

SVEUČILIŠTE U ZAGREBU  
FAKULTET ELEKTROTEHNIKE I RAČUNARSTVA

DIPLOMSKI RAD br. 2977

# Usporedba algoritama otkrivanja zajednica u društvenim mrežama

Daniel Marić

Zagreb, svibanj 2022.

*Umjesto ove stranice umetnite izvornik Vašeg rada.  
Da bi ste uklonili ovu stranicu obrišite naredbu \izvornik.*



# SADRŽAJ

1. Uvod	1
2. Društvene mreže i zajednice	3
3. Algoritmi otkrivanja društvenih zajednica	4
4. Skupovi podataka	5
5. Programsko ostvarenje	6
6. Vrednovanje i rezultati	7
7. Zaključak	8
Literatura	9

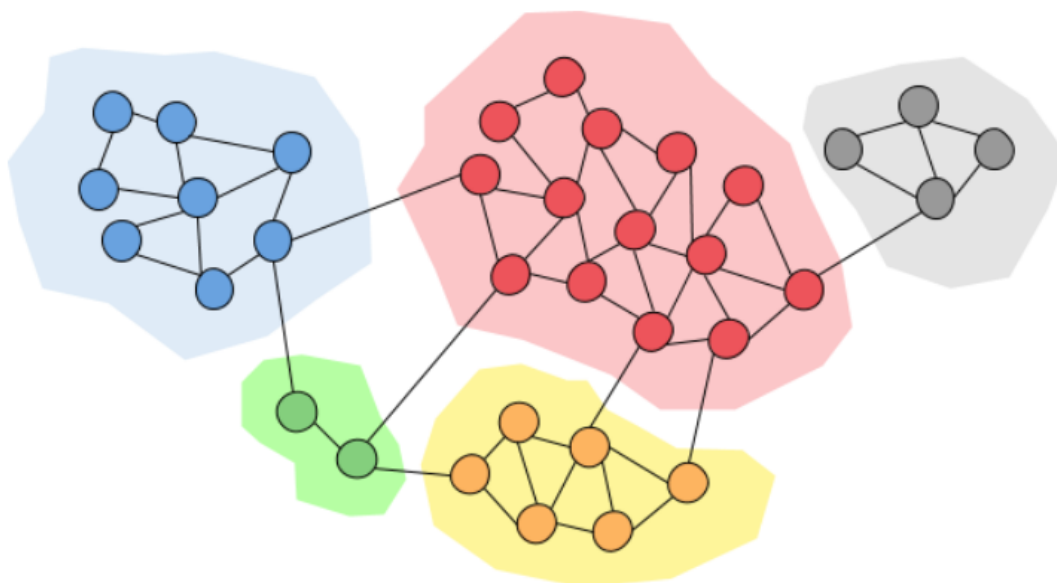
# 1. Uvod

Pojavom popularnih internetskih usluga za povezivanje korisnika stvorene su velike mreže društvenih zajednica. Generirane su velike količine podataka iz kojih je moguće izvući mnoštvo korisnih informacija. Takve zajednice sastoje se od puno manjih zajednica koje se po svojim karakteristikama razlikuju od ostalih. Takve zajednice potrebno je pronaći kako bi im se pristupilo na najbolji mogući način. Rješavanje ovog problema važno je i u drugim granama znanosti kao na primjer u sociologiji, biologiji ili računarskoj znanosti gdje su problemi predstavljeni na takav način, pomoću strukture grafa.

Upravo su grafovi najpogodnija struktura podataka za pristup ovome problemu gdje relevantne značajke, u primjeru društvenih mreža ljude, možemo prikazati pomoću čvorova dok će bridovi predstavljati veze između tih značajki. Više bridova među određenim značajkama značit će da tu mogu postojati obilježja zajednice, npr. u biologiji bi to mogla biti tkiva koja u organima obavljaju sličnu ulogu.

