

UNIVERZITET U BEOGRADU
ELEKTROTEHNIČKI FAKULTET



ANALIZA SOCIJALNIH MREŽA

Projektni zadatak

Verzija 1.0

Predmetni nastavnici:

dr Marko Mišić, docent

dr Jelica Protić, redovni profesor

Školska godina:

2019/2020.

Beograd, decembar 2019.

SADRŽAJ

SADRŽAJ	2
1. UVOD	3
2. CILJ	3
3. POSTAVLJENI PROBLEM	3
3.1. UPOREDNA ANALIZA NAUČNE PRODUKCIJE IZ OBLASTI RAČUNARSTVA NA UNIVERZITETU U BEOGRADU.....	3
3.2. SKUP PODATAKA ZA ANALIZU	4
3.3. MODELOVANJE MREŽE	5
3.4. ISTRAŽIVAČKA PITANJA I CILJEVI	6
3.5. PREPORUČENE METODE I ALATI.....	7
4. REZULTATI	7
5. PREDAJA, ODBRANA I VREDNOVANJE	8

1. UVOD

U okviru ovog dokumenta su data uputstva za izradu projektnog zadatka na predmetu Analiza socijalnih mreža (13M111ASM) u školskoj 2019/2020. godini. Studenti treba da pažljivo pročitaju ovo uputstvo pre izrade projektnog zadatka. Studenti projektni zadatak rade **samostalno ili u paru**.

2. CILJ

Cilj projektnog zadatka na predmetu Analiza socijalnih mreža je praktična primena stečenog teorijskog znanja iz predmeta na primeru jednog konkretnog istraživačkog problema. Kroz zadati istraživački problem, studenti treba da izvrše prikupljanje, obradu i preliminarnu analizu primarnog (sirovog) skupa podataka, izdvoje neophodne podatke i modeliraju problem mrežom odgovarajućeg tipa. Modeliranu mrežu treba da analiziraju alatima za obradu socijalnih mreža po izboru i izvrše vizuelizaciju mreže. Dobijene rezultate analize treba na odgovarajući način interpretirati u skladu sa postavljenim istraživačkim pitanjima.

3. POSTAVLJENI PROBLEM

U okviru ove sekcije je dat predlog projektnog zadatka za tekuću školsku godinu. Studenti mogu predložiti predmetnom nastavniku drugu temu. U tom slučaju, poželjno je priložiti i deo skupa podataka koji bi se analizirao, kako bi student na adekvatan način u saradnji sa nastavnikom postavio ciljeve istraživanja i istraživačka pitanja.

3.1. Uporedna analiza naučne produkcije iz oblasti računarstva na Univerzitetu u Beogradu

Tema projektnog zadatka u tekućoj školskoj godini je kvantitativna i kvalitativna analiza naučne produkcije iz oblasti računarstva na Univerzitetu u Beogradu (UB) na tri najvažnija fakulteta u njegovom sastavu koji se bave računarstvom, tačnije naučnih radova zaposlenih na katedrama fakulteta UB kojima je računarstvo primarna oblast. Cilj analize je da se kroz odgovarajuće bibliometrijske i naukometrijske analize, kao i kroz analizu kolaboracione, socijalne

mreže utvrdi stanje računarskih nauka na pojedinačnim fakultetima, kao i nivo saradnje između zaposlenih sa istog i različitih fakulteta koji se bave oblašću računarstva.

U okviru Univerziteta u Beogradu (UB) postoje tri fakulteta koji su matični za jednu od četiri oblasti računarstva. To su Matematički fakultet (MATF) za oblast računarskih nauka (*computer science*), Elektrotehnički fakultet (ETF) za oblast računarske tehnike (*computer engineering*) i softverskog inženjerstva (*software engineering*) i Fakultet organizacionih nauka (FON) za oblast informacionih sistema (*information systems*). U okviru navedenih fakulteta postoje katedre koje se bave pomenutim naučnim oblastima. To su Katedra za softversko inženjerstvo (FON_SI), Katedra za informacione tehnologije (FON_IT) i Katedra za informacione sisteme (FON_IS) sa Fakulteta organizacionih nauka, Katedra za računarstvo i informatiku (MATF_RTI) sa Matematičkog fakulteta i Katedra za računarsku tehniku i informatiku (ETF_RTI) sa Elektrotehničkog fakulteta.

Pored obaveza u nastavi, zaposleni na univerzitetu imaju obavezu pisanja odgovarajućih naučnih publikacija. Napredovanje u akademskoj karijeri svakog zaposlenog u mnogome zavisi od kvaliteta i obima naučne produkcije pojedinca, ali utiče i na vidljivost, reputaciju i kvalitet ustanova (fakulteta) kojima oni pripadaju, kao i samog univerziteta.

