

UNIVERZITET U BEOGRADU  
ELEKTROTEHNIČKI FAKULTET



## **ANALIZA SOCIJALNIH MREŽA**

Projektni zadatak

Verzija 1.0

**Predmetni nastavnici:**

dr Marko Mišić, vanredni profesor

dr Jelica Protić, redovni profesor

**Predmetni saradnik:**

Predrag Obradović, asistent

**Školska godina:**

2024/2025.

Beograd, januar 2025.

# SADRŽAJ

<b>SADRŽAJ</b> .....	<b>2</b>
<b>1. UVOD</b> .....	<b>3</b>
<b>2. CILJ</b> .....	<b>3</b>
<b>3. POSTAVLJENI PROBLEM</b> .....	<b>3</b>
3.1. ANALIZA KRIMINALNIH MREŽA SICILIJANSKE MAFIJE U OKVIRU OPERACIJE „MONTANJA“ .....	3
3.2. SKUP PODATAKA ZA ANALIZU .....	4
3.3. MODELOVANJE MREŽA.....	5
3.4. ISTRAŽIVAČKA PITANJA I CILJEVI .....	6
3.4.1. Osnovna karakterizacija modelovanih mreža (10/20 poena).....	7
3.4.2. Analiza mera centralnosti (4/20 poena).....	7
3.4.3. Detekcija komuna Luvenskom metodom (2/20 poena).....	8
3.4.4. Detekcija komuna spektralnim klasterisanjem (2/20 poena) .....	8
3.4.5. Poređenje modelovanih mreža (2/20 poena) .....	8
3.5. PREPORUČENE METODE I ALATI.....	8
<b>4. REZULTATI</b> .....	<b>8</b>
<b>5. PREDAJA, ODBRANA I VREDNOVANJE</b> .....	<b>9</b>
<b>LITERATURA</b> .....	<b>9</b>

# 1. UVOD

U okviru ovog dokumenta su data uputstva za izradu projektnog zadatka na predmetu Analiza socijalnih mreža (13M111ASM) u školskoj 2024/2025. 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 i predstaviti u obliku izveštaja.

## 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. Analiza kriminalnih mreža sicilijanske mafije u okviru operacije „Montanja“

Sicilijanska mafija, poznata i kao „Cosa Nostra“, nastala je sredinom 19. veka na Siciliji kao tajno društvo sa ciljem pružanja zaštite i rešavanja sporova u ruralnim područjima gde je zvanična policija bila slabo prisutna. Vremenom se razvila u visoko organizovanu kriminalnu mrežu uključenu u razne ilegalne aktivnosti, poput iznude, krijumčarenja i trgovine drogom. Uticaj mafije se duboko ukorenio u sicilijanskoj politici, ekonomiji i kulturi, čineći je moćnom silom kako na lokalnom, tako i na međunarodnom nivou. Njena hijerarhijska struktura, kodeks ćutanja („omerta“) i strateška

upotreba nasilja omogućili su joj da opstane i prilagodi se uprkos brojnim pokušajima policije da razbije njene operacije.

Predmet ovogodišnjeg projektnog zadatka je kvantitativna i kvalitativna analiza strukture i funkcionisanja kriminalnih mreža sicilijanske mafije, s posebnim osvrtom na mreže povezane sa „Operacijom Montanja“. U tom kontekstu, potrebno je analizirati komunikacije i fizičke interakcije osumnjičenih lica, kao i uloge pojedinaca unutar posmatranih mafijaških klanova. Fokus istraživanja će biti na dve glavne kriminalne grupe, porodici „Mistretta“ i klanu „Batanesi“, koje su bile predmet pomenute antimafijaške operacije.

### **3.2. Skup podataka za analizu**

Kao što je prikazano u referentnim radovima koji se bave analizom mreža kriminalaca [1][2][3][4], sicilijanska mafija se po strukturi i metodima operisanja razlikuje od drugih kriminalnih organizacija. Mafijaške mreže imaju specifične karakteristike, zbog distribucije i jačine veza, što ih čini izuzetno otpornim na egzogene poremećaje, pa ispoljavaju drugačiju dinamiku na njih se ne mogu direktno primeniti prethodna znanja o drugim socijalnim mrežama sličnih dimenzija.

