LUAREA DECIZIILOR ÎN SOCIAL MEDIA

Prof.univ.dr. Emil Stan
Universitatea Hyperion

Abstract:
Decision Making Process is very difficult in social networks. This it hapens due to complexity and fuzziness of images taken to social media. For optimization of decision we propose two types of models which simulate the characteristics and properties of social networks. Using these models, the decision making process can be improved.
Keywords: graph, random graph, decision making, edge, node, model, simulation.

În mai 2011, Facebook avea 721 de milioane de utilizatori. Modelând această rețea de socializare printr-un graf orientat obținem o structură uriașă, reprezentată de un graf cu 721 milioane de noduri. Un **utilizator de Facebook**, la momentul acela, avea în medie 190 de prieteni. Aceasta înseamnă că toți utilizatorii de Facebook, luați în considerare, au avut un total de 68,5 miliarde de prietenii (modelate ca arce din graf). **Luarea deciziilor** în astfel de rețele uriașe este un proces dificil, care comportă multe riscuri. **Simularea deciziilor** în astfel de grafuri este însoțită de multe riscuri, dată fiind dificultatea realizării unor modele care să fotografieze realitatea cât mai veridic. În elaborarea unor modele care să permită **elaborarea deciziilor optime în social – media** se pun în mod natural câteva întrebări. Care sunt *procesele determinante* care stau la baza inițierii acestor prietenii, a acestor muchii din graf? Mai important, cum pot aceste aparent independente prietenii să formeze această rețea complexă de socializare?

În social media există mai multe rețele de socializare care conțin milioane de noduri și miliarde de arce. Aceste rețele complexe au în compunere miliarde de prietenii, motivele pentru existența fiecăreia, de cele mai multe ori, fiind obscure. *Complexitatea acestor rețele*, precum și dificultățile legate de analiza independentă a fiecăreia dintre aceste prietenii, creează necesitatea elaborării unor modele care generează, pe o scară mai mică, grafuri similare cu rețelele din lumea reală. În ipoteza că aceste modele pot simula proprietățile observate în rețelele de socializare reale, analiza rețelelor din lumea reală se reduce la măsurarea diferitelor *proprietăți ale rețelei simulate cu ajutorul modelelor*, cu costuri reduse și eficiență sporită. În plus, aceste modele permit o mai bună înțelegere a fenomenelor observate în rețelele din lumea reală prin furnizarea de explicații matematice concrete. De asemenea, modelele permit efectuarea unor experimente controlate în rețele suficient de mici, pe când rețelele reale, datorită dimensiunilor acestora, nu sunt pretabile unor astfel de simulări. Vom discuta câteva modele principale de rețea care pot fi grupate în două categorii, anume modelul *grafului aleator* și modelul *small world* (lume mică).

Aceste modele sunt concepute pentru a modela cu exactitate proprietățile realității observate în rețelele din lumea reală. Înainte de a ne ocupa în amănunt de aceste modele, vom discuta **proprietățile rețelelor de socializare**.

1. Proprietățile ale rețelelor de socializare

Rețelele de socializare din lumea reală se bucură de o serie de caracteristici comune. În proiectarea modelelor rețelelor ne propunem să elaborăm modele care pot descrie cu exactitate aceste rețele, evidențiind aceste caracteristici comune. Pentru a determina aceste caracteristici, o practică comună este de a identifica atributele lor și de a arăta că măsurătorile acestor atribute sunt consecvente cu realitatea. În special, trei atribute de rețea trebuie să prezinte măsurători consecvente cu rețelele din lumea reală: distribuția gradelor, coeficientul de clustering și lungimea drumului mediu. **Distribuția gradelor DGR** denotă cum gradele nodurilor sunt aranjate printr-o rețea. **Coeficientul de clustering CCR** măsoară tranzitivitatea unei rețele. De asemenea, **lungimea drumului mediu LDM** reprezintă media distanțelor (lungimea celui mai scurt drum) între perechile de noduri. Vom discuta cum deciziile legate de aceste trei atribute comportă evoluția rețelelor reale.

