

Evolucija optimizacije broja međuklusterskih veza, u društvenim mrežama

Rava Filipović

ORAO a.d.

Bijeljina, Republika Srpska, BiH
 filipovicrava@yahoo.com

Sadržaj—U ovom radu je opisan razvoj efikasnih i efektivnih postupaka klasterovanja članova neke društvene mreže sa posebnim osvrtom na evoluciju optimizacije vrijednosti funkcije cilja inicijalnog rješenja, pri čemu su napravljena tabelarna i grafička poredjenja inicijalnih rješenja i njihovih poboljšanih rješenja (tabu heuristikom). Ovaj rad je implementiran u okviru rješavanja problema klasterovanja tako da se minimizira broj međuklusterskih veza, u društvenim mrežama.

Ključne riječi—klasteri; optimizacija; heuristike;

I. UVOD

Otkrivanje klastera¹ ili zajednica u velikim realnim grafovima, kao što su velike društvene mreže², je značajan problem od praktičnog interesa. Pojam optimizacija međuklusterskih veza, za potrebe velikih društvenih mreža, jeste praktičan problem koji je ponikao u fazi zahuktalog razvoja društvenih mreža, na bazi interneta i novih informacionih tehnologija. Razvoj heurističkih³ algoritama⁴, za potrebe velikih društvenih mreža, predstavlja jedan od onih zadataka u kome udruženi čovjek i računar daju novi kvalitet, koji je samo pre neku deceniju bio nezamisliv. Cilj ovog istraživanja je razvoj efikasnih i efektivnih postupaka klasterovanja članova neke društvene mreže.

II. KLASITERI U DRUŠTVENIM MREŽAMA

Socijalna ili društvena mreža je društvena struktura sastavljena od pojedinaca (individua), koji predstavljaju čvorove. Čvorovi su povezani jednim ili više specifičnih tipova međuzavisnosti, kao što su: prijateljstvo, vizije, ideje, finansijski interes, ... Društvene mreže su modelirane pomoću grafa. Čvorovi su individue unutar mreže, a veze su njihovi

međusobni odnosi. Osim što predstavljaju neku vrstu trenda, što doprinosi njihovoj popularnosti, one donose i nove mogućnosti u komunikaciji i olakšano sklapanje kontakata, kao i upotrebu i razmjenu različitih sadržaja, što još više doprinosi njihovom značaju. Društvene mreže pružaju razne mogućnosti i neosporno je da od njih ima dosta koristi. Otkriće tijesno povezanih grupa, po raznim oblicima veza, u ovim mrežama je od fundamentalnog i praktičnog interesa.

III. KLASTEROVANJE KAO PROBLEM OPTIMIZACIJE

Klasterovanje⁵, kao jedna od aktivnosti istraživanja podataka, rješava optimizacione⁶ probleme grupisanja elemenata na osnovu nekih veza (relacija) koje postoje između njih. Analizu podataka je lakše izvesti nad grupisanim podacima u klaster, jer se time omogućava: pojednostavljenje skupa podataka, proučavanje postupaka, principa grupisanja, identifikacija odnosa, ... Takođe, grupisanje skupa podataka u manje grupe (klaster) predstavlja izazov u otkrivanju pravila, modela i zakonitosti na osnovu kojih mogu da se donesu važne odluke. Uočena zapažanja mogu biti korištena u svrhu daljih analiza. Postoji veliki broj definicija klastera u literaturi, na primjer [1], [2], [5]. Sve one nose isti pojam podskup grafa (ovde shvata kao graf odnosa prijateljstava korisnika društvene mreže) u kome je koncentracija linkova između čvorova (korisnika) unutar klastera znatno veća nego koncentracija linkova između klastera.

IV. KONSTRUKTIVNE HEURISTIKE

Konstruktivne heuristike generišu samo jedno dopustivo rješenje problema. Pohlepni (greedy) algoritam je jedna konstruktivna heuristika. Ovi algoritmi su dosta brzi, jednostavni, koriste se za online probleme i daju dobre rezultate. Uopšteno rečeno pohlepne algoritme je lako razumjeti i implementirati. Jednostavni su i brzi. Zato, koliko su dobri? Za većinu problema optimizacije, ostvaruju rješenja, koja se iterativno poboljšavaju nekim od metoda heuristika, da bi se dobilo na kvalitetu rješenja.

