Rajko Nenadov

Heuristike za rešavanje NP-kompletnih problema na scale-free mrežama

Ovaj rad se bavi konstruisanjem heuristike za rešavanje dva NP-kompletna problema na scale-free mrežama. Posmatrani su problemi Hamiltonove konture i klike. U problemu Hamiltonove konture cilj je bio pronaći neku pravilnost u karakteristikama čvorova koji formiraju konturu, dok je kod problema klike fokus stavljen na nalaženje onih čvorova koji eventualno pripadaju rešenju, pa su onda posmatrani samo oni. Na osnovu dobijenih rezultata napravljene su "branch-and-bound" heuristike. Dobijena heuristika za Hamiltonovu konturu je primenljiva na mrežama malih dimenzija (35 čvorova), dok je heuristika za maksimalnu kliku primenljiva na mrežama većih dimenzija (4000 čvorova).

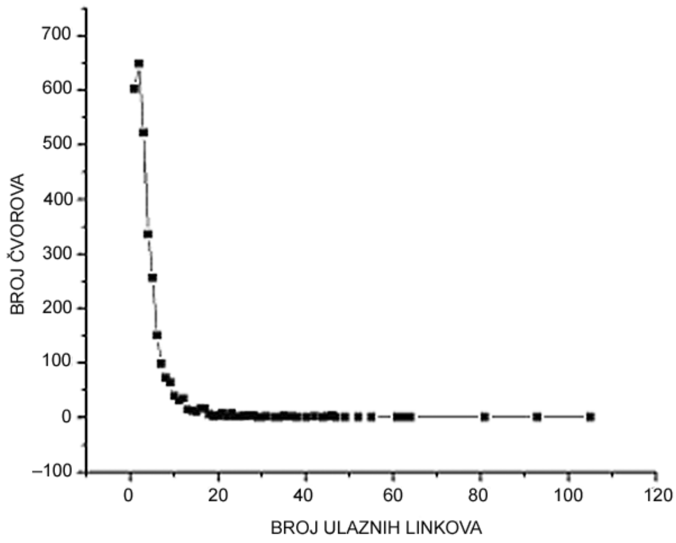
Uvod

NP-kompletni problemi su oni čije rešenje nije moguće naći u polinomnom vremenu. Ukoliko znamo nešto o prostoru nad kojim treba da se reši NP-problem, možemo iskoristiti karakteristike tog prostora da brže dođemo do rešenja. Konkretno u ovom slučaju se radi o mrežama koje imaju karakterističnu topologiju. Posmatrani su problemi nalaženja Hamiltonove konture i problem klike. Problem Hamiltonove konture jeste da se nađe put u mreži, tako da se na tom putu svi čvorovi posete tačno jedanput i da se iz poslednjeg čvora vrati u prvi, dok kod poluhamiltonove konture vraćanje u početni čvor nije potrebno. Problem klike se sastoji u nalaženju kompletnog podgrafa proizvoljne veličine.

Realne mreže se mogu modelirati BA modelom scale-free mreža (u daljem tekstu će se ponekad umesto reči mreža koristiti i reč graf). Autor modela je mađarski fizičar Barabaši (Barabasi 1999). Realne scale-free mreže su na primer: internet na fizičkom nivou (računari i ruteri su čvorovi, veza između dva čvora postoji samo ako su oni direktno povezani), internet na logičkom nivou (WWW – čvorovi su HTML stranice, veze su HTML linkovi), socijalne mreže, mreža citata, mreža glumaca. Do pre par godina smatralo se da model *random mreža* može u potpunosti opisati na-

Rajko Nenadov (1987), Novi Sad, Ilije Birčanina 57a, učenik 4. razreda Gimnazije "Jovan Jovanović Zmaj" u Novom Sadu

MENTOR: Miloš Savić, student Prirodno-matematičkog fakulteta u Novom Sadu, Departman za matematiku i informatiku



Slika 1.
Distribucija broja čvorova po broju linkova kod BA modela scale-free mreža

Figure 1.
Distribution of number of nodes by the number of limks at BA model of scale-free networks.

vedene realne mreže. Distribucija broja čvorova po broju linkova u modelu random mreža je Gausovska. Ovaj modelom se na primer može opisati mreža auto puteva. Empirijski je utvrđeno da WWW mrežu ne možemo opisati modelom random mreža. Zapravo, posmatrajući realne mreže, u kojoj postoje preferencijalni čvorovi, može se zaključiti da postoji par čvorova koji su povezani sa većinom ostalih, dok su ostali povezani sa svega jednim ili dva čvora (slika 1). Upravo to je glavna odlika mreža koje se opisuju scale-free modelom. Preferencijalnost označava verovatnoću da dati čvor bude povezan sa čvorom koji se kao novi ubacuje u mrežu. Čvorovi koji su preferencijalni u jednom trenutku, težiće da ostanu preferencijalni i u budućnosti, tj. prate princip "bogat postaje bogatiji". Dinamičnost znači da čvorovi mogu da se ubacuju i izbacuju iz mreže. Preferencijalnost za dati čvor i po Barabasijevom scale-free modelu se izračunava kao količnik broja veza kojima je jedan kraj u čvoru i i ukupnog broja veza. Postoje i drugi načini modeliranja scale-free mreža.

