

# Analiza socijalnih mreža

Projektni zadatak  
2017-2018.

Marko Mišić, Jelica Protić

13M111ASM

2017/2018.

# Uvod

---

- Projektni zadatak na predmetu  
Analiza socijalnih mreža se sastoji od dva zadatka
- Prvi projektni zadatak nosi 20 poena
  - Analiza i prezentovanje osnovnih rezultata jednog naučno-istraživačkog rada
- Drugi projektni zadatak nosi 40 poena
  - Analiza konkretnog istraživačkog problema
  - Realan skup podataka

# Cilj projektnog zadatka

---

- Praktična primena stečenog teorijskog znanja na primeru konkretnog istraživačkog problema
- Prikupljanje, obrada i preliminarna analiza primarnog (sirovog) skupa podataka
- Modeliranje problema mrežom odgovarajućeg tipa
- Analiza modelirane mreže alatima za obradu socijalnih mreža po izboru
- Vizuelizaciju mreže
- Interpretacija rezultata u skladu sa postavljenim istraživačkim pitanjima

# Prvi projektni zadatak

---

- Radovi za analizu su dostupni u arhivi na sajtu predmeta
  - Lozinka za pristup je saopštena na listi elektronske pošte
- Prilikom obrade konkretnog rada, odnosno problema treba obratiti pažnju na sledeće:
  - Oblast problema koji se rešava
  - Motivaciju za upotrebu tehnika za analizu socijalnih mreža i modeliranje problema mrežom
  - Način modeliranja problema mrežom
    - Tip grafa, način uspostavljanja veza i sl.
  - Primenjene tehnike i metode za analizu
    - Osnovna svojstva, centralnosti, klasterizacija...
  - Glavne rezultate analize
  - Vizuelizaciju (ako postoji)

# Drugi projektni zadatak

---

- Analiza naučne produkcije na ETF
  - U periodu od 2000. do 2016. godine
- ETF je nastavna, ali i naučno-istraživačka institucija
- Zaposleni na fakultetu imaju obavezu pisanja odgovarajućih naučnih publikacija
  - Radovi u časopisima
  - Radovi na međunarodnim i nacionalnim konferencijama
  - Knjige, disertacije, tehnička rešenja, patentni
- Napredovanje u akademskoj karijeri u mnogome zavisi od kvaliteta i obima naučne produkcije pojedinca
- Naučna produkcija na fakultetu se odvija u okviru većeg broja naučnih polja

# Skup podataka za analizu

---

- Primarni skup podataka čine podaci o:
  - Radovima u naučnim časopisima
    - Sa i bez *impact* faktora
  - Radovima na međunarodnim konferencijama
  - Radovima na domaćim konferencijama
  - Zaposlenima na fakultetu
    - U periodu 2000. do 2017. godine
- Primarni skup podataka je dobijen upitom u odgovarajuću bazu u fakultetskom informacionom sistemu
  - Dostupan je i na sajtu fakulteta

# Problemi primarnog skupa podataka

---

- Širi vremenski period
- Parsiranje autorstava
- Razrešavanje imena
  - Nejednak unos imena i prezimena autora
  - Osobe sa istim imenima na više katedri
    - Razrešavanje po oblasti?
  - Upotreba ćirilice i latinice
- Nejednak format podataka o radovima u časopisima i na konferencijama

# Obrada primarnog skupa podataka

---

- Potrebna obrada i čišćenje primarnog skupa podataka
  - Excel za osnovna filtriranja
  - Skripte (Python, Java, R...) za formiranje mreže
    - Upotreba *hash* mapa ili *dictionary*-ja
  - Ručno razrešavanje kod dvosmislenosti
- Izbor međuformata i izlaznog formata
  - CSV ili XLS kao međuformati
  - Neki od formata koji je podržan u alatima za analizu
    - GML, GEXF, UCINET DL
    - Varijante *edge* liste



# Modelovanje mreže

---

- Koautorska težinska mreža
  - Dva zaposlena na fakultetu će biti povezana ukoliko su napisala bar jedan rad zajedno
  - Težinu grane će predstavljati broj zajedničkih radova
    - Jedinična težina ili frakciono brojanje
- Pitanja:
  - Da li praviti odvojene mreže za radove u časopisima i na međunarodnim i domaćim konferencijama?
    - Jedna nasuprot tri mreže
  - Kako pamtiti pripadnost katedri?
    - Čuvati odgovarajući atribut čvora
  - Šta raditi sa koautorima koji nisu sa ETF?
    - Eliminirati ili ostaviti?

# Istraživačka pitanja (1)

---

- Koliki je prosečan broj koautora po svakom naučniku?
- Ko su najproduktivniji naučnici na ETF i kojoj oblasti pripadaju?
- Koje zajednice ili naučnoistraživačke grupe se mogu uočiti prilikom analize mreže?
- Koji naučnici predstavljaju centre okupljanja u okviru svojih oblasti?
- Koji naučnici predstavljaju jezgro mreže?

# Istraživačka pitanja (2)

---

- Da li naučnici iz različitih oblasti međusobno saraduju i u kojoj meri?
- Ko su naučnici koji povezuju različite grupe u okviru mreže?
- Kolika je gustina mreže?
- U kojoj meri je mreža povezana i centralizovana?
- Koje su prosečne distance u okviru mreže?
- Kakva je distribucija čvorova po stepenu i da li prati neku zakonomernost?

# Istraživačka pitanja (3)

---

- Da li postoje razlike između naučnoistraživačkih oblasti u smislu obima i učestanosti publikovanja u časopisima i na konferencijama?
- Da li postoji razlika u prosečnom broju autora po radovima u časopisima i na konferencijama?
- U kojim časopisima i na kojim konferencijama se u proseku najviše objavljuje?
- U kojim godinama su naučnici bili najproduktivniji po oblastima?

# Metode i alati za analizu

---

- Alati za analizu modelirane socijalne mreže
  - UCINET, Gephi
  - Programski jezik R
    - *sna* i *igraph* paketi
- Vizuelizacija mreže
  - Gephi, NodeXL
  - Programski jezik R
    - *igraph* paket

# Tehnički detalji

---

- Radi se samostalno ili u paru
- Predaja pred ispit elektronskom poštom
- Rezultati u obliku pisanog izveštaja
  - Preporučeni obim do 10 stranica teksta
  - Dopunske datoteke sa rezultatima rada, kodom i sl.
  - Šablon dostupan na sajtu predmeta
- Poeni važe jednu školsku godinu

# Literatura

---

- D. Hansen, B. Shneiderman, M. Smith, Analyzing Social Media Networks with NodeXL: Insights from a Connected World, Morgan Kaufmann, 2010.
- Charles Kadushin, Understanding Social Networks: Theories, Concepts and Findings, Oxford University Press, 2012
- Christina Prell, Social Network Analysis: History, Theory and Methodology, SAGE Publications Ltd, 2012.
- Hanneman, Robert A. and Mark Riddle, Introduction to social network methods, University of California, Riverside, 2005.