Analizirani skup podataka [5] kreiran u okviru istraživanja [1] i [2] proizašao je iz naloga za pritvor koji je 14. marta 2007. godine izdao istražni sudija Suda u Mesini, pred kraj velike antimafijaške operacije poznate kao „Operacija Montanja“. Ovu operaciju, završenu 2007. godine, sproveo je Tužilački ured u Mesini (Sicilija), uz angažman Specijalne operativne jedinice italijanske policije (R.O.S. – Reparto Operativo Speciale karabinjera, specijalizovane za antimafijaške istrage). Konkretno, ova istraga predstavljala je značajnu operaciju usmerenu na dva mafijaška klana poznata kao porodica „Mistretta“ i klan „Batanesi“. Od 2003. do 2007. godine, utvrđeno je da su ove porodice infiltrirale nekoliko privrednih aktivnosti, uključujući velike infrastrukturne projekte, putem kartela preduzetnika povezanih sa sicilijanskom mafijom.

Skup podataka se sastoji iz 3 *Excel* datoteke *Montagna\_Meetings\_Edgelist.csv*, *Montagna\_Phone\_Calls\_Edgelist.csv*, *Montagna\_Roles.csv*. „Meetings“ (Sastanci) obuhvata fizičke sastanke među osumnjičenima, dok se „Phone Calls“ (Telefonski pozivi) odnosi na telefonske razgovore među pojedincima. Osim toga, dostupan je i ograničeni pregled uloga posmatranih članova mafije i pojedinaca povezanih sa njima – „Roles“ (Uloge).

Čvorovi mogu pripadati različitim kategorijama. Neki čvorovi predstavljaju vođe (tj. „šefove“) ili vojnike (tj. „picciotti“, sicilijanska reč koja označava najniži rang u hijerarhiji mafije)

kriminalne organizacije. Drugi čvorovi su povezani s pojedincima (npr. prodavac voća ili pekar) koji su imali jedan ili više poziva sa članovima mafijaškog sindikata, ali nisu povezani ni sa jednom kriminalnom organizacijom.

Na osnovu primarnog skupa podataka po potrebi formirati sekundarni skup podataka (eng. *secondary dataset*) koji predstavlja prečišćenu verziju podataka za analizu. Prečišćavanje izvršiti prema potrebama zadatka i ciljevima istraživanja. Prilikom prečišćavanja se mogu izostaviti svi nepotrebni podaci.

### 3.3. Modelovanje mreža

Sekundarni skup podataka iz ove analize koristiće se za modelovanje odgovarajućih socijalnih mreža. U zavisnosti od toga da li se projekat radi samostalno ili u paru, postoje dve različite varijante projektnog zadatka. Ako studenti rade projekat samostalno ili u paru za maksimalno 20 od 40 poena, potrebno je da se fokusiraju na modelovanje samo mreža telefonskih poziva (TNet) i fizičkih sastanaka (SNet) i izvrše analize na osnovu tih mreža (potpoglavlje „Modelovanje mreža TNet i SNet“). S druge strane, studenti koji rade projekat u paru i žele da osvoje maksimalnih 40 poena, treba da modeluju i analiziraju pet agregiranih mreža (ONet1 - ONet5) i da koriste različite pondere u procesu agregacije kako bi dodatno istražili uticaj svake mreže na strukturu objedinjene mreže (potpoglavlje „Modelovanje mreža ONet1-5“).

#### *Modelovanje mreža TNet i SNet*

Studenti koji rade projekat samostalno, ili u paru za maksimalno 20 od 40 poena, potrebno je da se fokusiraju na modelovanje mreža telefonskih poziva (TNet) i fizičkih sastanaka (SNet). Ove mreže modelovati na osnovu dostupnih podataka, koristeći odgovarajuće tipove mreža (usmerena, neusmerena, težinska, itd.) u skladu sa postavljenim istraživačkim pitanjima i ciljevima analize.

- **Mreža fizičkih sastanaka (SNet):** U ovoj mreži, osumnjičeni predstavljaju čvorove, a veza između dva čvora se uspostavlja ukoliko su dva osumnjičena zabeležena na istom fizičkom sastanku.
- **Mreža telefonskih poziva (TNet):** U ovoj mreži, veza između dva čvora se uspostavlja ukoliko su dva osumnjičena ostvarila telefonski kontakt.

Za analizu, potrebno je koristiti samo mreže TNet i SNet. Agregacija ovih mreža u zajedničku mrežu (ONet) nije potrebna za ovaj nivo zadatka.