Najveći deo naučne produkcije se odvija kroz publikacije kao što su knjige, disertacije, radovi u naučnim časopisima i radovi na međunarodnim i domaćim naučnim konferencijama. Međutim, u smislu evaluacije, najviše se cene radovi u naučnim časopisima i radovi na naučno-stručnim konferencijama.

3.2. Skup podataka za analizu

U okviru ovog projektnog zadatka je potrebno analizirati naučnu produkciju na iz oblasti računarstva na Univerzitetu u Beogradu koja se odnosi na radove u časopisima i radove na naučno-stručnim konferencijama. Podaci za analizu (primarni skup podataka) su dostupni u vidu odgovarajućih Excel tabela u arhivi koja je priložena uz tekst projektnog zadatka.

Za prikupljanje podataka o naučnim publikacijama su korišćeni podaci o autorima koji su se vodili kao zaposleni u nastavnim zvanjima ili bili angažovani kao laboratorijsko osoblje na prethodno pomenutim katedrama u avgustu 2018., što je prikupljeno sa internet prezentacija ovih fakulteta. Podaci su dostupni u okviru Excel datoteke *UB_cs_authors.xlsx*. Ostali autori naučnih publikacija koji se javljaju u okviru skupa podataka za analizu nisu od interesa za analizu.

Primarni skup podataka za analizu (eng. *primary dataset*) je dobijen prikupljanjem podataka iz indeksne baze naučnih radova *Scopus* zaključno sa avgustom 2018. U skupu podataka se nalazi 1290 radova prikupljenih na osnovu dostupnih podataka o zaposlenima na ranije pomenuta tri fakulteta UB. Podaci su u određenoj meri prečišćeni, ali određeni podaci mogu biti nepotpuni ili u neodgovarajućem obliku. Podaci se sastoje od sledećih kolona:

- UB zaposleni – ime zaposlenog kome je pridružena odgovarajuća publikacija u *Scopus* bazi
- Naslov – naslov publikacije
- Godina – godina izdavanja
- Autori – autori rada
- Broj citiranja – broj citata drugih autora prema *Scopus* bazi
- Tip rada – tip publikacije
- Cite factor – *CiteScore* bibliometrijski faktor
- SJIR – *SCImago Journal Rank* bibliometrijski faktor
- SNIP – *Source Normalized Impact per Paper* bibliometrijski faktor
- Ime dokumenta – ime dokumenta (časopisa ili zbornika radova) u kome je publikacija objavljena

Na osnovu primarnog skupa podataka treba formirati sekundarni skup podataka (eng. *secondary dataset*) koji predstavlja prečišćenu verziju podataka za analizu. Prečišćavanje izvršiti prema potrebama zadatka i ciljevima istraživanja. U analizi su od interesa samo publikacije tipa *Article*, *Conference Paper*, *Article in Press*, *Review*, *Book Chapter*, dok se ostale mogu zanemariti. Prilikom prečišćavanja se mogu izostaviti svi nepotrebni podaci.

3.3. Modelovanje mreže

Sekundarni skup podataka je potrebno iskoristiti za modelovanje odgovarajućih socijalnih mreža. Potrebno je modelovati mrežu koautora (na nivou katedri, fakulteta i celog univerziteta) i mrežu časopisa. Prilikom modelovanja mreže implementirati odgovarajući tip mreže (usmerena, neusmerena, težinska i sl.) u skladu sa postavljenim istraživačkim pitanjima i ciljevima. Primarna mreža za analizu treba da bude mreža koautora na nivou univerziteta, a preostale treba iskoristiti u

funkciji odgovaranja na određena istraživačka pitanja. Po potrebi se mogu napraviti i analizirati i druge mreže na osnovu zadatog skupa podataka.

U okviru mreže koautora, zaposleni na fakultetima UB koji se bave računarstvom treba da predstavljaju čvorove mreže, a vezu između dva čvora treba uspostaviti ukoliko su dva zaposlena napisala bar jedan naučni rad zajedno. Težinu grane (veze) modelovati celovitim ili frakcionim brojanjem.

U okviru mreže časopisa, časopis treba da predstavlja čvor mreže, a vezu između dva čvora treba uspostaviti ukoliko postoji bar jedan autor koji je objavljivao u oba časopisa. Zbornike radova sa konferencija ne treba uzimati u obzir prilikom modelovanja ove mreže.

3.4. Istraživačka pitanja i ciljevi

Prilikom obrade primarnog i sekundarnog skupa podataka pogodno je kao smernice koristiti prethodno definisana istraživačka pitanja. U okviru ove sekcije je postavljen jedan broj takvih pitanja, a studenti treba da na osnovu analize problema i samih podataka definišu dodatna pitanja ili specijalizuju navedena čime mogu bliže usmeriti samu analizu.

