Autor:

Piotr Turecki ***pturecki@wp.pl***

Przedmiot:

Modelowanie rzeczywistości

Temat:

Modele sieci o własnościach zaobserwowanych w rzeczywistych sieciach (small-world oraz scale-free)

Prowadzący:

Dr hab. Witold Dzwinel

1. *Wstęp*

Złożone systemy często modeluje się przy pomocy sieci. Celem tworzenia modeli naturalnych sieci jest umożliwienie badania ich struktury a także badania wielu zjawisk, które sieci dotyczą, np. rozprzestrzeniania się epidemii, odporności na awarie, itp.

Badanie struktury sieci ma duże znaczenie przy konstrukcji algorytmów, które na takich sieciach mają działać (np. routing). W takich przypadkach szczególnie ważne jest stworzenie modelu sieci, który jak najlepiej odpowiada rzeczywistej sieci (np. sieci WWW) ponieważ z praktycznego punktu widzenia algorytmy muszą być zoptymalizowane tylko dla tych przypadków, które istnieją w rzeczywistym świecie, a nie dla wszystkich możliwych matematycznych modeli.

Opracowano wiele modeli sieci. Od najprostszych czyli grafów losowych, poprzez modele dobrze opisujące sieci small-world, następnie sieci scale-free world a kończąc na modelach, które posiadają obie te właściwości. Badania eksperymentalne wykazały, że większość sieci istniejących w rzeczywistości posiada właściwości small-world oraz scale-free world, przez co ostatni model jest obecnie najistotniejszy z praktycznego punktu widzenia.

2. *Modelowanie sieci*

Sieć składa się z węzłów oraz połączeń między nimi. Węzły lub połączenia mogą psiadać dodatkowe informacje, np. wagi, nazwy, itp. Powszechnie używa się sieci do opisu wielu złożonych systemów w wielu dziadzinach nauki – inżynierii, biologii, socjologii, lingwistyce i innych.

Modelowanie złożonego systemu jako zbioru węzłów i połączeń oczywiście nie może zachować całej informacji istniejącej w oryginalnym systemie. W każdym modelu występują ograniczenia i obowiązkiem osoby modelującej jest wybranie tych informacji, które mają wpływ na postawiony cel. Im prostszy model tym łatwiej go zbadać i – niestety – tym mniejsze ma zastosowanie praktyczne.

Dla przykładu można modelować sieci symulujące:

- Sieć WWW (optymalizacja, odporność na awarie, itp.)
- Komunikację w ludzkiej populacji (rozprzestrzenianie się epidemii)
- Sieć reprezentującą interakcje zachodzące w ludzkich komórkach (zrozumienie metabolizmu)
- Sieć elektryczną (odporność na awarie, optymalizacja)
- Sieć komunikacji w mieście (optymalizacja)
- i wiele innych

3. *Historia*

Wraz z kolejnymi badaniami eksperymentalnymi nad rzeczywistymi sieciami, które pokazywały ich nowe właściwości, tworzone były nowe modele sieci, które coraz lepiej zachowywały te właściwości.

Grafy losowe

Badania nad grafami losowymi zaczęły się najwcześniej bo już w latach 60-tych. Nie polegały one na modelowaniu rzeczywistych sieci, tylko na teoretycznym rozważaniu własności takich grafów. Niedługo okazało się też, że odkryte własności nie potwierdzają się do końca w sieciach rzeczywistych.

Small-world

W 1998 roku Watts i Strogatz wykazali, że sieć połączeń w niektórych rzeczywistych sieciach nie jest ani całkiem losowa, ani całkiem uporządkowana. Nazwali je sieciami small-world. Sieci te wykazują się wysokim współczynnikiem skupienia (jak połączenia w regularnej kracie) oraz małą średnią odległością między węzłami (jak w losowych gafach). Watts i Strogatz zaproponowali też prosty model opisujący jak konstruować sieci o takich własnościach (model WS).

