Modelowanie Procesów Stochastycznych

Tytuł projektu: **Sieci BA**

Autor: Aleksandra Buczek

SIEĆ BA

sieć ewoluująca – rośnie i zmienia się w czasie

model sieci o potęgowym rozkładzie stopni wierzchołków: $\boldsymbol{P(k) \sim k^{-\gamma}}$

sieć wzrasta zgodnie z regułą preferencyjnego dołączania węzłów - nowy węzeł chętniej tworzy połączenie z lepiej usieciowionymi węzłami

reguła PAR jest liniowa – prawdopodobieństwo wylosowania węzła, z którym łączy się nowy węzeł, jest proporcjonalne do stopnia węzła wylosowanego: $\pi(k_i) \sim k_i$

w granicy t -> ∞ rozkład P(k) wynosi:

$$\boldsymbol{P(k) = \frac{2m^2}{k^3}} \quad (0)$$

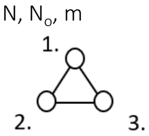
IMPLEMENTACJA ALGORYTMU BA: (etap 1. projektu)

parametry sieci BA:

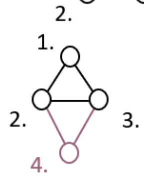
* N_o – liczba węzłów w klastrze początkowym, $N_o >= 1$
* m – liczba połączeń z innymi węzłami utworzonych przez nowy dodany węzeł, $m <= N_o$
* N – końcowa liczba węzłów sieci

proces tworzenia sieci BA:

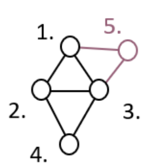
0) Podanie parametrów: N, N_o, m
1) Segment początkowy:



2) Nowy węzeł:

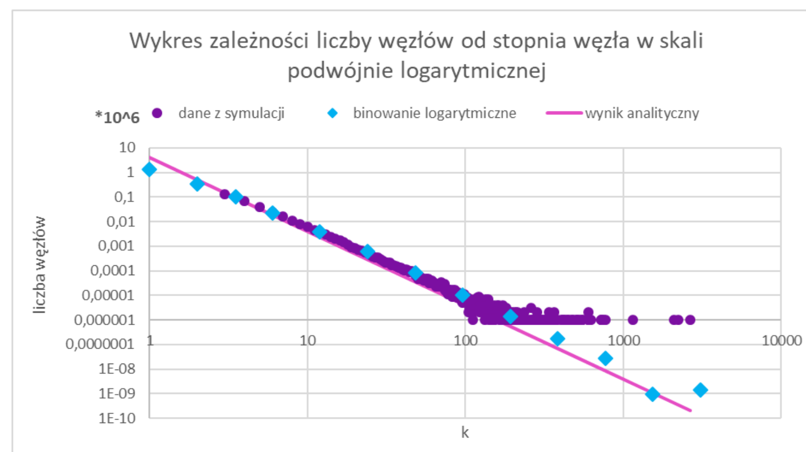


3) Kolejne węzły...



... aż do dodania N węzłów

W ramach projektu przeprowadzono symulację sieci BA dla: $N_o = m = 1$ oraz $N = 2 \cdot 10^6$:



W skali podwójnie logarytmicznej rozkłady potęgowe przyjmują postać: $lnP(k) = -\gamma\, lnk + lnC$, gdzie: C – stała. Odpowiada to liniowemu równaniu: $y = ax + b$.
Na wykresie można zaobserwować zależność liniową. Po dopasowaniu prostej (wyznaczonej poprzez binowanie logarytmiczne) do danych z symulacji, otrzymano wartość **γ = 2.82**, czyli zbliżoną do wyniku analitycznego, w którym **γ = 3**.