¹Klaster (engl. Cluster) je mala grupa ili gomila nečega.

²Društvena mreža (takođe i socijalna mreža) je društvena struktura sastavljena od pojedinaca (ili organizacija) koji se nazivaju "čvorovi", a koji su povezani jednim ili više specifičnih tipova „veza“, kao što su: prijateljstvo, vizije, ideje, finansijski interes, srodstvo, zajednički interes, finansijska razmjena, nedopadanje, ili odnosi povjerenja, znanja ili prestiža.

³Heuristika - nauka o metodama i principima pronalaženja novog.

⁴Algoritam je konačna i precizno definisana procedura, niz dobro definisanih pravila, kojom se ulazne vrijednosti transformišu u izlazne ili se opisuje izvršavanje nekog postupka.

⁵„Klastering“ (Clustering): Klasterovanje se odnosi na grupisanje podataka po sličnosti.

⁶Optimizacija ili matematičko programiranje je grana matematike koja proučava maksimiziranje i minimiziranje funkcija.

V. RJEŠENJE PROBLEMA

A. Definisan problem klasterovanja

Proučavan problem klasterovanja se može definisati na sledeći način: Dat je skup od N osoba i matrica prijateljstava između osoba. Rasporediti datih N osoba u K klastera tako da u svakom klasteru bude bar $N/(2 \cdot K)$ osoba, ali ne više od $(2 \cdot N)/K$ osoba i da ukupan broj parova osoba koji su prijatelji a pri tome se nalaze u različitim klasterima (tzv. međuklasterska prijateljstva) bude što manji (minimalan).

Treba naglasiti da su za potrebe ovog rada uzeti testni rezultati (rada "Empirijski rezultati Tabu pretraživanja za problem podjele u klastere, u društvenim mrežama", koji je nastavak ovog dijela istraživanja proučavanog problema podjele u klastere), kako bi se napravila tabelarna i grafička poredjenja evolucije optimizacije međuklasterskih veza i veza unutar klastera inicijalnog rješenja, što je ujedno i cilj ovog problema proučavanja.

B. Podjela u klastere

Na početku algoritma se generiše jedna permutacija cijelih brojeva (Permutacija osoba: 1 17 3 20 24 6 12 4 26 8 28 5 29 23 15 25 22 14 18 19 7 21 30 10 13 11 27 9 16 2) od 1 do N (gdje je N broj osoba). Skup od N osoba se particioniše u K disjunktih podskupova, koji predstavljaju upravo tih K klastera, na sledeći način. Redom osoba po osoba se smješta u slučajno odabran klaster (jedan od željenog broja klastera K), pri čemu se vodi računa prvo o minimalnom, a posle i o maksimalnom mogućem broju osoba u klasterima. Intuitivno, posle velikog broja eksperimenata, je postavljen dodatni uslov da svaki klaster mora imati bar $N/(2 \cdot K)$ osoba i ne više od $(2 \cdot N)/K$ osoba. Tako, prvo se gleda koji slučajno odabran klaster ima manji broj osoba od minimalno dozvoljene vrijednosti i u njega se stavlja redom odabrana osoba iz generisane permutacije cijelih brojeva. Postupak se ponavlja sve dotle dok jedan od klastera ne dostigne minimalnu vrijednost broja osoba, pri čemu taj klaster prestaje da se puni osobama. Na isti način, postupak se ponavlja sve dotle dok preostali klasteri (jedan po jedan) ne dostignu svoju minimalnu vrijednost. Postupak se nastavlja smještanjem svake preostale osobe (generisane permutacije cijelih brojeva skupa od N osoba) u slučajno odabrane klastere (koji su dostigli minimalnu vrijednost broja osoba) do maksimalno mogućeg broja osoba u njima. Kada slučajno odabran klaster upunjem minimalno mogućim brojem osoba dostigne maksimalnu vrijednost (broja osoba) tada taj klaster prestaje da se puni osobama. Postupak se ponavlja sve dotle dok se ne potroše preostale osobe iz generisane permutacije. Ovim grupisanje skupa od N osoba u željen broj klastera K je završen. Konstruisano je jedno moguće rješenje, nazvano početno. Ovom polaznom rješenju se dodjeljuje broj međuklasterskih veza kao i broj veza unutar svakog klastera. Početno rješenje i njemu pridružen broj međuklasterskih veza se pamte i postaju tekući. Za ilustrovan primjer dobijenih inicijalnih rješenja korištena je testna instanca⁷ $N=30$ osoba i definisan broj klastera je $K=4$. Svako