1.1 Distribuția gradelor DGR

Să luăm în considerație distribuția bogăției între membrii unei populații. Cele mai multe persoane fizice au o cantitate medie de capital, în timp ce câteva sunt considerate extrem de bogate. De fapt, observăm că mult mai multe persoane dispun de o valoare medie de capital decât cele mai bogate. Similar, să luăm în considerație distribuția populației într-o suprafață a unui județ. Puține zone metropolitane sunt dens populate, iar cele mai multe localități au o dimensiune medie a populației. În mass-media sociale, observăm același fenomen regulat atunci când măsurăm popularitatea sau a spectaculozitatea pentru entități. De exemplu, multe site-uri sunt vizitate aproximativ de o mie de ori pe lună, în timp ce puține sunt vizitate mai mult de un milion de ori pe zi. Cei mai mulți utilizatori de social media sunt activi pe câteva site-uri, în timp ce câteva persoane sunt active pe sute de site-uri.

Produsele cele mai vândute sunt mult mai modeste ca preț de vânzare, comparativ cu cele scumpe.

De asemenea, există multe persoane cu câțiva prieteni și o mână de utilizatorii cu mii de prieteni. Ultima observație este legată direct de **gradul nodului în social media**. Gradul unui nod în social media reprezintă de multe ori numărul de prieteni pe care un individ îl are. Astfel, distribuția numărului de prieteni denotă distribuția gradelor de rețea. Se pare că în toate observațiile efectuate, distribuția valorilor urmează o *distribuție de tip Pareto*. **Diagrama Pareto a fost numită după Vilfredo Pareto, un economist italian din secolul al XIX-lea, care a realizat un studiu despre bogăție și sărăcie în Europa, la începutul secolului al XX-lea. Acesta a descoperit că cei bogați sunt puțini la număr, iar cei săraci sunt majoritari. Acest principiu se bazează pe distribuția inegală a lucrurilor în univers. Aceasta este legea „puținului semnificativ în fața abundentului lipsit de substanță.”**

Principiul Pareto este legea nescrisă, conform căreia ,,20% dintre probleme au 80% din impact.” 20% din probleme reprezintă ,,puținul semnificativ”, iar restul alcătuiesc ,,abundentul lipsit de substanță.”

Din punctul de vedere al calității, diagrama a fost introdusă de profesorul J.M. Juran, pentru a face distincția între:

- Problemele esențiale, care sunt puține la număr, dar care au rezultate importante și
- Problemele secundare, care sunt multe la număr, dar care au puține rezultate.

Diagrama Pareto este un instrument folosit atunci când cercetăm un proces ce oferă informații împărțite pe categorii, pentru a putea număra repetiția unei anumite categorii. Informațiile sunt aranjate în ordine și, de aceea, problemele esențiale pot fi identificate și corectate mai întâi.

De exemplu, să notăm cu k gradul unui nod (de exemplu, numărul de prieteni pe care

un individ îi are). Fie p_k fracțiunea de persoane cu gradul k, adică, frecvența de observare k raportată la numărul total de observații. Apoi, în distribuția dată de legea putere

p_k = a*k**(-b);(1)

în care b este exponentul puterii, iar a este o constantă a rețelei de socializare.

Luând logaritmul ambelor părți ale ecuației (1), avem

ln p_k = (-b) ln k + ln a (2)

Ecuația 2 arată că reprezentând grafic logaritmul unei distribuții de putere obținem o linie dreaptă cu panta *-b*. Acest grafic relevă, de asemenea, o metodologie pentru a verifica dacă o rețea prezintă o distribuție DGR dată de o lege putere [6].

În reprezentarea grafică, axa Ox reprezintă logaritmul gradului și axa Oy reprezintă logaritmul fracțiunii persoanelor cu acest grad (adică, ln p_k). Obținerea unei linii demonstrează observarea unei tendințe liniare, ceea ce conduce la acceptarea modelului distribuției legii putere.

Etapele care sunt specifice validării acestui model sunt [1]

1. Se determină măsura popularității fiecărui individ pentru întreaga rețea.

De exemplu, putem lua numărul de prieteni într-o rețea socială ca o măsură a popularității. Vom nota valoarea măsurată cu k.

2. Se determină p_k, fracțiunea de persoane cu popularitatea k.

3. Se trasează un grafic log-log, în cazul acesta. axa Ox reprezintă ln k și axa Oy

reprezintă ln p_k.

4. În cazul în care există o distribuție DGR lege putere, ar trebui să se respecte o linie dreaptă.

Modelul prezentat a fost aplicat pentru realizarea graficelor log-log pentru numărul de prieteni pe rețelele din lumea reală. În toate rețelele s-a constatat un trend liniar ce denotă o distribuție de lege putere.