Rješavanje problema koji su predstavljeni grafovima je vrlo složeno, vremenski i prostorno. Ovakvi grafovi nisu jednostavnog oblika, ali u njima postoje određene pravilnosti koje se mogu iskoristiti. U tu svrhu razvijeno je mnogo algoritama za otkrivanje zajednica koji različitim pristupima pokušavaju pronaći rješenje ovog problema. Pojedini algoritmi su bolji od drugih na jednom tipu društvenih mreža ili lošiji na drugom te se zato koriste evaluacijske mjere kojima se procjenjuje koliko je dobro rješenje koje je algoritam pronašao. Što više algoritama se testira s različitim društvenim mrežama i evaluacijskim mjerama dobit ćemo bolji uvid u to kada je koji bolje koristiti. Najpoznatiji algoritam među njima je Newman-Girvanov algoritam koji će se nešto detaljnije opisati uz još nekoliko njih. Sličnim problemima bavili su se radovi Lancichinettija i Fortunata [1] iz 2016. te [2] iz 2009. godine.

Unatrag posljednjih nekoliko godina uvedeni su zakoni o zaštiti osobnih podataka te je sada znatno teže dobiti pristup korisnim informacijama. Zato se



**Slika 1.1:** Primjer grafa nepreklapajućih društvenih zajednica.

koriste posebni algoritmi za generiranje umjetnih skupova podataka koji će za zadane parametre generirati graf pomoću kojih se mogu provoditi istraživanja.

U nastavku rada bit će opisana struktura i svojstva društvenih mreža i zajednica, algoritmi koji pronalaze društvene zajednice, skupovi podataka koji su korišteni u sklopu rada, programsko rješenje koje pokreće i evaluira rješenja algoritama te će se prikazati rezultati i do kojih se došlo.

## 2. Društvene mreže i zajednice

Društvene mreže može se pronaći gdje god postoji sustav koji sadrži entitete koji su međusobno povezani. Primjera je mnogo, a neki od njih su: društvene web platforme, email mreža, web stranice koje sadrže poveznice prema dugima, uređaji koji su povezani preko internetske mreže i slično. Kako bi se skupina entiteta mogla nazvati društvenom zajednicom među njima mora postojati nekakav tip odnosa. Može biti jednosmjerni ili dvosmjerni te mogu postojati težine kojima se odnosu daje veća ili manja značajnost. Društvene mreže imaju složenu organizacijsku strukturu te se može pretpostaviti svojstvo lokalnosti koje kaže da ako jedan entitet ima veze prema neka druga dva entiteta onda je vjerojatnost da ta druga dva entiteta imaju vezu veća od prosječne.

Društvene mreže imaju karakteristično svojstvo grupiranja u strukturu zajednice. Ako se čvorovi mreže mogu podijeliti u nepreklapajuće ili preklapajuće zajednice tako da broj veza između članova zajednice značajno premašuje broj veza između bilo koje dvije zajednice znači da mreža ima strukturu društvenih zajednica. Mreže koje imaju takvu strukturu često se mogu prikazati i kao hijerijske strukture. U ovom radu obradit će se mreže koje sadrže nepreklapajuće strukture sa vezama koje nemaju određene težine.

Proces pronalaska društvenih zajednica jedan je od glavnih zadataka u analizama društvenih mreža. Detekcija zajednica može biti vrlo korisna u raznim primjenama kao što je primjerice pronalaženje grupa kojima bi se mogle slati reklame za određene proizvode koji bi ih mogli zanimati umjesto da se svakom pojedincu šalju posebno. Još jedan primjer bio bi preporuka određenih sadržaja koji bi se mogli prikazivati grupama koje pokazuju zanimanja prema sličnim interesima. Primjera ima još mnogo, ali iz ova dva je već vidljivo da se korisne informacije mogu zaključivati iz društvenih mreža. Kako bi društvene mreže pohranili i analizirali u računalu potrebna je prikladna struktura podataka koja će u ovom slučaju biti graf.

Grafovi...

### 3. Algoritmi otkrivanja društvenih zajednica

...



## 4. Skupovi podataka

...

## 5. Programsko ostvarenje

Python i networkx

## 6. Vrednovanje i rezultati

Rezultati.

## 7. Zaključak

Zaključak.

# LITERATURA

- [1] Santo Fortunato i Darko Hric. Community detection in networks: A user guide. *Physics reports*, 659:1–44, 2016.
- [2] Andrea Lancichinetti i Santo Fortunato. Community detection algorithms: a comparative analysis. *Physical review E*, 80(5):056117, 2009.

## Usporedba algoritama otkrivanja zajednica u društvenim mrežama

### Sažetak

Sažetak na hrvatskom jeziku.

**Ključne riječi:** Ključne riječi, odvojene zarezima.

### Title on english

### Abstract

Abstract.

**Keywords:** Keywords.