowych w ujęciu Barabasieg-Albert, które zostanie szczegółowo opisane w dalszej części dokumentu (3). Generator poza spełnianiem swojej podstawowej funkcji musi charakteryzować się przenośnością oraz jak najmniejszą złożonością obliczeniową i pamięciową. Dla poprawnie działającego generatora kolejnym krokiem w realizacji zadania jest zbadanie rozkładów stopni wierzchołków grafów i porównanie ich z modelami teoretycznymi. Pełna realizacja zadania zakłada istnienie możliwości zapisu wygenerowanego grafu do ustalonego formatu, co pozwoli odwtorzyć sam graf oraz przebieg eksperymentów numerycznych.
Dokumentacja ta jest wolna od opisu implementacji generatora, eksperymentów numerycznych oraz sposobu ich wizualizacji. Kwestie te zostaną omówione w kolejnych wersjach dokumentacjach, tj. odpowiednio: wersji 2.0 oraz 3.0.

2 Grafy losowe

Stosowanie teorii grafów do modelowania zjawisk zachodzących w realnym świecie jest oparte w dużej mierze na grafach losowych. Podyktowane jest to faktem, że zjawiska te i towarzyszące im zdarzenia wykazują w skali makroskopowej charakter stochastyczny. Przykłady dziedzin, w których stosowane są grafy losowe do modelowania pewnych zjawisk są następujące:

- sieci połączeń handlowych
- sieci `WWW`
- sieci neuronowe (rekurencyjne)
- sieci społecznościowe (np. `Facebook`)

Zdefiniowanie grafu losowego wymaga z kolei zdefiniowania struktur jak *przestrzeń grafów losowych* $\mathcal{G}$, która jest wyposażona w unormowaną miarę $\mathbb{P}(\bullet)$ mówiącą o prawdopodobieństwie wylosowania grafu G o pewnych właściwościach [2].
Zadanie to ze względu na złożoną strukturę obiektów jakimi są grafy nie jest tak intuicyjne jak określenie przestrzeni probabilistycznej dla zdarzeń, które można reprezentować liczbami. Z tego względu istnieje szereg alternatywnych modeli, które podejmują się rozwiązania tego zadania. Pokrótce zostanie omówiony najstarszy i najprostszy model wprowadzony przez Erdős'a i Rényi'e jeszcze w latach 60. ubiegłego wieku [1].

2.1 Model E-R

Model ten jest oparty o dwójkę parametrów (n, p): parametr $n \in \mathbb{N}$ oznacza liczbę wierzchołków generowanego grafu G, a $p \in (0, 1)$ stanowi o prawdopodobieństwie zdarzenia polegającego na zaistnieniu krawędz między każdą parą z n^2 wierzchołków grafu G.
Na podstawie powyższego łatwo widać, że rozkład stopni wierzchołków w grafie zadany jest przez rozkład dwumianowy z funkcją gęstości prawdopodobieństwa

$$p(n, k; p) = \binom{n-1}{k} p^k (1-p)^{n-k-1} \tag{1}$$

implikuje to fakt, że średni stopień wrzechołka $\mathbb{E}deg(v)$ wynosi $(n-1)p$. Ponadto, prawdopodobieństwo wylosowania grafu E-R o e krawędziach i n wierzchołkach wynosi $\binom{\binom{n}{2}}{m} p^m (1-p)^{\binom{n}{2}-m}$.

Na tej podstawie liczba wszystkich możliwych grafów E-R o n wierzchołkach wynosi

$$\sum_{e=0}^{\binom{n}{2}} \binom{\binom{n}{2}}{e} = 2^{\binom{n}{2}} \tag{2}$$

przy czym $\binom{\binom{n}{2}}{e}$ oznacza liczbę e-elementowych kombinacji zbioru utworzonego ze wszystkich par zbioru n-elementowego.
Niestety, model taki nie jest najlepszym kandydatem do *naśladowania* obiektów rzeczywistych. Przy $p << 1$ rozkład stopni wierzchołków dany jest rozkładem Poissona, tj. rozkładem, który stosowany jest do zdarzeń rzadkich występujących w określonym przedziale czasu. Grafy generowane w tym modelu nie są w stanie dobrze odwzorowywać *hub*-ów, tj. skupisk.
Modele, które są wolne powyższej opisanych wad grafów opartych o model E-R, oparte są o rozkłady potęgowe i zostaną opisane w następnej sekcji (3).

3 Model Barabasiego-Albert

TODO: nawiązanie do poprzedniego rozdziału i napisanie motywacji modeli BA w kontekście zasady maksymalnej e

3.1 Sieć bezskalowa

TODO: zdefiniowanie sieci bezskalowej

3.2 Preferencyjne dołączanie wierzchołków

TODO: opisanie na czym polega ten mechanizm

3.3 Rozkład stopni wierzchołków

TODO: wyprowadzenie zależności na rozkład stopni wierzchołków i napisanie, że są różne

3.3.1 Model czasu ciągłego

TODO: wyprowadzenie r.s.w. dla modelu czasu ciągłego

3.3.2 Model równania *master*

TODO: to samo co wyżej

Literatura

[1] P. Erdős and A Rényi. On the evolution of random graphs. In *PUBLICATION OF THE MATHEMATICAL INSTITUTE OF THE HUNGARIAN ACADEMY OF SCIENCES*, pages 17–61, 1960.

[2] Agata Fronczak. Wykładnicze grafy przypadkowe: teoria, przykłady, symulacje numeryczne. 2014.