

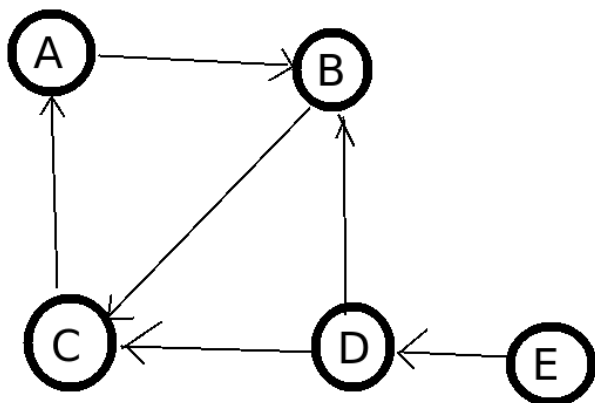
# AVSP završni ispit 2017

13. lipnja 2017.

## 1 1. zadatak

Za zadani graf na Slici 1 napisati jednadžbe toka ranga/utjecaja (engl. *rank*) za sve čvorove u grafu u grafu (engl. *Flow Equation Formulation*).

- Primijeni li se metoda uzastopnog potenciranja s ciljem izračunavanja vektora ranga  $\mathbf{r}$  na zadani graf, hoće li dobiveni rezultati biti vjerodostojni? Detaljno objasniti odgovor.
- Modificirati zadani graf uvođenjem teleportirajućih poveznica (engl. *teleports*). Napisati vrijednost *Google* matrice  $A$  ako vjerojatnost da će slučajni šetač (engl. *random walker*) slijediti poveznicu iz grafa iznosi  $\beta = 0.9$ .



Slika 1: Graf čvorova.

## 2 2. zadatak

U Tablici 1 zadana je matrica ocjena korisnika za pojedine filmove (engl. *user-item matrix*). Prazna polja u matrici predstavljaju ocjene koje nedostaju. Korištenjem algoritama suradničkog filtriranja (engl. *Collaborative Filtering*) potrebno je izračunati preporuku/ocjenu za korisnika U4 i film M5.

- Koristi se *Item-Item* pristup suradničkog filtriranja.
- Koristi se *User-User* pristup suradničkog filtriranja.

Tablica 1: *User-item matrix*.

x	U1	U2	U3	U4	U5
M1	2	1	1		3
M2		3	3		4
M3			3	1	2
M4	4	2	5	2	2
M5	2	5	4	?	1

### 3. zadatak

Razvijen je novi *Recommender System* za filmove, s preko  $10^6$  naslova. Budući da je sustav u ranoj fazi korištenja, ima relativno malo korisničkih ocjena, reda veličin  $10^3$ .

- Koji model (**content-based**, *item-based*, *user-based*) biste koristili u ovoj fazi i zašto?
- Korisnik A je filmovima *Godfather* i *Goodfellas* dao ocjenu 5/5. Koji mu od sljedeća tri filma sustav vjerojatno neće preporučiti i zašto: *Godfather 2*, *Shawshank Redemption*, *Finding Nemo*?
- Nakon nekog vremena, sustav ima dovoljno ocjena. Prosječna je ocjena filma 3.5. Prosječna ocjena koju korisnik A daje filmovima je 3.8. Prosječna ocjena za film *Deadpool* u sustavu je 4. Koliko iznosi *baseline estimation* za ocjenu korisnika A za film *Deadpool*?

### 4. zadatak

Skup podataka:

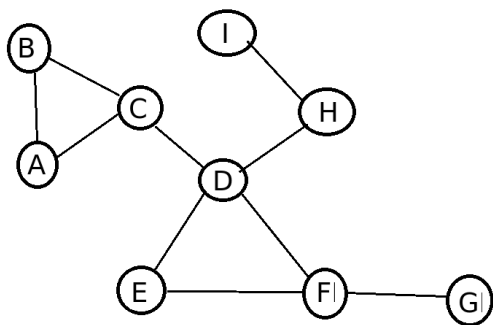
(1, 1), (1, 3), (1, 5), (2, 2), (2, 4) i (3, 3) grupiran je primjenom algoritma BFR (Bradley, Fayyad, Reina) u jednu grupu C1. Nakon toga, u skup je dodana točka (4, 4).

- Navedite stanje strukture podataka algoritma **prije** dodavanja točke.
- Definirajte i objasnite Mahalanobisovu udaljenost.
- Izračunajte  $MD((4,4))$  i navedite stanje strukture podataka **nakon** dodavanja točke.

### 5. zadatak

Zadan je graf društvene mreže na Slici 2.

- Navedite izlaz Girvan-Newman algoritma nakon 3 koraka.
- Nacrtajte AGM (*Affiliation Graph Model*) za mrežu nakon provođenja algoritma u zadatku a).
- Navedite definiciju društvene mreže.



Slika 2: Graf čvorova.

## 6 6. zadatak

U sustavu oglašavanja nalaze se 2 oglašivača:

- Oglašivač A ima početni budžet 5 i nudi jedan novac za oglase X i Y.
- Oglašivač B ima početni budžet 6 i nudi dva novca, ali samo za oglase Y.

Pretpostavite da sustav dodjeljuje oglasni prostor primjenom algoritma BALANCE i da se u izjednačenom slučaju oglasi dodjeljuju prema prioritetu  $A \rightarrow B$  – dakle, A ima veći prioritet.

- Kojim će se redoslijedom oglasi prikazivati korisnicima?
- Odredite *competitive ratio*.
- Navedite općeniti *competitive ratio* za 2 oglašivača s **istim** ograničenim budžetom.

## 7 7. zadatak

Razmatramo algoritam DGIM. Odredite i obrazložite:

- prostornu složenost **vremenske oznake** (engl. *timestamp*) pretinca,
- prostornu složenost **veličine** pretinca,
- ukupnu prostornu složenost cijelog algoritma,
- vremensku složenost odgovaranja na upit.