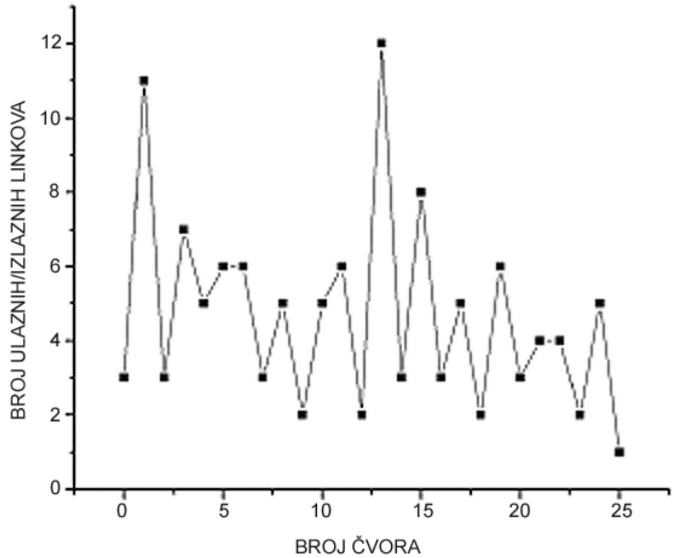
Karakteristike mreža

Za scale-free mreže su uočene tri bitne karakterisike. To su small-world koeficijent, klastering koeficijent, i distribucija broja čvorova po broju linkova. Small-world koeficijent za čvor i se računa kao aritmetička sredina minimalnih rastojanja od čvora i do svih ostalih, i označava koliko je prosečno "hopova", to jest presedanja, potrebno da bi se došlo od tog čvora do bilo kog drugog. Klastering koeficijent za čvor se računa kao količnik broja veza kojima je jedan kraj u datom čvoru i ukupnog broja linkova. Klastering koeficijent za ceo graf numerički pokazuje koliko on odstupa od kompletnog grafa.

Heuristika za traženje Hamiltonove i polu-Hamiltonove konture

Vreme za traženje Hamiltonove konture raste eksponencijalno sa porastom broja čvorova. Optimizaciju rešenja treba tražiti na sledeći način: za dati čvor i (u kome smo trenutno) se određuje koji će nam njegovi susedi možda dati bolje rešenje i samo oni se posmatraju. Drugim rečima, tražimo neku "branch-and-bound" heuristiku. Pored Hamiltonove, obratićemo pažnju i na traženje poluhamiltonove konture. Ideja koja se koristi za traženje heuristike je sledeća: na dovoljno velikom broju uzoraka nađemo sve Hamiltonove i poluhamiltonove konture (ukoliko na datoj mreži ne postoji ni jedna od ove dve konture, onda posmatramo najduži mogući put u kome se svaki čvor pojavljuje samo jednom) i potom gledamo kako se karakteristike čvorova na tim putevima menjaju. Računarskom simulacijom je ustanovljeno je da se preko distribucije broja čvorova po broju linkova može doći do heuristike.

Čvorovi su posmatrani onim redom kojim se pojavljuju u konturi i za svaki čvor je gledan njegov broj veza. Utvrđeno je da se takav niz može grubo aproksimirati testerastom ("cik-cak") funkcijom (slika 2.) Taj rezultat je poslužio kao glavna ideja za nalaženje heuristike, pošto znamo kako treba da se menjaju karakteristike čvorova na putu koji mi konstruišemo. Problem na koji se nailazi je odstupanje u pravilnosti. Na grafikonu na slici 2 vidi se da se iz čvora 9 prelazi u čvor 10 koji ima veći broj linkova od njega, ali se onda takođe i iz čvora 10 prelazi u čvor 11 koji isto ima veći broj linkova od njega. Posmatrajući sve Hamiltonove konture u mreži



Slika 2.
Broj veza čvorova u konturi redom

Figure 2.
Number of relations between the knots in conture, respectively.

dolazi se do zaključka da se ponekad mora napraviti ovakvo odstupanje od šablona. Dozvoljeno odstupanje će nam biti jedan parametar u heuristici.