pojedinačno rješenje je predstavljeno distribucijom cijelih brojeva (osoba) po klasterima, pri čemu im je pridružen odgovarajući broj međuklasterskih prijateljstava i broj prijateljstava unutar klastera:

```
Permutacija osoba: 1 17 3 20 24 6 12 4 26 8 28 5 29 23 15 25 22 14 18 19 7 21 30 10 13 11 27 9 16 2
Cluster[0]= 1 24 6 23 15 22 21
Cluster[1]= 17 4 28 29 30 16
Cluster[2]= 3 8 5 14 18 19 27 9
Cluster[3]= 20 12 26 25 7 10 13 11 2
Ukupan broj međuklasterskih prijateljstava je 111
Ukupan broj prijateljstava u Cluster[1] = 7
Ukupan broj prijateljstava u Cluster[2] = 6
Ukupan broj prijateljstava u Cluster[3] = 7
Ukupan broj prijateljstava u Cluster[4] = 10
```

```
Permutacija osoba: 16 13 5 7 3 2 27 17 30 10 8 1 22 29 21 15 11 12 23 20 6 9 19 26 28 4 24 25 14 18
Cluster[0]= 16 10 8 25 18
Cluster[1]= 13 17 1 23 20 9 28
Cluster[2]= 5 3 30 11 6 4 24
Cluster[3]= 7 2 27 22 29 21 15 12 19 26 14
Ukupan broj međuklasterskih prijateljstava je 104
Ukupan broj prijateljstava u Cluster[1] = 5
Ukupan broj prijateljstava u Cluster[2] = 7
Ukupan broj prijateljstava u Cluster[3] = 5
Ukupan broj prijateljstava u Cluster[4] = 20
```

...

```
Permutacija osoba: 30 2 3 4 27 10 1 5 21 13 19 6 29 23 9 15 7 22 14 8 20 18 25 24 26 12 11 28 16 17
Cluster[0]= 30 21 13 7 14 24 26
Cluster[1]= 2 5 6 29 20 12 11
Cluster[2]= 3 27 10 23 9 15 16 17
Cluster[3]= 4 1 19 22 8 18 25 28
Ukupan broj međuklasterskih prijateljstava je 110
Ukupan broj prijateljstava u Cluster[1] = 4
Ukupan broj prijateljstava u Cluster[2] = 7
Ukupan broj prijateljstava u Cluster[3] = 12
Ukupan broj prijateljstava u Cluster[4] = 8
```

```
Permutacija osoba: 29 27 3 14 19 11 18 1 13 17 20 28 9 10 2 16 25 15 21 22 23 12 7 30 4 6 5 26 24 8
```

```
Cluster[1]= 29 11 28 9 10 15 12 30 6 26 24 8
Cluster[1]= 27 18 13 2 22 5
Cluster[2]= 3 17 20 16 21 7 4
Cluster[3]= 14 19 1 25 23
Ukupan broj međuklasterskih prijateljstava je 105
Ukupan broj prijateljstava u Cluster[1] = 15
Ukupan broj prijateljstava u Cluster[2] = 6
Ukupan broj prijateljstava u Cluster[3] = 9
Ukupan broj prijateljstava u Cluster[4] = 6
```

U tabeli, koja sledi, dat je broj međuklasterskih veza i broj veza unutar svakog klastera inicijalnih rješenja .

TABELA I. BR.MED. V. I BR. V.UNUTAR KLASTERA INICIJALNOG RJ.