Trzy lata później wprowadzono pojęcie efektywności określające jak efektywnie jest wymieniana informacja w sieci. Używając nowej miary wykazano, że sieci small-world są lokalnie i globalnie efektywne. Dodatkowo badanie sieci pod względem efektywności umożliwiło rozszerzenie badań na sieci, w których nie zawsze można się dostać z jednego węzła do drugiego oraz na sieci z wagami, które są bliższe rzeczywistości.

Scale-free

Jednocześnie inne badania wykazały, że rzeczywiste sieci posiadają jeszcze jedną ważną własność: ilość połączeń (k) wychodzących z węzła jest proporcjonalna do jej potęgi (tzw. power-law)

$$P(k) \sim k^{-y}$$

Wykazano, że wykładnik jest niemal stały i wynosi –3 (tzn. parametr y = 3). Nad takimi sieciami pracował najpierw Price (jeszcze latach 60-tych, ale jego model nie był bardzo dokładny) oraz inspirowani jego pracą Barabasi i Albert (wiele lat później – bo po koniec lat 90-tych). Zaproponowali oni też model pozwalający skonstruować sieć o właściwościach scale-free (model BA). Wykazali też, że sieci scale-free wykazują dużą odporność na błędy, tzn. usunięcie losowo wybranych węzłów nie powoduje zakłóceń w komunikacji sieci oraz że są wyjątkowo nieodporne na ataki, tzn. usunięcie węzłów, które pełnią kluczową rolę w komunikacji. Wszystkie te własności zaobserwowano w rzeczywistych sieciach.

Small-world i scale-free

Model BA pozwala na stworzenie sieci, w której liczba połączeń węzła jest proporcjonalna do ich potęgi. Niestety sieci te nie posiadają własności small-world.

Sieci scale-free posiadają jedynie pierwszą własność sieci small-world a mianowicie średnia odległość między węzłami jest mała. Nie posiadają wysokiego współczynnika skupienia – równie ważnej własności small-world. W 2002 roku Klemm i Eguiluz zaproponowali następny model sieci (model KE), który umożliwia stworzenie sieci, w której rozkład połączeń zachowuje własność scale-free i jednocześnie w sieci występuje mała średnia odległość między węzłami oraz występuje wysoki współczynnik skupienia. Model ten obecnie najlepiej opisuje własności rzeczywistych sieci.

4. Właściwości sieci

1. Efekt small-world

Występuje wtedy gdy średnia odległość między węzłami jest stosunkowo mała (np. w porównaniu z ilością węzłów). Średnią odległość obliczamy jako:

$$l = \frac{1}{\frac{1}{2}n(n+1)} \sum_{i \geq j} d_{ij}$$

Gdzie n jest ilością węzłów a d_{ij} jest odległością między węzłami „i” oraz „j”. Często, dodatkowo ten wynik mnoży się przez $\frac{n-1}{n+1}$ co nie ma istotnego znaczenia dla praktycznych celów.

Obliczenie średniej odległości stwarza problemy jeśli sieć składa się więcej niż z jednej części. Dlatego czasami oblicza się ją na wzór harmonicznej średniej:

$$l^{-1} = \frac{1}{\frac{1}{2}n(n+1)} \sum_{i \geq j} d_{ij}$$

Bardziej precyzyjna definicja efektu small-world mówi, że efekt ten występuje jeśli wartość „l” rośnie logarytmicznie lub wolniej wraz ze wzrostem węzłów w sieci przy jednakowej liczbie połączeń wychodzących z węzła. Logarytmiczny wzrost został pokazany dla wielu sieci występujących w rzeczywistości (a także dla losowych grafów).