Rețele de distribuție care prezintă distribuția gradelor cu legea putere drept sunt adesea numite rețele independente de scară. Având în vedere că majoritatea rețelelor sociale sunt independente de scară, suntem interesați de modele care pot genera rețele sintetice cu odistribuție a gradelor după legea putere.

1.2 Coeficientul de clustering CCR

În rețelele sociale reale, prietenii sunt foarte tranzitivi. Cu alte cuvinte, prietenii unei persoane sunt adesea prieteni unul cu altul. Aceste prietenii formează triade de prietenii care sunt frecvent observate în rețelele sociale. Aceste triade conduc la rețele cu medie [locală] mare a coeficienților de clustering CCR. În mai 2011, Facebook a avut o medie a coeficienților de clustering de 0,5 pentru persoanele fizice care au avut doi prieteni; gradul lor a fost 2 [2]. Acest lucru indică faptul că 50% dintre toți utilizatorii cu doi prieteni aveau caracteristica că aceștia doi erau de asemenea prieteni. Prezentăm în continuare media coeficienților de clustering CCR pentru mai multe rețele sociale din lumea reală și web:

Web ……………………………………... 0,08
Facebook (pentru 100 prieteni) ………………0,14
Flikr……………………………………...0,31
LiveJournal……………………………………0,33
Orkut……………………………………..0,17
YouTube……………………………………….0,13

1.3 Lungimea drumului mediu LDM

În rețelele reale, oricare doi membri ai rețelei sunt, de obicei, conectați prin intermediul unui celui mai scurt drum. Cu alte cuvinte, durata lungimea drumului mediu LDM este mică. Acest lucru este cunoscut sub numele de fenomenul de lume mică. În bine-cunoscute experiment *lume mică* efectuat în 1960 de către Stanley Milgram [2]. Milgram a arătat că oamenii din întreaga lume sunt legați între ei printr-o cale de cel mult șase persoane (de exemplu, cele șase grade de separare). În mod similar, observăm mici lungimi ale drumului mediu în rețelele sociale. De exemplu, în mai 2011, lungimea drumului mediu LDM între indivizii din Facebook a fost pe graful atașat de 4,7. Această medie a fost de numai 4,3 pentru persoanele fizice din SUA în aceeași perioadă de timp, [3]. Prezentăm mai jos lungimea drumului mediu LDM pentru lumea reală a rețelelor sociale și web.

Web ……………………………………... 16,1
Facebook (pentru 100 prieteni) ………………4,67
Flikr……………………………………...5,67
LiveJournal……………………………………5,88
Orkut……………………………………..4,25
YouTube……………………………………….5,11

Aceste trei proprietăți – distribuția gradelor după legea putere, coeficientul de clustering CCR mare și lungimea drumului mediu LDM scurtă sunt în mod constant observate în rețelele din lumea reală. De aceea, modelele proiectate vor fi modele bazate pe ipoteze simple cu privire la modul de formare al prieteniilor, astfel ca aceste modele să genereze rețele independente de scară, cu coeficientul de clustering CCR mare și lungimea drumului mediu mică. Vom începe cu cel mai simplu model de rețea, modelul grafului aleator.

2. Graful aleator

Ipoteza care stă la baza proiectării modelului **graful aleator** se referă la modul în care pot fi formate prieteniile:

Arcele (de exemplu, prieteniile) între noduri (de exemplu, persoane fizice) sunt formate

aleatoriu.

De fapt, prieteniile în rețelele de socializare reale sunt departe de a fi aleatoare. Prin asumarea ipotezei prieteniilor aleatoare, simplificăm procesul de formare al prieteniilor în rețelele din lumea reală, în speranța că aceste prietenii aleatorii, în cele din urmă, vor conduce la crearea de rețele care prezintă caracteristici comune celor observate în lumea reală a rețelelor de socializare. Formal, putem presupune că pentru un graf cu un număr fix de noduri n, oricare dintre cele C_n^2 muchii pot fi formate în mod independent, cu probabilitatea p.