PROCEDURA TWORZENIA SIECI BA, A ŁAŃCUCH MARKOWA: (etapy 2. i 3. projektu)

Dla sieci BA jednym ze sposobów wyznaczenia rozkładu stopni wierzchołków jest skorzystanie z metody równania master. Wykorzystywany jest rozkład prawdopodobieństwa stopni węzłów, gdzie $k_i(t)$ to zmienna losowa. Przyjmując, że $p_i(k, t_i, t)$ to prawdopodobieństwo, że węzeł dodany do sieci BA w kroku czasowym t_i, w chwili t ma stopień wynoszący k_i, można zapisać:

$$p_i(k, t_i, t+1) = \sum_{l=0}^{m} p(l;m) p_i(k-l, t_i, t), \quad (1)$$

gdzie p(l;m) - prawdopodobieństwo otrzymania przez węzeł l (l ≤ m) połączeń (opisane rozkładem Bernoulliego). Równanie to pokazuje prawdopodobieństwo, że węzeł uzyska nowe krawędzie i zwiększy swój stopień o l, zmieniając w ten sposób swój stan. Dla wszystkich węzłów w sieci BA rozkład prawdopodobieństwa wynosi:

$$P(k, t_i, t) = \frac{1}{t}\sum_{t_i=1}^{t} p_i(k, t_i, t). \quad (2)$$

Na podstawie powyższych równań oraz po uwzględnieniu granicy t -> ∞, otrzymano:

$$P(k, t_i, t+1) = \left(1 - \frac{k}{2t}\right) P(k, t_i, t) + \left(\frac{k-1}{2t}\right) P(k-1, t_i, t). \quad (3)$$

Znajdująca się w algorytmie tworzenia sieci BA reguła preferencyjnego dołączania powoduje,że:

* kolejne stopnie wierzchołka $k_i(t)$ zmieniają się w krokach czasowych: t = i, i + 1, i + 2,..., jak: m, m + 1, m + 2...; w danym czasie zmiana ta związana jest z pewną losowością,
* stopień węzła ewoluuje na skutek dołączania do sieci nowych węzłów, pojawiających się z kolejnymi krokami czasowymi; jest tym samym zależny od czasu, z wartością stopnia zależną od poprzedniej.

Na podstawie tych obserwacji, można stwierdzić, że procedura tworzenia sieci BA jest procesem stochastycznym, a stopnie wierzchołków (i wraz z nimi ich rozkład) tworzą łańcuch Markowa z prawdopodobieństwami przejść zależnymi od czasu, uzyskanymi z równania (3) i wynoszącymi:

$$\begin{cases} 1 - \dfrac{k}{2t}, \textit{gdy stopień węzła nie ulega zmianie}, \\ \dfrac{k}{2t} \quad , \textit{gdy stopień węzła zwiększa się o } 1. \end{cases} \quad (4)$$

Prawdopodobieństwa te są zmienne w czasie, co oznacza, że jest to niejednorodny łańcuch Markowa.
W granicy t -> ∞, czas można przyjąć jako zmienną ciągłą:

$$P(k, t_i, t+1) - P(k, t_i, t) = \frac{dP(k, t_i, t)}{dt}. \quad (5)$$

Podstawiając powyższe równanie do (3) oraz przyjmując, że dla dużych czasów $P(k,t_i,t)$ -> P(k), otrzymano:

$$P(k) = \left(\frac{k-1}{k+2}\right) P(k-1). \quad (6)$$

Po zastosowaniu rekurencji oraz dla warunku początkowego: $P(m) = \frac{2}{m+2}$, wyznaczono równanie na rozkład stopni wierzchołków w sieci BA:

$$P(k) = \frac{2m(m+1)}{k(k+1)(k+2)}. \quad (7)$$

Można zauważyć, że w granicy k >> m, rozkład ten przyjmuje postać równoważną do równania (0), czyli postaci analitycznej. Uzyskana z symulacji wartość wykładnika γ = 2.82, pokazuje, że również w tym przypadku rozkład stopni wierzchołków dąży do omawianej postaci. Dla tych podejść otrzymano równoważne wyniki.

Podsumowanie:

Występująca w sieci BA reguła preferencyjnego dołączania ma znaczący wpływ na sposób, w jaki sieć ewoluuje. Nowe węzły wybierają z większym prawdopodobieństwem bardziej usieciowione, starsze węzły niż te, o mniejszym stopniu. W rezultacie, węzły o wyższym stopniu mają większą szansę na dalsze podwyższanie tego parametru. Jest to zjawisko występujące w wielu rzeczywistych sieciach złożonych, znane pod nazwą efektu świętego Mateusza (lub rich get richer). Pojawia się np. w sieci cytowań, czy sieci kontaktów społecznych. Węzły wybierane są losowo, co wprowadza do procesu tworzenia sieci BA aspekt stochastyczny.