2. Skupienie (lub przechodniość)

W wielu sieciach wykazano, że jeśli węzeł A jest połączony z węzłem B, i węzeł B jest połączony z węzłem C, to istnieje zwiększone prawdopodobieństwo, że węzeł A jest połączony z węzłem C. W języku sieci społecznych: przyjaciel twojego przyjaciela, prawdopodobnie jest też twoim przyjacielem. Z punktu widzenia topologii sieci, przechodniość oznacza występowanie trójkątów (trzy węzły połączone każdy z każdym). Współczynnik skupienia może być zdefiniowany jako:

$$C = \frac{3 * liczba\_trojkatow}{liczba\_polaczonych\_trojek}$$

gdzie liczba połączonych trójek oznacza ilość węzłów, które są połączone z dwoma innymi węzłami. Współczynnik C zawiera się w przedziale od 0 do 1. Powyższa definicja jest równoważna definicji:

$$C = \frac{6 * liczba\_trojkatow}{liczba\_sciezek\_o\_dlugosci\_2}$$

Alternatywną definicję wprowadzili Watts i Strogatz. Definiują oni lokalną wartość:

$$C_i = \frac{liczba\_wezlow\_polaczonych\_z\_i}{liczba\_trojek\_o\_centrum\_w\_i}$$

a następnie współczynnik skupienia dla całej sieci obliczają jako:

$$C = \frac{1}{n}\sum_i C_i$$

Dla wierzchołków, z których wychodzi zero lub jedno połączenie przyjmuje się $C_i = 0$.

3. Rozkład stopnia

Stopień węzła w sieci jest to liczba połączeń jakie z niego wychodzą. Definiuje się p_k jako ilość wierzchołków o k połączeniach podzieloną przez ilość wszystkich wierzchołków. Inaczej mówiąc p_k jest prawdopodobieństwem, że losowo wybrany wierzchołek jest stopnia k. Można sporządzić wykres p_k w zależności od k. Kształt tego wykresu nazywamy rozkładem stopnia (ang. degree distribution) sieci.

Alternatywnym sposobem przedstawiania rozkładu stopnia jest sporządzenie wykresu funkcji skumulowanej:

$$P_k = \sum_{k'=k}^{\infty} p_{k'}$$

która oznacza prawdopodobieństwo, że stopień jest większy bądź równy k.

4. Sprężystość

Sprężystość (ang. resilience) sieci jest to odporność na usuwanie niektórych węzłów. Można mówić o odporności na usuwanie losowych wierzchołków, bądź o odporności na usuwanie wierzchołków należących do pewnej klasy (np. o największej liczbie wychodzących połączeń). Parametr ten jest istotny dla sieci, w których istotna jest komunikacja, a tak jest w niemal wszystkich przypadkach.

5. „Mixing patterns”

Chodzi tutaj o prawdopodobieństwo wystąpienia połączenia między danymi dwoma typami wierzchołków. W większości sieci występują wierzchołki kilku typów, a prawdopodobieństwo wystąpienia połączenia między nimi często zależy od ich typów.

Na przykład w sieci określającej zależności między roślinami, roślinożercami i mięsożercami będzie dużo połączeń między pierwszymi dwoma klasami, ostatnimi dwoma klasami ale będzie bardzo mało połączeń między pierwszą i ostatnią klasą.

Do określenia tego typu zależności można użyć tzw. znormalizowanej „mixing matrix” (e). Niech E_{ij} oznacza liczbę połączeń między węzłami typów „i” oraz „j”, a E niech będzie macierzą o elementach E_{ij}. Wspomnianą macierz otrzymujemy jako:

$$e = \frac{E}{\|E\|}$$

gdzie $\|x\|$ oznacza sumę wszystkich elementów macierzy x. Elementy e_{ij} określają procent połączeń które są pomiędzy typami „i” oraz „j”.

