

Analiza socijalnih mreža

Pojam i predstavljanje socijalnih mreža

Marko Mišić, Jelica Protić

13M111ASM

2022/2023.

Teorija grafova

- Deo matematike sa dugom istorijom i primenama u drugim naukama
 - Problem konjićevog puta (*Knight's tour*) 9. vek
 - *Euler*-ov problem sedam mostova u *Königsberg*-u (1736. godina)
 - Kirhofovi zakoni (1845. godina)
 - Teorija hemijske strukture (1857. godina)
 - Upotreba u računarstvu (50-te godine XX veka)
 - Najkraća rastojanja, minimalno obuhvatno stablo, maksimizacija protoka

Teorija mreža (1)

- Teorija mreža (eng. *Network science*) bavi se proučavanjem usmerenih i neusmerenih grafova kojima se modeluju odnosi između diskretnih objekata (entiteta)
 - Kombinuje znanja iz teorije grafova, statističke fizike, *data mining*-a, vizualizacije informacija i sociologije
- Mreža se predstavlja pomoću skupa čvorova i grana
 - Diskretni objekti predstavljaju čvorove
 - Odnosi se modeluju granama grafa
 - U graf se unose i dodatne osobine (težine grana i sl.)

Teorija mreža (2)

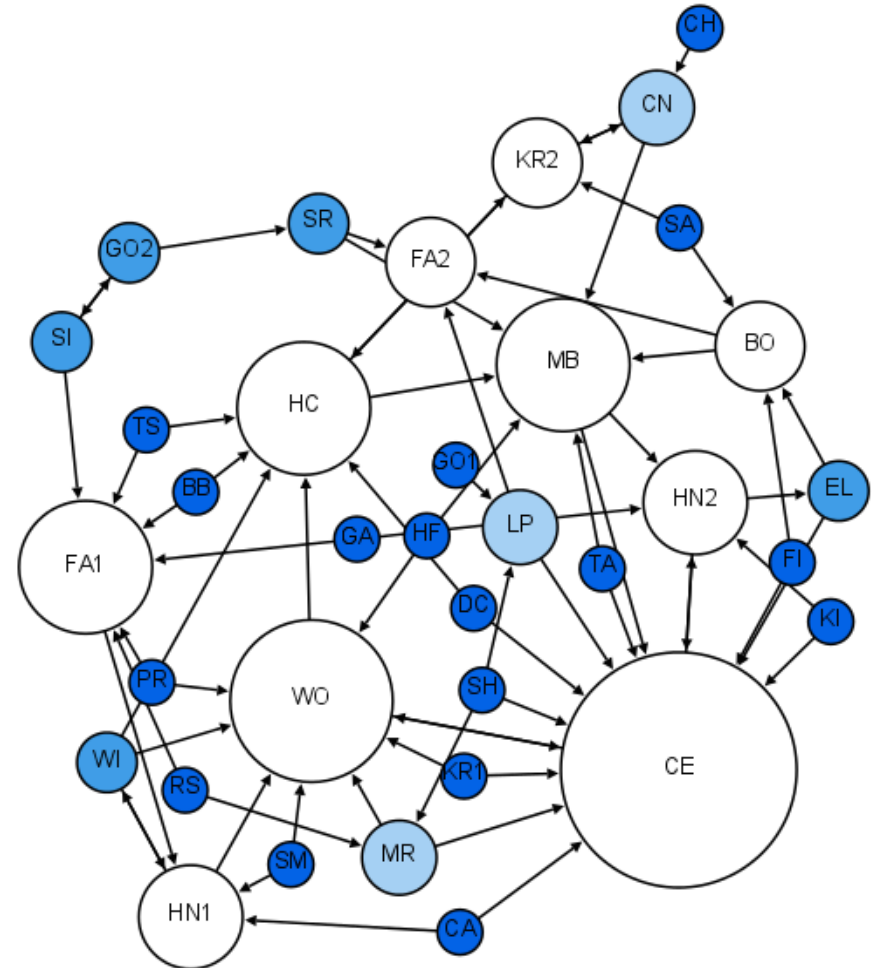
- „*The study of network representations of physical, biological, and social phenomena leading to predictive models of these phenomena*“
 - US National Research Council definicija
- Primena u prirodnim i društvenim naukama
 - Fizika, hemija, računarstvo, elektrotehnika, biologija, ekonomija, sociologija
- Prve ideje iz 1930. (Moreno) u sociologiji
- *Erdős–Rényi random graph* model (1959.)
- *Barabási scale-free* mreže (1999.)