Acest tip de graf se numește un graf aleatoriu, iar modelul obținut îl notăm G (n, p). Acest model a fost propus pentru prima dată, în mod independent de Edgar Gilbert [3] și

Solomonoff și Rapoport [2]. Un alt mod de a genera aleatoriu grafice este să se presupună că atât numărul de noduri n și cât și numărul de muchii m sunt fixe. Cu toate acestea, avem nevoie pentru a determina care m muchii sunt selectate din setul de C_n^2 posibile. Fie Ω mulțimea de grafuri cu n noduri și m muchii. Pentru a genera un graf aleator, putem selecta unul dintre grafurile din Ω. Numărul de grafuri cu n noduri și m muchii (adică K, cardinalul mulțimii Ω) este combinări de m luate câte C_n^2. Probabilitatea de selecție a unui anumit graf este 1/K.

Se poate gândi probabilitatea de selecție a unui graf făcând o analogie la p, probabilitatea de selecție a unei muchii în G (n, p).

Acest al doilea model a fost introdus de Paul Erdos și Alfred Renyi [3] și se notează ca modelul G (n; m). La limită, ambele modele acționeze în mod similar.

Numărul mediu de muchii din G (n, p) este C_n^2p. Acum, dacă stabilim $C_n^2p= m$, la

limită, ambele modele acționează la fel, deoarece acestea conțin același număr de muchii. Să observăm că modelul G (n, m), conține un număr fix de muchii, iar al doilea model G (n, p) este posibil să nu conțină nici-o muchie sau toate posibilele muchii.

Matematic, modelul G (n, p) este aproape întotdeauna simplu de analizat; prin urmare, restul acestei secțiuni se ocupă cu determinarea proprietăților acestui model. Notăm că există multe grafuri cu noduri n și m muchii (de exemplu, generate de G (n, m)). Același argument este valabil și pentru G (n, p), multe grafuri pot fi generate de model. Prin urmare, atunci când se măsoară proprietățile grafurilor aleatorii, măsurile sunt calculate pe toate graficele care pot fi generate de model și apoi se află media. Acest lucru este deosebit de util atunci când suntem interesați în medie, și nu specific, de comportamentul grafurilor mari.

În G (n, p), numărul de muchii nu este fixat; prin urmare, vom examina mai întâi unele proprietăți matematice privind numărul preconizat de muchii care sunt conectate la un nod, numărul mediu de muchii observate în graf și probabilitatea de a observa m muchii într-un graf aleatoriu generat de procesul G (n, p).

O primă astfel de proprietate ne arată că numărul mediu de muchii conectate la un nod (gradul mediu) în G (n, p) este *(n - 1) p*.

Într-adevăr, un nod poate fi conectat la cel mai mult *n - 1* noduri (prin *n-1* muchii). Toate muchiile sunt selectate, în mod independent, cu probabilitatea *p*. Prin urmare, în medie

(n - 1) p dintre ele sunt selectate. În literatura de specialitate, gradul mediu este adesea denumit folosind notația c sau k. Din moment ce am folosi frecvent k pentru a indica valorile gradului, folosim c pentru a indica gradul mediu dintr-un graf aleator

$c = (n - 1) p$; (3)

sau echivalent,

$p= c / (n - 1)$ (4)

O a doua proprietate se referă la numărul mediu de muchii din G (n, p) care este C_n^2p.

Urmând aceeași linie de argumentare, pentru că muchiile sunt selectate independent și deoarece avem un maxim de C_n^2 muchii, numărul mediu de muchii este C_n^2p.

O ultimă proprietate pe care o scoatem în evidență este că într-un graf generat de modelul G (n; p), probabilitatea de observare a m muchii este distribuția binomială

$P(m) = C_m^r p^{**}m (1-p)^{**}q$ (5)

Unde $r = C_n^2$, iar $q = r-m$.

Având în vedere aceste afirmații de bază, vom analiza cum grafurile aleatoare evoluează atunci când adăugăm noi muchii.

2.1 Evoluția grafurilor aleatoare

În grafurile aleatoare, atunci când nodurile evoluează și formează legături, după un timp, o mare parte dintre noduri vor fi conectate între ele, în sensul că între acestea există o cale. Această fracțiune mare din mulțimea nodurilor face o componentă conectată, de obicei numită *componenta gigant* sau cea mai mare componentă conectată CC. Putem dirija comportamentul modelului graf aleator prin selectarea corespunzătoare a valorii p.

In G (n; p), atunci când p = 0, dimensiunea celei mai mari componente conectate este 0 (nici există nici măcar două perechi de noduri care sunt conectate), iar când p = 1, dimensiunea acesteia este n (toate perechile sunt conectate).