Prawdopodobieństwo tego, że wierzchołek typu „i” jest połączony z typem „j” wynosi:

$$P(j\,|\,i) = \frac{e_{ij}}{\sum_j e_{ij}}$$

Taka definicja spełnia warunki normalizacji, tzn:

$$\sum_{ij} e_{ij} = 1, \qquad \sum_j P(j\,|\,i) = 1$$

Współczynnik „wymieszania” (ang. assortative mixing) można zdefiniować następująco:

$$Q = \frac{\sum_i P(i\,|\,i) - 1}{N - 1}$$

Taka definicja ma pożądane właściwości, tzn. wynosi 1 gdy połączenia występują tylko pomiędzy wierzchołkami tego samego typu, i 0 gdy połączenia są całkowicie losowe. Współczynnik ten ma jednak wady (np. dla asymetrycznej macierzy otrzymujemy dwie różne wartości).

Inną definicją współczynnika wymieszania (posiadającą również pożądane właściwości wspomniane wyżej) jest:

$$r = \frac{trace(e) - \|e^2\|}{1 - \|e^2\|}$$

5. *Grafy losowe*

Celem tego opracowania jest wprowadzenie w tematykę modelowania rzeczywistych sieci. Do tego celu stworzono dokładniejsze modele niż losowe grafy, jednak ze względu na historyczne pierwszeństwo zostaną one krótko opisane.

Najprostszym modelem, zdefiniowanym niezależnie przez dwa zespoły: Solomonoff i Rapoport oraz Erdos i Renyi, jest następujący model:

Mając dany zbiór „n” wierzchołków, rozpatrujemy każdą możliwą parę i tworzymy między nimi połączenie z prawdopodobieństwem „p”.

Prawdopodobieństwo, że tworzona sieć będzie mieć „m” krawędzi wynosi: $p^m(1-p)^{M-m}$, gdzie $M = 1/2n(n-1)$ jest maksymalną możliwą liczbą krawędzi. Rozkład stopnia jest rozkładem Poissona i wyraża się wzorem:

$$p_k = \binom{n}{k} p^k (1-p)^{n-k} \cong \frac{z^k e^{-z}}{k!}$$

gdzie $z = p(n-1)$ dla dużych n.

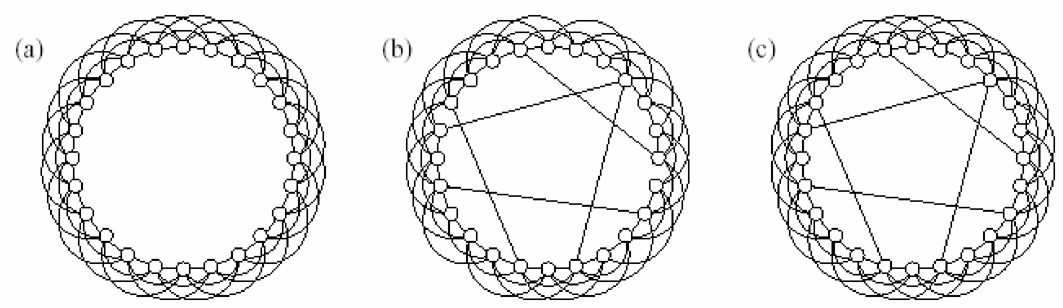
Grafy losowe dobrze symulują jedną z cech (niestety właściwie tylko jedną) sieci rzeczywistych – efekt small-world. Wartość parametru „l” zdefiniowaną wcześniej można obliczyć ze wzoru $l = \log n / \log z$, co pozwala powiedzieć, że występuje efekt small-world.

Istnieje wiele odmian tego modelu, mających na celu upodobnienie ich do rzeczywistych sieci.

6. Model small-world

Model ten został zaproponowany przez Watts i Strogatz.

Model small-world może być zbudowany z dowolnej regularnej sieci. Budowa polega na wybraniu małej liczby krawędzi, a następnie na losowej zmianie jednego wierzchołka takiej krawędzi. Nowe połączenia tworzone są tak, żeby nie występowały dwa takie same połączenia, ani połączenia samego do siebie. Krawędź jest przepinana za zadanym prawdopodobieństwem „p".