Teorija mreža (3)

- Pojam kompleksne mreže (*complex network*) definiše graf sa netrivialnim topološkim odlikama
 - Ne pojavljuju se u jednostavnim mrežama, poput rešetke (*lattice, mesh*) ili *random* grafa
 - Pojavljuju se u računarskim, biološkim, tehničkim, neuronskim i socijalnim mrežama
- Tipični primeri:
 - *Scale-free* mreže
 - *Small-world* mreže
- Mlada i perspektivna grana nauke
 - Intenzivnija izučavanja od 2000. godine

Morenov sociogram

- Socijalna struktura razreda u osnovnoj školi
 - Dečaci su prijatelji samo dečaka
 - Devojčice su prijatelji samo devojčicama
 - Jedan jedini dečak je rekao da mu se sviđa jedna devojčica
 - Osećaj nije bio uzvraćen



Pojam socijalne mreže

- Socijalna mreža $\neq$ *Facebook, LinkedIn, Twitter!*
 - Socijalna mreža je širi pojam od pojedinačnog servisa na internetu
- U najširem smislu, socijalna mreža modeluje odnose i interakcije između različitih entiteta (aktera)
- Akteri mogu biti entiteti koji predstavljaju:
 - Pojedinačne osobe
 - Društvene grupe, organizacije i institucije
 - Naučne radove
 - Geografska područja (zemlje, gradove)
- Često se nazivaju kolaboracionim mrežama

Tipična veličina

- Male socijalne mreže
 - 10^2 do 10^3 čvorova (tipično pre 2000. godine)
- Komunikacione mreže
 - 10^5 do 10^8 čvorova (oko 2000. godine)
- Društvene mreže na internetu
 - Prosečno oko 10^8 čvorova
 - Današnja Facebook mreža ima više od 10^9 korisnika
 - Mreža veb stranica ima više od 10^{12} stranica
- Veliki porast u kratkom vremenu!

Tipične primene (1)

- U širokom dijapazonu oblasti
 - Upravljanje i menadžment
 - Poboljšanje organizacione strukture u kompanijama
 - Ekonomija
 - Poboljšanje prodaje
 - Zdravlje
 - Epidemiologija, neuronauke, analiza DNK
 - Farmakologija
 - Transport i logistika
 - Fizika i hemija

Tipične primene (2)

- Određivanje kruga prijatelja ili saradnika neke osobe
- Pronalaženje osoba od uticaja na nekom polju
- Otkrivanje načina širenja informacija unutar mreže
- Analiza *World Wide Web*-a
- Bibliometrijske i naukometrijske analize
 - Analizira naučne produkcije pojedinih naučnika ili organizacija

Tipične primene (3)

- Određivanje interakcija između različitih čestica
 - Fizika čestica i statistička fizika
- Analiziranje proteinskih i drugih struktura
 - Biološke nauke, hemija i farmakologija
- Modelovanje gena i njihovih međusobnih povezanosti
 - Genetika
- Praćenje i kontrola širenja zaraznih bolesti
 - Epidemiologija
- Planiranje mreže puteva, transport i logistika
- Istraživanje i optimizacija računarskih i telekomunikacionih mreža

Vrste socijalnih mreža (1)

- Ne postoje značajne razlike u odnosu na tradicionalne grafove
- Čvorovi grafa (*nodes*) opisuju aktere u okviru socijalne mreže
- Grane grafa (*edges, links*) predstavljaju socijalne relacije:
 - Prijateljstvo, poslovne i rodbinske veza, članstvo u organizacijama i sl.
 - Interakcije poput razmene poruka, robno-novčanih tokova i sl.
 - Neka zajednička osobina članova mreže

Vrste socijalnih mreža (2)

- Relacija između aktera može, a ne mora biti simetrična
- Kod usmerenih mreža, relacija nije simetrična
 - Uobičajeno se označava usmerenom strelicom prilikom vizuelizacije grafa
 - Pošiljaoci i primaoci poruka
 - Praćenje na *Instagram*-u i *Twitter*-u
- Kod neusmerene mreže, relacija važi u oba smera
 - Mreža prijatelja na *Facebook*-u, članstvo u organizacijama
- Postoje i mešoviti grafovi

Vrste socijalnih mreža (3)