Prezentăm în continuare dimensiunea celei mai mari componente conectate (DCC) pentru grafuri aleatoare cu 10 noduri și diferite valori pentru p. Tabelul va conține, de asemenea, informații cu privire la gradul mediu C, dimensiunea diametrului DD, dimensiunea celei mai mari componente DCC și lungimea drumului mediu LDM a grafului aleator.

După cum se arată în tabelul de mai jos, când p devine mai mare, graful devine mai dens. Când p este foarte mic, se constată următoarele:

p	C	DD	DCC	LDM
0	0	0	0	0
0,055	0,8	2	4	1,5
0,11	1	6	7	2,66
1	9	1	10	1

Din acest tabel putem observa:

1. Nici-o componentă gigant nu se observă în grafic.
2. Sunt formate mici componente conectate izolate.

3. Diametrul este mic, deoarece toate nodurile sunt în componente izolate, în care există puține noduri.

Când p devine mai mare, se întâmplă următoarele:
1. O componentă gigant începe să apară.
2. Componentele izolate devin conectate.
3. Valorile diametrului cresc.

In acest moment, nodurile sunt conectate între ele prin căi lungi (vezi p = 0,11 în tabel). Atunci când p continuă să crească, graficul aleatoriu își schimbă din nou proprietățile. Pentru valori mai mari, diametrul începe să scadă, nodurile sunt conectate între ele prin intermediul diferitelor drumuri (care sunt din ce mai scurte). Punctul în care valoarea diametrului începe să scadă în graful aleator se numește *tranziție de fază*. La punctul de tranziție de fază sunt observate două fenomene:

1. Componenta gigant, care abia începuse să apară, începe să crească.
2. Diametrul, care a atins valoarea sa maximă, începe în scădere.

Este dovedit faptul că în grafurile aleatoare tranziție de fază se produce atunci când c = 1; adică, p = 1 / (n - 1).

Până acum am discutat despre generarea și evoluția grafurilor aleatoare. Cu toate acestea, avem nevoie, de asemenea, să analizăm modul în care grafurile aleatorii efectuează imitarea proprietăților pe care le au rețelele din lumea reală. Multe exemple indică faptul că grafurile aleatorii pot modela lungime drumului mediu LDM din lumea reală a rețelelor exact, dar nu reușesc să genereze o distribuție realistă a gradelor DGR sau coeficienții de clustering CCR preciși. Vom discuta aceste proprietăți în continuare.

2.2 Proprietățile grafurilor aleatoare

Când calculăm distribuție gradelor, estimăm probabilitatea de observare

P (d_v = d) pentru nodul v. Se demonstrează că pentru un graf generat de G (n, p), nodul v are gradul d, d< n-1, cu probabilitatea dată de o distribuție binomială a gradelor.

Aceasta presupune că n este fixat. Putem generaliza acest rezultat prin calcularea

distribuției gradelor graficelor aleatoare la limită (de exemplu, n→ ∞). În această situație

P (d_v = d) =(e**(-c))(c**d)/d! (6)

care este de fapt distribuția Poisson cu media c. Astfel, la limită,

grafurile aleatoare au distribuția Poisson, care diferă de distribuția putere, lege observată în rețelele reale.

Într-un graf aleator generat de G (n; p), coeficientul de clustering CCR mediu

pentru nodul v este p.

C (v) = (Numărul de perechi conectate ale vecinilor lui v) / (Numărul de perechi ale vecinilor lui v)

Nodul v poate avea diferite grade în funcție de muchiile care se formează random. Astfel, putem calcula valoarea medie pentru C (v):

M (C(v)) = ∑ **M**(C(v) | dv = d) P(dv = d) (7)

Suma de mai sus se face după toate nodurile rețelei.

Pentru un graf aleator, avem

M(C(v) | dv = d) = p (8)

Substituind ecuația 8 în ecuația 6, vom obține

M (C (v)) = p (9)

De asemenea, coeficientul de clustering general CCR al unui graf aleator generat de G (n; p) este p. Așa cum am stabilit coeficientul de clustering general CCR al unui graf aleatoriu definește probabilitatea ca doi vecini al aceluiași nod să fie conectați. În grafuri aleatorii, pentru oricare două noduri, această probabilitate este aceeași și este egală cu probabilitatea p, (probabilitatea de a obține două noduri conectate). Să observăm că în grafuri aleatoare, coeficientul de clustering local este echivalent cu coeficientul de clustering global.