Na rysunku a) przedstawiona jest jednowymiarowa regularna sieć gdzie każdy wierzchołek połączony jest z dwoma sąsiadami. Na rys b) niektóre połączenia zostały losowo zmienione. Na rys c) połączenia zmienione są jak na rys b) ale oryginalne połączenia nie zostały usunięte.

Dla uproszczenia matematycznych obliczeń rezygnuje się z dwóch podanych warunków, tzn: połączenia mogą się dublować oraz mogę łączyć wierzchołki z samymi sobą, co uzyskuje się poprzez zmianę obu (a nie jednego) wierzchołków wybranej krawędzi.

Właściwości sieci tworzonych przy pomocy tego modelu są następujące (k oznacza ilość sąsiednich wierzchołków z jakimi jest połączony każdy wierzchołek w wejściowej regularnej sieci, natomiast L oznacza ilość wierzchołków):

1. Współczynnik skupienia

Dla obu wersji (tzn. uproszczonej i oryginalnej) wynosi on:

$$C = \frac{3(k-1)}{2(2k-1)}(1-p)^3$$

2. Rozkład stopnia

$$p_j = \sum_{n=0}^{\min(j-k,k)} \binom{k}{n} (1-p)^n p^{k-n} \frac{(pk)^{j-k-n}}{(j-k-n)!} e^{-pk}$$

dla $j \geq k$, i $p_j = 0$ dla j < k.

3. Średnia długość ścieżki

Jeśli $p \to 0$ czyli model staje się „large-world” średnia długość ścieżki wynosi $L/4k$ (czyli oczywiście nie spełnia założeń efektu small-world).

Dokładnego wzoru na średnią długość ścieżki na razie nie odkryto. Wykazano, że spełnia warunek:

$$l = \xi g(L/\xi)$$

gdzie ξ zależy od „p”, a g(x) jest nieznaną funkcją, która zależy tylko od rozmiaru systemu i topologii wejściowej sieci, ale nie zależy od „L”, ξ ani „p”.

Wykazano, że $\xi \sim p^{-\tau}$ dla małych „p”, przy czym τ jest stałą.

Najdokładniejszym oszacowaniem na średnią długość ścieżki jest (wyprowadzone z poprzedniego wzoru):

$$l = \frac{L}{k} f(Lkp)$$

przy czym funkcję f(x) dobrze przybliża funkcja:

$$f(x) = \frac{1}{2\sqrt{x^2 + 2x}} \tanh^{-1} \sqrt{\frac{x}{x+2}}$$

7. Sieci scale-free

Sieci posiadające własność scale-free tworzy się poprzez rozbudowę już istniejącej sieci (początkowo może to być jedynie kilka wierzchołków).

1. Model Price'a

W tym modelu sieć tworzona jest następująco: niech początkowa (skierowana) sieć posiada „n" wierzchołków. Niech p_k będzie ułamkiem wierzchołków, które posiadają k wejściowych połączeń, więc $\sum_k p_k = 1$. Każdy nowy wierzchołek dodany do sieci posiada ustaloną i stałą dla siebie (ale mogącą się różnić dla kolejnych dodawanych wierzchołków) liczbę połączeń wyjściowych (stopień wyjściowy). Średnia stopni wyjściowych wszystkich nowych wierzchołków jest stała – oznaczmy ją „m". Jest to też średnia stopni wejściowych wierzchołków w sieci: $\sum_k kp_k = m$.

Prawdopodobieństwo wystąpienia połączenia między nowym wierzchołkiem, a jednym ze starych jest proporcjonalne do stopnia wejściowego starego wierzchołka. Stwarza to problem od razu na starcie, ponieważ na początku mamy dany zbiór niepołączonych wierzchołków (tzn. stopień wejściowy i wyjściowy wynosi 0). Prostym i powszechnie stosowanym rozwiązaniem jest przyjęcie, że prawdopodobieństwo to jest proporcjonalne do $k + k_0$ gdzie k_0 jest pewną stałą, która zazwyczaj wynosi 1 (tzn. w matematycznych rozważaniach). Z tą poprawką prawdopodobieństwo to wynosi:

$$\frac{(k+1)p_k}{\sum_k (k+1)p_k} = \frac{(k+1)p_k}{m+1}$$

Jeśli oznaczymy przez $p_{k,n}$ zmianę wartości p_k przy sieci mającej „n" wierzchołków poprzez dodanie do niej nowego wierzchołka to zmiana ta wyniesie:

$$(n+1)p_{k,n+1} - np_{k,n} = [kp_{k-1,n} - (k+1)p_{k,n}]\frac{m}{m+1}\text{, dla } k \geq 1$$

lub

$$(n+1)p_{0,n+1} - np_{0,n} = 1 - p_{0,n}\frac{m}{m+1}\text{, dla } k = 0$$

Szukając stacjonarnego rozwiązania tzn. przyjmując $p_{k,n+1} = p_{k,n} = p_k$ otrzymuje się:

$$p_k = [kp_{k-1} - (k+1)p_k]m/(m+1)$$

$$p_0 = 1 - p_0 m/(m+1)$$

przekształcając to rozwiązanie dochodzi się do zależności:

$$p_k \sim k^{-(2+1/m)}$$

co się mniej więcej zgadza z wartościami zaobserwowanymi w rzeczywistych sieciach. Warto zwrócić uwagę na fakt, że rozwiązanie takie otrzymano bez pomocy komputerów (lata 60-te), na których można by symulować taki model i badać jego własności.

2. *Model BA (Barabasi i Albert)*

Model ten posiada tylko jedną, ale istotną różnicę odróżniającą go od poprzedniego modelu. W modelu BA rozpatruje się sieci nieskierowane, tzn. nie ma różnicy między stopniem wejściowym a wyjściowym (są takie same). Z jednej strony takie podejście ogranicza zastosowania (wiele sieci jest skierowanych), z drugiej omija się dzięki temu problem „pierwszego połączenia” rozwiązanego w poprzednim modelu poprzez wprowadzenie stałej k_0.

W modelu tym prawdopodobieństwo, że nowy wierzchołek połączy się ze starym jest analogiczne jak w poprzednim modelu:

$$\frac{kp_k}{\sum_k kp_k} = \frac{kp_k}{2m}$$

Zmiana wartości p_k przy dodaniu nowego wierzchołka wynosi – dla k > m:

$$p_k = \frac{1}{2}(k-1)p_{k-1} - \frac{1}{2}kp_k$$

a dla k = m (sposób budowy sieci powoduje, że k nie może być mniejsze od m):

$$p_k = 1 - \frac{1}{2}mp_m$$

Przekształcając rozwiązanie dostajmy:

$$p_k = \frac{2m(m+1)}{(k+2)(k+1)k}$$

co dla dużych wartości k daje zależność charakterystyczną dla sieci scale-free czyli $p_k \sim k^{-3}$.

3. Ogólniejsze modele BA

Następne modele bazują na modelu BA, ale pozwalają na budowanie sieci o zadanej funkcji $p_k \sim k^{-\alpha}$ (tzn ustalając pewne wartości wejściowe można sterować wartością parametru α a tym samym lepiej dostosować tworzony model do danych eksperymentalnych).

A) przyjmując, że prawdopodobieństwo połączenia się z wierzchołkiem o stopniu k jest proporcjonalne do $k + k_0$ ($-m < k_0 < \infty$) otrzymuje się:

$$\alpha = 3 + k_0 / m$$

B) przyjmując, że liczba krawędzi nowego wierzchołka zmienia się z czasem wraz ze wzrostem rozmiaru sieci według wzoru $m = n^a$ gdzie a jest pewną stałą, a prawdopodobieństwo połączenia z wierzchołkiem o stopniu k wynosi $k + Bn^a$ gdzie B jest stałe otrzymuje się:

$$\alpha = 2 + \frac{B(1+a)}{1-Ba}$$

4. Modele z „kopiowaniem wierzchołków"

Sieć budowana jest w następująco: dodawane są nowe wierzchołki oraz nowe skierowane krawędzie – które są łączone z losowymi wierzchołkami, bądź informacja do których wierzchołków mają być przyłączone „kopiowana" jest z innego, losowo wybranego wierzchołka. Jeśli wierzchołek, z którego kopiowana jest informacja posiada mniej krawędzi niż potrzeba, wybierany jest następny wierzchołek i z niego są dalej kopiowane połączenia, i tak aż wszystkie połączenia nowego wierzchołka będą „kogoś" wskazywały.