Drugi problem na koji se nailazi je što ne znamo iz kog čvora treba da počnemo pretragu. Može se intuitivno zaključiti da je nepotrebno kretati u pretragu iz čvora koji ima veliki broj veza, tako da ćemo pretragu ograničiti na polaženje samo iz prvih k čvorova sa najmanjim brojem veza; tako dolazimo do drugog parametra. U slučaju da je broj dozvoljenih odstupanja jednak broju čvorova iz kojih krećemo pretragu, i to jednako broju čvorova mreže, dobijamo običnu pretragu bez odsecanja. Postoji šansa da se heuristika poboljša, to jest da se za datu mrežu utvrdi optimalna vrednost parametara tako da se pretraga maksimalno skrati, ali da rešenje bude tačno. Uz dobro izabrane parametre, pretraga bi mogla da se ubrza bar za 50%, a verovatno i više (to je broj dobijen iz pretraživanja na malim mrežama, što nije dovoljno za neko tačnije određivanje parametara).

Pseudo-kod programa koji traži Hamiltonovu konturu:

```
NadjiHamiltona(k)
   i=0;
   while not PronasaoPut do
        Pretraga(i,0,gore,0)
        i=i+1
end;

Pretraga(trenutniCvor,dubina,smer,odstup)
   if dubinaukupan_broj_cvorova then PronasaoPut
   else
       za sve cvorove sused za koje je broj ulaznih
       linkova veci, jednak, odnosno manji (zavisno
       od smera)od broja ulaznih linkova trenutnog
       cvora radi Pretraga(sused,dubina+1,suprotan
       smer,odstup)

       if odstupzvoljen_odstup then
           za sve cvorove sused za koje je broj ulaz
           nih linkova suprotan od onog koji nam treba
           radi Pretraga(sused,dubina+1,smer,odstup+1)
end;
```

Heuristika za traženje maksimalne klike

Ispostavilo se da je traženje heuristike za poboljšanje nalaženja maksimalne klike dosta jednostavnije nego traženje heuristike za Hamiltonove konture. Na eksperimentalnom uzorku mreža koje su veličine od 10 do 4000 čvorova, utvrđeno je da je maksimalna klika za sve mreže veličine 4 (važno je primetiti da su te scale-free mreže pravljene po Barabasijevom modelu). Dosta jednostavna heuristika, koja mnogo ubrzava posao, a koja je zasnovana na tome da je maksimalna klika veća (to će se pokazati na većim mrežama) ili jednaka 4, je sledeća: iz mreže treba izbaciti sve čvorove ciji je broj linkova manji od 4, i onda na tako smanjenoj mreži naći kliku nekim standardnim postupkom (npr. backtrack). Time se odstranjuje preko 50% čvorova. Testirano je i da li se pronalaze sve klike ako se odstrani još više čvorova. Ako izbacimo prvih 65% čvorova koji imaju najmanji broj veza, broj nađenih klika je u proseku za 1 manji od ukupnog broja klika. Za pronalaženje bilo koje maksimalne klike, moguće je odstraniti i do 80% čvorova. Time je omogućeno brzo traženje veličine klike na velikim mrežama.

Zaključak

Ovaj rad se bavi ispitivanjem karakteristika NP kompletnih problema na scale-free mrežama. Traženje Hamiltonove konture pretragom uz korišćenje dobijene heuristike, primenljivo je samo na manjim mrežama (na kojima daje dobre rezultate), zato što vremenska složenost nije polinomna. Za razliku od toga, heuristika za traženje maksimalne klike je znatno efikasnija, i omogućava pretragu na većim mrežama.

Reference

Barabasi A. 1999. *Scale-free Networks*.

Reka A., Barabasi A. 2002. *Statistical mechanics of complex networks*. University of Notre Dame

Živković M. *Algoritmi*. 2000. Beograd: Matematički fakultet.

Rajko Nenadov

Heuristics for Solving NP-Complete Problems on Scale-Free Networks

This paper describes construction of heuristics for solving two NP-complete problems on scale-free networks. Those two are problems of Hamilton cycle and clique. Considering the Hamilton cycle problem, the goal was to find some regularity in the characteristics of the nodes that form the cycle, while the focus of the max clique problem was in finding the nodes that may be members of the solution, and then consider only them. Based on this results, two “branch-and-bound” heuristics were made. First (Hummilton) can be used for smaller networks, about 35 nodes, while the other, for finding maximum cliques, is applicable to much bigger networks (4000 nodes).