În grafice aleatoare, coeficientul de clustering este egal cu probabilitatea

p; prin urmare, prin selectarea adecvată a lui p, putem genera rețele cu

un coeficient de clustering mare. Să observăm că selectarea unui p mare nu este de dorit

deoarece acest lucru va genera un graf foarte dens, care este nerealist, deoarece

în lumea reală, rețelele sunt adesea rare. Astfel, grafurile aleatoare sunt considerate, în general, incapabile de a genera rețele cu CCR mari, fără a compromite alte proprietăți necesare.

Lungimea drumului mediu LDM într-un grafic aleator este

LDM= (ln V)/(ln c) (10)

In grafurile aleatoare, dimensiunea diametrului mediu D tinde spre LDM la limită. Folosind acest fapt, avem

C**D = V (11)

2.3 Modelarea rețelelor de socializare cu grafuri aleatoare

Având în vedere o anumită rețea de socializare din lumea reală, putem simula proprietățile acesteia cu ajutorul unui model bazat pe un graf aleatoriu. Putem calcula gradul mediu c în rețeaua dată. De la c, se poate trece la probabilitatea de conexiune p. Aceasta poate fi poate fi calculată (p = c / (n-1). Utilizarea p și a numărului de noduri din rețeaua dată n permite elaborarea unui model de graf aleator G (n; p). Acesta va simula caracteristicile rețelei reale. Grafurile aleatoare modelează bine lungimile drumului mediu LDM; cu toate acestea, atunci când se analizează tranzitivitatea, modelul grafic aleatoriu subestimează drastic coeficientul de clustering CCR. Pentru a rezolva această problemă, vom studia *modelul lumea mică*.

3. **M*odelul lumea mică***

Ipoteza de bază din modelul graf aleator este următoarea: conexiunile în rețelele din lumea reală sunt formate la întâmplare. Deși nerealist, graful aleator poate modela lungimea drumului mediu LDM în rețelele reale în mod corespunzător, dar subestimează coeficientul de clustering. Pentru a diminua această problemă, Duncan J. Watts și Steven Strogatz în 1997 au propus modelul lumea mică.

În interacțiunile reale, multe persoane au un număr limitat de și de multe ori un număr fix de conexiuni. Persoanele pot conecta cu părinții lor, cu frați, surori, bunici și profesori etc. Astfel, în loc de asumarea unor conexiuni aleatoare, așa cum am făcut în modelele grafurilor aleatoare, se poate presupune un model în care oamenii au același număr de vecini (prieteni). Și acest nou model este nerealist; cu toate acestea, modelează mai exact coeficientul de clustering al rețelelor din lumea reală. În ceea ce privește teoria grafurilor, această ipoteză este echivalentă cu *generarea unei rețele regulate,* mai precis o latice inelară, ceea presupune încorporarea persoanelor fizice într-o rețea uniformă. Un inel este un caz special de rețea în care există un anumit tipar pentru modul în care nodurile sunt conectate unul la celălalt. În particular, într-un inel regulat cu c grade, noduri sunt conectate la c/2 precedenți și în c / 2 vecini care urmează.

Rețelele regulate pot modela tranzitivitatea bine; cu toate acestea, LDM este prea mare. Mai mult decât atât, coeficientul de clustering are valoarea ¾ care este fixă și nu poate fi armonizată cu valorile coeficienților din rețelele din lumea reală. Pentru a depăși aceste probleme, modelul propus introduce un parametru β care controlează aleatoriul în model. Modelul începe cu o latice regulată regulat și se demarează un proces de adăugare de muchii pe baza parametrului β. Atunci când β este o, avem o latice regulată, iar când β este 1 avem un graf aleatoriu.

Procedura creează noi muchii printr-un proces numit reconectare. Reconectarea înlocuiește o muchie existentă între nodurile v_i și v_j cu una care nu exista între v_i și v_k cu probabilitatea β. Cu alte cuvinte, o muchie este deconectată de la unul din punctele finale v_j și conectată la un nou punct final v_k. Nod v_k este selectat uniform.

Rețeaua generată folosind această procedură are câteva proprietăți interesante.

În funcție de valoarea parametrului β, aceasta poate avea un CCR mare și, de asemenea, o LDM scurtă. Distribuția gradelor, cu toate acestea, încă nu se potrivește cu cea a rețelelor de lumea reală.