Br. med. veza	Br. veza unutar klastera			
	K1	K2	K3	K4
111	7	6	7	10
104	5	7	5	20
110	4	7	12	8
107	7	6	16	5
112	5	7	14	3
109	6	3	5	18
103	17	7	1	13
111	6	8	12	4
109	8	10	5	9
114	4	5	4	14
98	6	6	25	6
111	17	3	4	6
108	11	10	2	10
100	4	22	9	6
108	2	9	6	16
113	6	5	6	11
110	15	4	5	7
105	8	2	8	18
110	9	6	6	10
105	15	6	9	6

Na inicijalno rješenje se primijeni lokalno pretraživanje tabu heuristikom (opisano u radu "Empirijski rezultati Tabu pretraživanja za problem podjele u klastere, u društvenim mrežama", kao nastavku ovog rada), a potom se algoritam restarta od početka. Dvadeset restarta algoritma je jedan od kriterija zaustavljanja rada algoritma. Kao nastavak (prethodno) ilustrovanog primjera inicijalnih rješenja slede

⁷ Instanca - konkretan test primjer.

njihova poboljšana rješenja predstavljena distribucijom cijelih brojeva (osoba) po klasterima, pri čemu im je pridružen odgovarajući broj međuklusterskih prijateljstava i broj prijateljstava unutar klastera:

Cluster[0]= 3 18 19 15 23 16 24 17 11 7 25 26 20
Cluster[1]= 13 30 29 10 4 12
Cluster[2]= 21 28 27 14 5 2 6 22
Cluster[3]= 8 9 1
Ukupan broj međuklusterskih prijateljstava je 45
Ukupan broj prijateljstava u Cluster[1] = 64
Ukupan broj prijateljstava u Cluster[2] = 11
Ukupan broj prijateljstava u Cluster[3] = 21
Ukupan broj prijateljstava u Cluster[4] = 0

Cluster[0]= 16 10 8
Cluster[1]= 24 11 3 25 18 20 23 1 17 7 15 19
Cluster[2]= 4 30 13 9 29
Cluster[3]= 2 27 22 28 5 6 14 26 12 21
Ukupan broj međuklusterskih prijateljstava je 39
Ukupan broj prijateljstava u Cluster[1] = 2
Ukupan broj prijateljstava u Cluster[2] = 59
Ukupan broj prijateljstava u Cluster[3] = 10
Ukupan broj prijateljstava u Cluster[4] = 31

...

Cluster[0]= 18 17 3 11 16 19 7 24 1 20 23 25 15
Cluster[1]= 13 8 30 10 21 12 9 4 29
Cluster[2]= 27 5 14
Cluster[3]= 28 2 22 26 6
Ukupan broj međuklusterskih prijateljstava je 37
Ukupan broj prijateljstava u Cluster[1] = 71
Ukupan broj prijateljstava u Cluster[2] = 20
Ukupan broj prijateljstava u Cluster[3] = 3
Ukupan broj prijateljstava u Cluster[4] = 10

Cluster[0]= 4 30 10 9 29 13
Cluster[1]= 28 6 26 5 22 2 27 14
Cluster[2]= 1 24 3 15 11 18 7 19 16 20 17 25 23
Cluster[3]= 21 8 12
Ukupan broj međuklusterskih prijateljstava 27
Ukupan broj prijateljstava u Cluster[1] = 15
Ukupan broj prijateljstava u Cluster[2] = 26
Ukupan broj prijateljstava u Cluster[3] = 71
Ukupan broj prijateljstava u Cluster[4] = 2

U sledećoj tabeli dat je broj međuklusterskih veza i broj veza unutar svakog klastera popravljenih inicijalnih rješenja (dalje u tekstu popravljenog ili popravljenih rješenja).

TABELA II. BR.MED. V. I BR. V.UNUTAR KLASTERA POP. INIC. RJ.