### ***Modelovanje mreža ONet1-5***

Studenti koji rade projekat u paru za maksimalnih 40 poena, potrebno je da agregacijom podataka o telefonskim pozivima i fizičkim sastancima kreiraju pet objedinjenih mreža (ONet1-ONet5). Suštinski, mreža ONet se dobija linearnom kombinacijom mreža TNet i SNet:

$$w_{ij}^O = a \cdot w_{ij}^T + b \cdot w_{ij}^S$$

gde su  $w_{ij}^T$  i  $w_{ij}^S$  težine veza između čvorova  $i$  i  $j$  u mrežama TNet i SNet, dok su  $a$  i  $b$  koeficijenti ponderisanja koji odražavaju relativni značaj svake mreže u agregaciji. Potrebno je kreirati 5 mreža ONet1 – ONet5 za sledeće vrednosti koeficijenata ponderisanja:

$$(a, b) \in \{(0,1), (0.25, 0.75), (0.5, 0.5), (0.75, 0.25), (1,0)\}$$

Obratiti pažnju da ONet1 i ONet5 odgovaraju mrežama sastanaka i telefonskih poziva, respektivno, mreže ONet2 i ONet4 su mreže u kojim se veći značaj pridaje vezama nastalim na osnovu sastanaka, odnosno na osnovu telefonskih poziva, dok se u mreži ONet2 telefonski pozivi i sastanci tretiraju ekvivalentno.

Izabrane vrednosti koeficijenata  $a$  i  $b$  omogućavaju detaljno ispitivanje uticaja svake mreže na strukturu objedinjene mreže i identifikaciju ključnih veza i čvorova unutar kriminalne organizacije. Objedinjene mreže omogućavaju sveobuhvatnu analizu integrisanjem svih informacija, pružajući širu perspektivu o obimu i strukturi interakcija unutar posmatrane kriminalne mreže, dok će testiranje različitih vrednosti težinskih koeficijenata doprineti robusnosti zaključaka i identifikaciji ključnih obrazaca interakcija.

### **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, nakon što odgovore na ova pitanja, na osnovu analize problema i samih podataka definišu dodatna pitanja ili specijalizuju navedena čime mogu bliže usmeriti samu analizu. Odgovore na pitanja treba dati u formi specificiranoj u poglavlju 4.

Pitanja su grupisana u kategorije po tematici koju obrađuju i tehnikama analize koje se u njima sprovode. Ako nije drugačije navedeno, svako od pitanja se odnosi na svaku od konstruisanih mreže, ukoliko je za nju smisljeno i primenljivo. Studentima se predlaže da prvo sve analize sprovedu za jednu od mreža, pa zatim pređu na sledeću i da svoje odgovore na taj način strukturiraju i izlože u izveštaju o projektu.

Da bi se odgovorilo na postavljena pitanja, potrebno je primeniti odgovarajuće mere i metode za analizu mreže ili statističke metode. Mreže 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. Tamo gde se očekuje odgovor u obliku neke vrste rangiranja, navesti listu od 5 do 10 najrelevantnijih rezultata.

#### **3.4.1. Osnovna karakterizacija modelovanih mreža (10/20 poena)**

- 1) Kolika je gustina mreže?
- 2) Kolike su prosečne distance u okviru mreže i dijametar mreže?
- 3) U kojoj meri je mreža povezana i centralizovana? Navesti broj i veličine povezanih komponenata i proceniti da li postoji gigantska komponenta.
- 4) Koliki je prosečni, a koliki globalni koeficijent klasterizacije mreže? Kakva je raspodela lokalnog koeficijenta klasterizacije njenih čvorova? Da li je klasterisanje izraženo ili ne? Odgovor dati upoređivanjem sa slučajno generisanim *Erdos-Renyi* i *scale free* mrežama istih dimenzija.
- 5) Proceniti da li mreža iskazuje osobine malog sveta.
- 6) Izvršiti asortativnu analizu po stepenu čvora i dati odgovor da li je i koliko izraženo asortativno mešanje. Priložiti i vizuelizaciju.
- 7) Da li mreža ispoljava fenomen kluba bogatih (eng. *rich club phenomenon*)? Analizu sprovesti poređenjem sa adekvatnom mrežom dobijenom Havel-Hakimi algoritmom.
- 8) Kakva je distribucija čvorova po stepenu i da li prati *power law* raspodelu?