- 1) Koliki je prosečan broj koautora po svakom autoru?
- 2) Ko su najproduktivniji naučnici iz oblasti računarstva i kom fakultetu i katedri pripadaju?
- 3) Koje katedre i fakulteti su najproduktivniji posmatrajući zasebno naučnu produkciju u časopisima i na konferencijama?
- 4) Koje zajednice ili naučnoistraživačke grupe se mogu uočiti prilikom analize mreže?
- 5) U kojoj meri autori imaju tendenciju da pišu publikacije sa istim koautorima?
- 6) Koji autori predstavljaju centre okupljanja u okviru svojih katedri i fakulteta?
- 7) Koji autori predstavljaju jezgro mreže?
- 8) Da li autori sa različitih fakulteta i katedri međusobno sarađuju i u kojoj meri?
- 9) Kakva je saradnja među autorima sa istog fakulteta?
- 10) Ko su autori koji povezuju različite grupe u okviru mreže?
- 11) Kolika je gustina svake od modelovanih mreža?
- 12) U kojoj meri su mreže povezane i centralizovane?
- 13) Koje su prosečne distance u okviru modelovanih mreža?
- 14) Kakva je distribucija čvorova po stepenu i da li prati neku zakonomernost?

- 15) Da li mreža koautora iskazuje osobine malog sveta?
- 16) Da li postoje razlike između fakulteta i katedri u smislu obima i učestanosti publikovanja u časopisima i na konferencijama?
- 17) Da li postoji razlika u prosečnom broju autora po radovima u časopisima i na konferencijama?
- 18) U kojim časopisima i na kojim konferencijama se u proseku najviše objavljuje?
- 19) Da li se izdvajaju i koje grupe (zajednice) časopisa u okviru mreže časopisa? Da li autori imaju tendenciju da objavljuju u časopisima sličnog profila?
- 20) U kojim godinama su autori bili najproduktivniji po fakultetima i katedrama?

Da bi se odgovorilo na postavljena pitanja, potrebno je primeniti odgovarajuće mere i metode za analizu mreže ili statističke metode. Mrežu bi trebalo karakterisati kako kroz osnovna svojstva mreže, tako i kroz složenije mere centralnosti i metode za detekciju komuna. Mere i metode izabrati prema adekvatnosti spram postavljenog problema.

3.5. Preporučene metode i alati

Za analizu modelirane socijalne mreže se preporučuje korišćenje programskih jezika Python (NetworkX biblioteka) i R (*sna* i *igraph* paketi) ili softverskih alata Gephi, UCINET, ili Pajek. Obrada primarnog skupa podataka se može obaviti pomoću MS Excel alata ili pisanjem odgovarajućih skripti u programskom jeziku po izboru. Ukoliko nije moguće drugačije, razrešavanje eventualnih dvosmislenosti u primarnom skupu podataka izvršiti ručno.

Vizuelizacija mreže se može obaviti korišćenjem alata Gephi, NodeXL ili kroz podršku u okviru programskih jezika Python (*matplotlib*, *graphviz* i *graph-tool* biblioteke) i R (*igraph* paket).

4. REZULTATI

Projektni zadatak se predaje u vidu pisanog izveštaja koji sadrži rezultate sprovedene analize i pisana objašnjenja uočenih fenomena. Uz izveštaj se dostavljaju i odgovarajuće dopunske datoteke, kao što su tabele sa rezultatima analize, izvorni programski kod skripti ili programa korišćenih u analizi, datoteke koje sadrže produkovane vizuelizacije i sl. Potpuno odsustvo dopunskih datoteke koje predstavljaju rezultate rada može povući umanjeње broja poena na

projektom zadatku. Za pisanje izveštaja se može koristiti šablon koji se nalazi u odgovarajućoj sekciji na sajtu predmeta. Preporučeni obim izveštaja je do 10 stranica teksta.

5. PREDAJA, ODBRANA I VREDNOVANJE

Projektni zadatak se predaje elektronskim putem najkasnije do termina ispita u odgovarajućem ispitnom roku na način kako to bude specificirao predmetni nastavnik. Na odbranu je potrebno doneti štampanu verziju izveštaja. Po pravilu, projektni zadatak se brani pred predmetnim nastavnikom u ispitnom roku u kome student želi da polaže ispit. Ukoliko student želi da brani zadatak u nekom drugom terminu, treba o tome da blagovremeno obavesti predmetnog nastavnika, radi eventualnog dogovora. Ukoliko se projektni zadatak radi u paru, studenti zajedno brane projektni zadatak.

Projektni zadatak nosi 40 poena. Poeni sa jednom odbranjenog projektnog zadatka važe jednu školsku godinu. Postoji mogućnost da se dobro urađeni projektni zadaci prošire u završni, master rad.