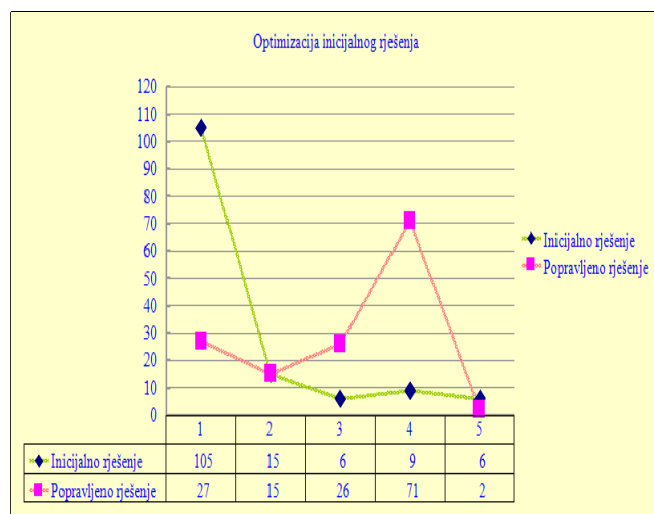
Br. med. veza	Br. veza unutar klastera			
	K1	K2	K3	K4
45	64	11	21	0
39	2	59	10	31
46	16	31	45	3
47	22	16	55	1
52	3	34	49	3
37	15	3	77	9
35	75	23	2	6
72	13	32	21	3
33	30	73	3	2
40	73	4	3	21
63	4	5	53	16
38	71	3	28	1
33	28	74	3	3
61	1	59	19	1
43	0	59	11	28
32	20	77	10	2
54	2	65	2	18
49	15	1	20	56
37	71	20	3	10
27	15	26	71	2

Ako se uporedi jedno inicijalno i njegovo popravljeno rješenje (iz tabela I i II) uočava se znatna razlika između rješenja, po broju međuklusterskih veza i broju prijateljskih veza unutar klastera. Broj međuklusterskih veza inicijalnog

rješenja je znatno veći od njegovog popravljenog rješenja. Dok, broj prijateljskih veza u klasterima inicijalnog rješenja je manji od broja međuklusterskih veza njegovog popravljenog rješenja. Isto se odnosi za preostala inicijalna rješenja i njihova popravljena rješenja.

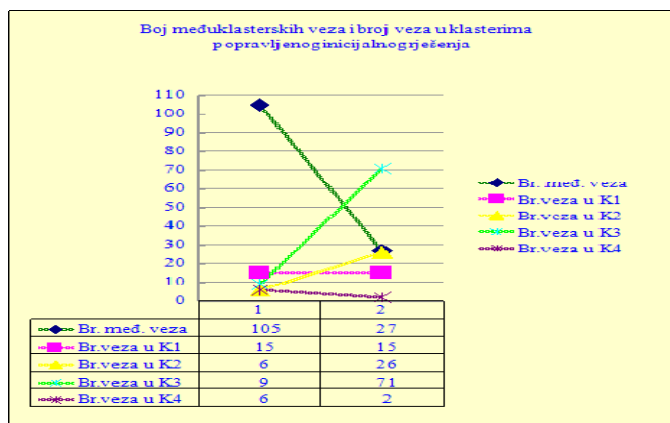
Potom su napravljena grafička poredjenja Sl.1 jednog inicijalnog rješenja i njegovog popravljenog rješenja, pri čemu podaci su uzeti po redovima tabela I i II.

Na Sl.1 jedno inicijalno rješenje (zelena linija) je imalo 105 međuklusterskih veza smještanjem skupa osoba (N=30) u definisan broj klastera (K=4). Prateći dalje tu liniju vidi se da je broj veza unutar klastera znatno manji od broja međuklusterskih veza, tako da je (zelena) linija skoro prišla x osi po pitanju broja veza unutar klastera. Druga linija (roze boje) odnosi se na popravljeno inicijalno rješenje. Prva tačka na ovoj liniji je daleko niža i odnosi se na broj međuklusterskih veza. Minimiziranjem broja međuklusterskih veza popravljenog rješenja (roze linija) se svelo na 27 međuklusterske veze. Broj međuklusterskih veza se znatno smanjio sa 105 na 27. Dalje, prateći istu liniju po drugim tačkama, broj prijateljskih veza unutar svaka od četiri klastera je naglo porastao, popravljanjem vrijednosti funkcije cilja inicijalnog rješenja. Prema tome, optimizovanjem broja međuklusterskih veza inicijalnog rješenja znatno je povećan broj prijateljskih veza unutar klastera Sl.1.