Proprietăți ale modelului lumii mici

Distribuția gradelor pentru modelul lumii mici este destul de similară cu distribuția gradelor la distribuția Poisson pentru grafurile aleatoare. În practică, în graficul generat de modelul lumii mici, cele mai multe noduri au grade similare din cauza laticii de bază. În contrast, în rețelele de lumea reală, gradele sunt distribuite pe baza distribuției legii de putere, în care caz cele mai multe noduri au grade mici și numai câteva au grade mari.

Coeficientul de clustering CCR pentru o latice regulată este 3 (c-2) / 4(c-1) și pentru un graf aleator este p = c / (n-1).

CCR pentru o rețea a lumii mici este o valoare între aceste două, în funcție de β. Notând CCR pentru laticea regulată C(0) și CCR pentru un model lume mică, cu β= p cu C(p), atunci relația dintre cele două valori pot fi calculată analitic; acesta a fost dovedit că este

C (p) = ((1 - p)** 3)C (0) (12)

Intuiția din spatele acestui raport este că deoarece CCR enumeră numărul de triade închise într-un graf, suntem interesați în triade care au mai rămas conectate după procesul de refacere. Pentru ca o triadă să rămână conectată, este necesar ca toate cele trei muchii să nu fie reconectate, cu probabilitatea (1 - p). Deoarece procesul se efectuează în mod independent pentru fiecare muchie, probabilitatea de observare a triadei este de (1-p)**3 ori probabilitatea de observare într-o rețea regulată. Să observăm că trebuie, de asemenea, să se ia în considerare noile triade care sunt formate prin procesul de refacere a conexiunilor; deoarece probabilitatea este nominală putem să o neglijăm.

Valoarea lui C (p) rămâne ridicată până când p ajunge la 0,1 (10% din muchii reconectate) și apoi scade rapid până la o valoare aproape de zero când p se apropie de 1. Întrucât CCR trebuie să fie mare este de preferat ca β≤0,1.

Aceeași procedură se poate face pentru LDM. Lungimea drumului mediu într-o rețea regulată este n/(2c), să-l notăm L(0). Notăm L (p), ca lungimea drumului mediu cale pentru o lume mică unde β = p. Spre deosebire de C (p), nici o formulă analitică pentru compararea

L (p) la L (0) există; însă, relația poate fi calculată empiric pentru diferite valorile ale lui p.

Se obține că LDM scade mai repede decât CCR și devine stabil când în jur de 1% din muchii sunt reconectate. Întrucât avem nevoie de mici LDM în grafurile generate este de preferat β≥0,01.

4. Concluzii

În lucrarea s-au propus câteva modele care să ușureze **procesul de luare a deciziilor** în rețelele social media. S-au introdus câteva caracteristici și proprietăți ale acestor modele pentru a putea compara calitatea acestor modele și în același timp pentru a analiza posibilitatea de a obține un grad de încredere înalt în judecățile efectuate pe baza modelelor de simulare create. Modelul de dorit pentru o rețea din lumea reală ar trebui să genereze grafuri cu coeficient de clustering CCR mare, concomitent cu lungimea drumului mediu LDM scurtă. Având în vedere o rețea din lumea reală în care gradul mediu C și CCR sunt date, ne-am stabilit C (p) = C și determinăm un β optim pe baza formulei anterioare. Având în vedere c și n (dimensiunea rețelei lumea reala), am putea astfel simula modelul lumea mică.

Desigur, modelul lumea mică este capabil de a genera o distribuție realistă a gradelor în grafurile simulate, permițând **luarea unor decizii** pe bază științifică, nu numai pe baza experienței sau intuiției.

Bibliografie

1. Newman Mark, *The structure and dynamics of network*, Princeton Press, Princeton, NJ, 2009
2. Sobel M., Becker M.P., (Editori), *Sociological Methodology*, Basil Blackwell, London, 2010
3. Stan Emil, *Conceperea și Proiectarea Produsului Jurnalistic*, Editura Victor, București, 2011
4. Stan Emil, *Tehnologia Informației*, Kullusis Press, București, 2010
5. Wolton Dominique, *Sauver la Communication*, Flamarion, Paris, 2009
6. Zhao S., *Do Internet users have more social ties?* J. of Comp. Mediated Communication, **11,** (3), 2009