#### **3.4.2. Analiza mera centralnosti (4/20 poena)**

- 9) Sprovesti analize centralnosti po stepenu, bliskosti i relacionoj centralnosti. Dati pregled najvažnijih aktera po svakoj od njih.
- 10) Ko su najvažniji akteri po centralnosti po sopstvenom vektoru? Šta nam to govori o njima?
- 11) Na osnovu prethodna dva pitanja predložiti i konstruisati heuristiku (kompozitnu meru centralnosti) za pronalaženje najvažnijih aktera i pronaci ih. Obratiti pažnju na tip mreže koji se analizira (usmerena ili neusmerena) i, shodno tome, prilagoditi koliko različite mrežne metrike utiču na heuristiku.

### 3.4.3. *Detekcija komuna Luvenskom metodom (2/20 poena)*

- 12) Sprovesti klasterisanje Luvenskom metodom (maksimizacijom modularnosti) u alatu *Gephi* za tri različite vrednosti parametra rezolucije. Konstruisati vizuelizacije i diskutovati izbor parametra rezolucije na dobijeno klasterisanje (broj i veličina klastera).
- 13) Koje zajednice (komune) se mogu uočiti prilikom analize mreže, a koji akteri su ključni brokeri? Da li postoji neko objašnjenje za detektovane komune, recimo na osnovu dostupnih informacija o ulogama aktera?

### 3.4.4. *Detekcija komuna spektralnim klasterisanjem (2/20 poena)*

- 14) Sprovesti spektralnu analizu i proceniti potencijalne kandidate za broj komuna u mreži. Uporediti rezultat sa dendogramom konstruisanim *Girvan-Newman* metodom.
- 15) Ko su akteri koji se mogu okarakterisati kao ključni brokeri (mostovi) u mreži? Šta ih čini brokerima? Porediti odgovor sa ključnim čvorovima dobijenim analizom centralnosti.

### 3.4.5. *Poređenje modelovanih mreža (2/20 poena)*

- 16) Sprovesti poređenje rezultata obrade modelovanih mreža u smislu definisanja glavnih aktera, karakterističnih obrazaca interakcije i sl.

## 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, Cytoscape 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 umanjene broja poena na projektnom 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 ili saradnikom 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 maksimalno 40 poena, a razlike u zahtevima između različitih verzija projekta su opisane u sekciji 3.3. Raspodela poena po tematskim oblastima je prikazana u sekciji 3.4. Studenti ne moraju realizovati sve zahteve u okviru projekta pre odbrane. Ne postoji minimalan broj poena koji je potrebno osvojiti na projektnom zadatku da bi se položio ispit.

Poeni sa jednom odbranjenog projektnog zadatka važe jednu školsku godinu. Postoji mogućnost da se dobro urađeni projektni zadaci prošire u završni, master rad. Upit u vezi sa takvom mogućnošću studenti mogu uputiti predmetnom nastavniku ili saradniku.

## **LITERATURA**

- [1] Cavallaro, L., Ficara, A., Meo, P. D., Fiumara, G., Catanese, S., Bagdasar, O., Song, W., & Liotta, A. (2020). Disrupting resilient criminal networks through data analysis: The case of Sicilian Mafia. PLOS ONE, 15(8), e0236476. <https://doi.org/10.1371/journal.pone.0236476>
- [2] Ficara, A. et al. (2020). Social Network Analysis of Sicilian Mafia Interconnections. In: Cherifi, H., Gaito, S., Mendes, J., Moro, E., Rocha, L. (eds) Complex Networks and Their Applications

VIII. COMPLEX NETWORKS 2019. Studies in Computational Intelligence, vol 882. Springer, Cham. [https://doi.org/10.1007/978-3-030-36683-4\\_36](https://doi.org/10.1007/978-3-030-36683-4_36)

- [3] Ferrara, E., De Meo, P., Catanese, S., & Fiumara, G. (2014). Detecting criminal organizations in mobile phone networks. *Expert Systems with Applications*, 41(13), 5733-5750.
- [4] Agreste, S., Catanese, S., De Meo, P., Ferrara, E., & Fiumara, G. (2016). Network structure and resilience of Mafia syndicates. *Information Sciences*, 351, 30-47.
- [5] Criminal Network: The Sicilian Mafia. "Montagna Operation" Dataset, available on: <https://zenodo.org/records/3938818>, poslednji pristup 06.01.2025.