Slika 1. Optimizacija inicijalnog rješenja

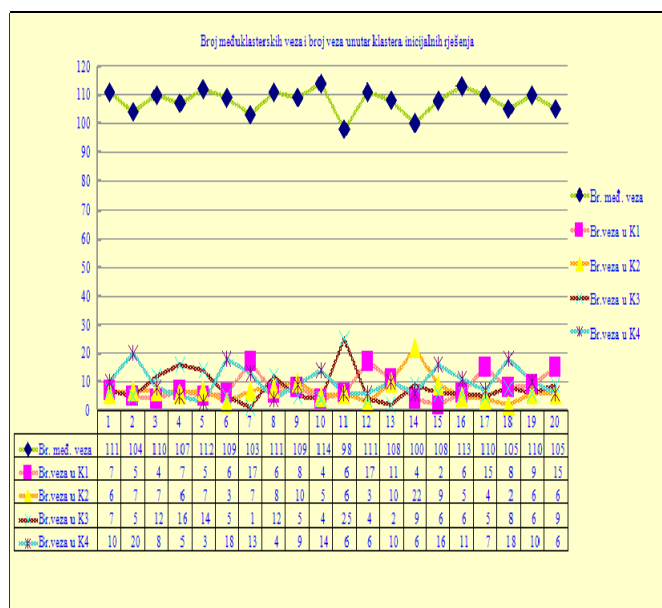
Dalje, napravljena su grafička poredjenja jednog inicijalnog rješenja i njegovog popravljenog rješenja Sl.2, pri čemu podaci su uzeti po kolonama (inicijalnog i njegovog popravljenog rješenja) tabela I i II. Na Sl.2 prateći pojedinačno liniju po liniju, koje se odnose na broj međuklusterskih veza i broj veza unutar svakog klastera, najviša tačka zelene linije se odnosi na broj međuklusterskih veza inicijalnog rješenja, a najniža tačka (iste linije) predstavlja njegovo popravljeno rješenje. Što se tiče ostalih linija one se odnose na broj veza unutar klastera inicijalnog i njegovog popravljenog rješenja. Odma se da uočiti nagli pad broja međuklusterskih veza inicijalnog rješenja i nagli rast broja veza unutar klastera poboljšanog rješenja.



Slika 2. Optimizovan br. međ. veza i br. veza unutar klastera

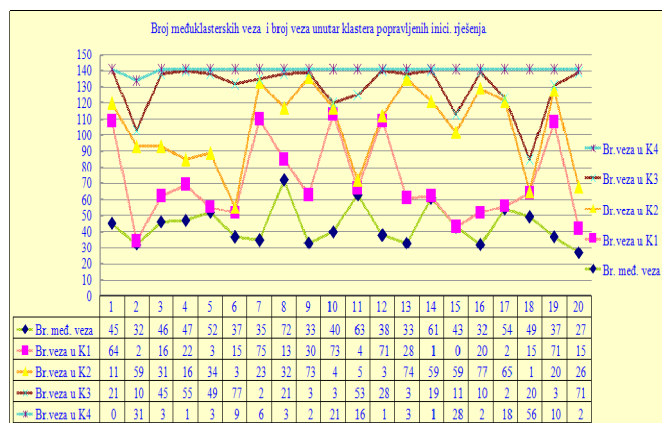
Praveći tabelarna poredjenja tabele I i tabele II mogu se uporediti karakteristike broja međuklasterских veza i veza unutar klastera inicijalnih rješenja i njihovih popravljenih rješenja, pri čemu se da uočiti znatno poboljšanje kvaliteta vrijednosti funkcije prilagođenosti reprezentacije polaznih rješenja. Potom su napravljena grafička poredjenja inicijalnih rješenja i njihovih popravljenih rješenja na Sl.3 i Sl.4 (pri čemu su podaci iz datih tabela uzeti po redovima datih tabela) i na Sl.5 i Sl.6 (pri čemu su podaci uzeti po kolonama datih tabela).

Na grafu Sl.3 odma se uočava linija (zelene boje) koju prate visoke vrijednosti funkcije cilja inicijalnih rješenja, dok se preostale linije sa znatno nižim vrijednostima odnose na broj veza u klasterima.