W takiej sieci $\alpha = (2-a)/(1-a)$ gdzie a jest współczynnikiem określającym proporcję między ilością krawędzi dodawanych losowo do ilości krawędzi kopiowanych z innego wierzchołka. Dla małych wartości a (od 0 do ½) wartość α mieści się w przedziale od 2 do 3.

8. *Model KE (Klemm-Eguiluz)*

Sieci stworzone przy pomocy tego modelu posiadają obie własności: scale-free oraz small-world. Wszystkie poprzednie modele scale-free nie cechowały się niskim współczynnikiem skupienia, przez co nie można ich było zaliczyć do kategorii sieci small-world.

W modelu KE każdy wierzchołek posiada dodatkową informację o jego stanie. Informacja ta określa czy wierzchołek jest aktywny czy nieaktywny.

Zaczynając od początkowej sieci o rozmiarze m, która jest kompletnie połączona i wszystkie wierzchołki są w stanie aktywnym, dodajemy nowe powtarzając poniższe 3 punkty:

- Dodanie nowego wierzchołka o m krawędziach

- Dla każdej z m krawędzi nowego wierzchołka wybiera się z prawdopodobieństwem „y” czy krawędź ma łączyć się z aktywnym, czy z nieaktywnym wierzchołkiem. Dla drugiego przypadku wierzchołek do którego ma być przyłączona krawędź jest wybierany na tej samej zasadzie jak w klasycznym modelu BA.

- Jeden z m aktywnych wierzchołków staje się nieaktywny. Prawdopodobieństwo, że do deaktywacji zostanie wybrany wierzchołek „i” wynosi:

  $$p_d = \frac{k_i^{-1}}{\sum_j k_j^{-1}}$$

  Nowy wierzchołek staje się aktywny.

Rozkład stopnia wynosi $p_k = 2m^2 k^{-3}$.

9. *Bibliografia*

Najważniejsze prace, z których korzystałem:

- „The structure and function of complex networks” – M. E. J. Newman
- “Statistical mechanics of complex networks” – Reka Albert, Albert-Laszlo Barabasi
- “Properties of nonuniform random graph models” – Satu Virtanen
- “Efficiency of scale-free networks: Error and attack tolerance” – Paolo Crucitti, Vito Latora, Massimo Marchiori, Andrea Rapisarda

Inne prace (z których niekoniecznie korzystałem, ale z którymi się zapoznałem przy okazji pisania tej pracy):

- “Random evolution in massive graph” – William Aiello, Fan Chung, Linyuan Lu
- “Scale-free characteristics of random networks: The topology of the world wide web” – Albert-Laszlo Barabasi, Reka Albert, Hawoong Jeong
- “Popularity based random graph models leading to a scale-free degree sequence” – Pierce G. Buckley, Deryk Osthus
- “Optimization in complex networks” – Ramon Ferrer I Cancho, Ricard V. Sole
- “The small-world phenomenon: An algorithmic perspective” – Jon Kleinberg
- “Mean-field theory for scale-free random networks” – Albert-Laszlo Barabasi, Reka Albert, Hawoong Jeong
- “Models of the small world” – M. E. J. Newman
- “Small worlds – The structure of social networks” – M. E. J. Newman
- “A random graph model for power law graphs” – William Aiello, Fan Chung, Linyuan Lu
- “A brief introduction to scale-free networks” - William J. Reed