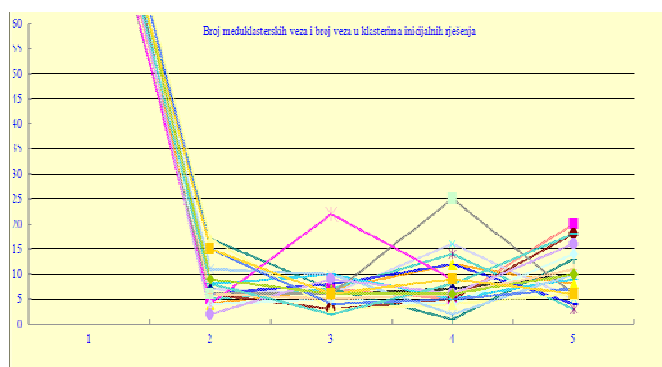
Slika 3. Inicijalna rješenja

Na Sl.4 liniju (zelene boje) prate znatno niže vrijednosti funkcije cilja popravljenih rješenja, dok se preostale linije, sa znatno višim vrijednostima, odnose na broj veza u klasterima. Optimizacijom broja međuklasterских veza polaznih rješenja broj međuklasterских veza se znatno smanjio, a broj veza unutar klastera se naglo povećao.



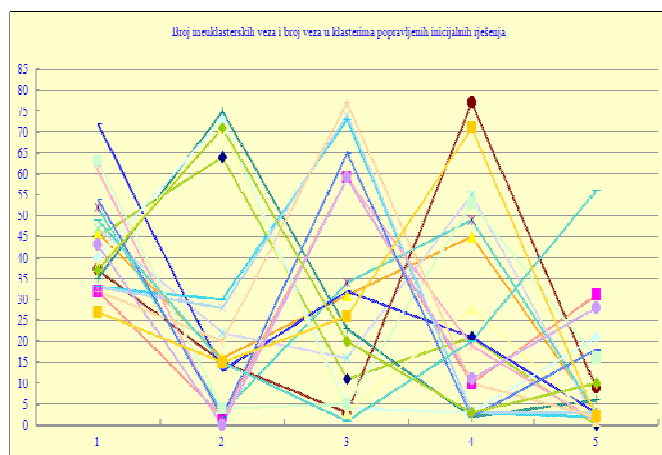
Slika 4. Optimizovana inicijalna rješenja

Sl.5 daje grafički prikaz (pri čemu su podaci uzeti po kolonama datih tabela) inicijalnih rješenja čije vrijednosti funkcije cilja broja međuklasterских veza su visoke, a vrijednosti broja veza unutar klastera su niske.



Slika 5. Br. međ. veza i br. veza unutar klastera inicijalnih rješenja

Sl.6 daje grafički prikaz (pri čemu su podaci uzeti po redovima datih tabela) popravljenih rješenja čije vrijednosti funkcije cilja broja međuklasterских veza su minimizirane, a vrijednosti broja veza unutar klastera su znatno porasle. Najkvalitetnije popravljeno rješenje prati žuta linija, sa najmanjim brojem međuklasterских veza.



Slika 6. Optimizovan br. međ. veza i br. veza unutar klastera

C. Naučni doprinos

Najvažniji novi rezultati dobijeni istraživanjem prikazanim u ovom radu su:

- Definirana varijanta problema podjele u klastere koja je realnija (više odgovara praktičnim potrebama), ali zato kompleksnija za rješavanje.
- Definiranje odgovarajuće funkcije cilja.
- Definisan nov način kodiranja prezentacije rješenja cijelim brojevima.
- Grupisanje ili izdvajanje sličnih objekata iz neke veće cjeline u manje grupe, čime se pojednostavljuje analiza podataka.
- Razvoj efikasnih i efektivnih postupaka klasterovanja.

Glavni doprinos je razvoj tabu algoritma za prethodno uvedenu varijantu problema klasterovanja. Razvoj ovog algoritma je zahtijevao da se razvije:

- Novi način predstavljanja potencijalnih rješenja.
- Postupak za formiranje inicijalnih rješenja.

Realizovani algoritam samostalno razvijen po performansama prevazilazi druge algoritme za klasterovanje i u zadovoljavajućem vremenu pronalazi izuzetno dobru podjelu u klastere. Takođe, treba naglasiti (kao doprinos) da su generisane test instance, koje se mogu koristiti za testiranje (poredjenje) metoda koje će u budućnosti biti razvijene.

Kao što se može vidjeti iz eksperimentalnih rezultata, predložen tabu algoritam je veoma uspješan pri rješavanju problema podjele u klastere. Zbog svega gore navedenog, naučno istraživanje opisano u ovom radu daje doprinos oblastima kombinatorne optimizacije i lokacijskih problema. Proučavan problem je razvijen u C# i testiran na primjerima iz literature, realnog života i primjerima generisanim na slučajan način.

ZAKLJUČAK

U mnogim životnim oblastima, istraživači se često sreću sa situacijama u kojima je potrebno objekte razvrstati u grupe homogenih objekata bilo da su individue, firme, proizvodi ili čak njihova ponašanja. Strateška rješenja bazirana na identifikaciji grupa pomažu u donošenju važnih odluka. Ova ista potreba se sreće u različitim područjima, od fizičkih do društvenih nauka.

Za proučavani problem, bilo je neophodno razmotriti i osmisliti efikasne i efektivne postupke smještanja objekata u klastere, neke društvene mreže. Efikasnost podrazumijeva dobijanje odgovarajuće podjele objekata u manje grupe (klastere) u što kraćem vremenu, a efektivnost se odnosi na kvalitet dobijenog rješenja sa što manjim brojem međuklusterskih veza. Međutim, postoji mnogo faktora koji utiču na efikasnost postupka smještanja osoba u klastere (broj osoba, broj klastera, maksimalan i minimalan moguć broj osoba u klasterima, ukupno utrošeno vrijeme, ...), od kojih

mnogi zavise jedni od drugih. Evolucija optimizacije vrijednosti funkcije cilja inicijalnih rješenja se najbolje očituje na tabelarnim i grafičkim poredjenjima inicijalnih rješenja i njihovih popraavljenih rješenja. Doprinos rada je program koji rješava problem podjele u klastere implementiran u okviru rješavanja problema klasterovanja tako da se minimizira broj međuklusterskih veza, u društvenim mrežama. Proučavan optimizacioni problem podjele u klastere je šire primjenljiv u praktične svrhe. Tako isti programski kod se može lako prilagoditi različitim primjenama u svim oblastima ljudskih aktivnosti (biologije, anatomije, hemije, fizike, lingvistike, ekonomije, računarstva, sporta, muzike, ...).

Proučavani problem, postupka podjele u klastere je bio složen i zahtijevao je izvođenje velikog broja eksperimenata. Intuitivnim postupcima smještanja osoba u klastere došlo se intuitivno do parametara, koji su minimizirali broj međuklusterskih veza. Dalja istraživanja će se kretati u pravcu primjene drugih heuristika i njihovih kombinacija, na istom proučavanom problemu, u cilju što efikasnijeg rješavanja proučavanog problema klasterovanja.

LITERATURA

- [1] M. Girvan and M. E. J. Newman, Community structure in social and biological networks. *Proc. Natl. Acad. Sci. USA* 99, 7821-7826 (2002).
- [2] M. E. J. Newman, Fast algorithm for detecting community structure in networks. Preprint cond-mat/0309508 (2003).
- [3] Josep M. Pujol, Javier Béjar, and Jordi Delgado, Clustering algorithm for determining community structure in large networks, *Software Department, Technical University of Catalonia, Jordi Girona 1-3 A0-S106, 08034, Barcelona, Spain*, 2006.
- [4] J. Leskovec, K. Lang, A. Dasgupta, and M. Mahoney, Community structure in large networks: Natural cluster sizes and the absence of large well-defined clusters. arXiv:0810.1355, October 2009.
- [5] S. Fortunato, Community detection in graphs. arXiv:0906.0612, June 2009.
- [6] J. Leskovec, K. Lang, M. Mahoney, Empirical Comparison of Algorithms for Network Community Detection, *ACM WWW International conference on World Wide Web (WWW)*, 2010.
- [7] Glover, F. and Laguna, M., 1997, *Tabu Search*, Kluwer, Boston, MA.
- [8] J. Yang, J. Leskovec, Defining and Evaluating Network Communities based on Ground-truth, *IEEE International Conference On Data Mining (ICDM)*, 2012.

ABSTRACT

This paper describes the development of efficient and effective methods of clustering members of a social network, with special emphasis on the evolution of the optimization objective function value of the initial solution, with built in tabular and graphical comparisons of initial solutions and their improved solutions (tabu heuristics). This work has been implemented within the framework of solving the problem of clustering to minimize the number of connections between clusters in social networks.

EVOLUTION OF THE OPTIMIZATION THE NUMBER OF CONNECTIOS BETWEEN CLUSTERS IN SOCIAL NETWORKS

Rava Filipović