- Relacija koju mreža modeluje može biti eksplicitno ili implicitno pristuna
- U direktnim mrežama, relacije su eksplicitno prisutne
 - Prijateljstvo na *Facebook*-u, praćenje na *Instagram*-u
 - Zajedničko autorstvo naučnog rada
- U indirektnim mrežama, relacije su implicitne
 - Prisustvo aktera istom događaju ili sceni (primer iz romana „Jadnici”)
 - *Co-occurrence* ili *co-presence* u nekoj situaciji
 - Implicitno neusmerena relacija
 - Pripadnost istoj organizaciji (studenti ETF-a)

Vrste socijalnih mreža (4)

- Relacije u mreži mogu biti višestruke
 - *Simplex* vs. *multiplex* pristup relacijama
- Mreža može predstavljati samo jedan tip relacija između aktera u mreži (*simplex* mreža)
 - Prijateljstvo, komunikaciju i sl.
- U okviru jedne mreže se može agregirati više tipova relacija
 - Komunikacija i praćenje na mrežama
 - Prijatelji i rodbina
- Problem vizuelizacije i gubitka informacija usled agregacije

Vrste socijalnih mreža (5)

- Relacija može nositi informaciju o njenom intenzitetu
 - Učestanost interakcije, broj razmenjenih poruka, broj koautorstava, razdaljinu, itd.
- Netežinski i težinski grafovi
 - *Binary*, *Signed* i *Valued* grafovi
- Kod binarnih grafova, prisustvo ili odsustvo grane čuva informaciju o relaciji
- Označeni grafovi su uobičajeno mešoviti
 - Dva tipa usmerenih grana i neusmerena grana
- Težina se tipično označava na grani prilikom vizuelizacije kod težinskih grafova
 - Mogu biti i usmereni i neusmereni

Vrste socijalnih mreža (6)

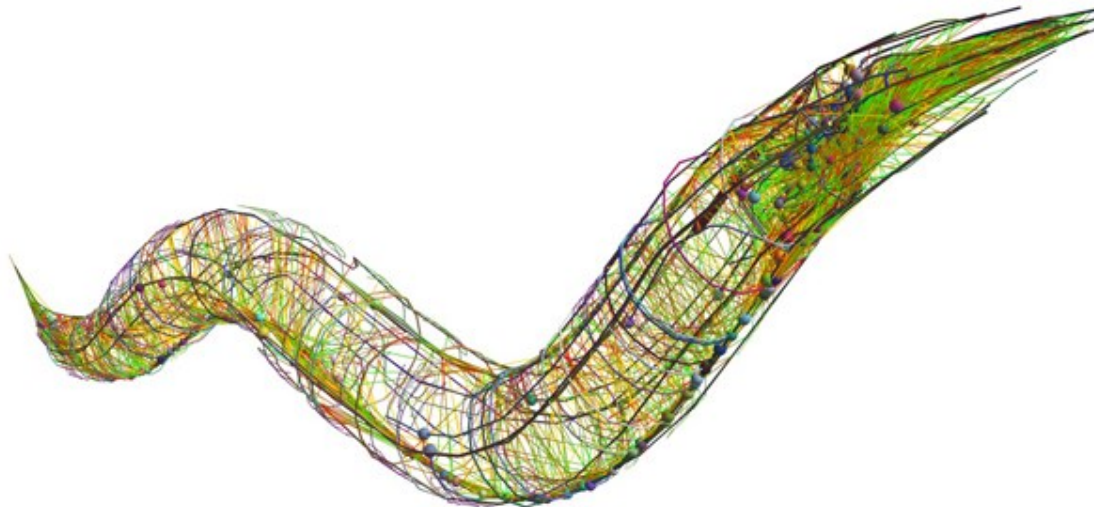
- Bipartitne mreže (*bipartite, two-mode networks*)
 - Dve vrste čvorova koji predstavljaju različite entitete
 - Relacije moguće samo između entiteta različitog tipa
 - Igrači i pozicije u timu
 - Najčešće proučavana vrsta multimodalnih mreža
 - Potreba za projekcijama, načinima za agregaciju
- Ego mreže
 - Fokusirane na izabranog aktera (*ego*) i njegove neposredne veze (*alters*)
 - Svi imaju vezu sa ego čvorom, a dozvoljene su i međusobne veze *alter-a*
 - Moguće i ego mreže sa više nivoa udaljenosti od ega

Tipični primeri (1)

- Primeri u nastavku su većinom sa INA kursa na FRI Ljubljana
- Izabrani primeri iz različitih domena
 - Neuronauke i medicina
 - Sport
 - Veb
 - Društvene mreže
 - Kolaboracije u nauci
 - Ekologija

Tipični primeri (2)

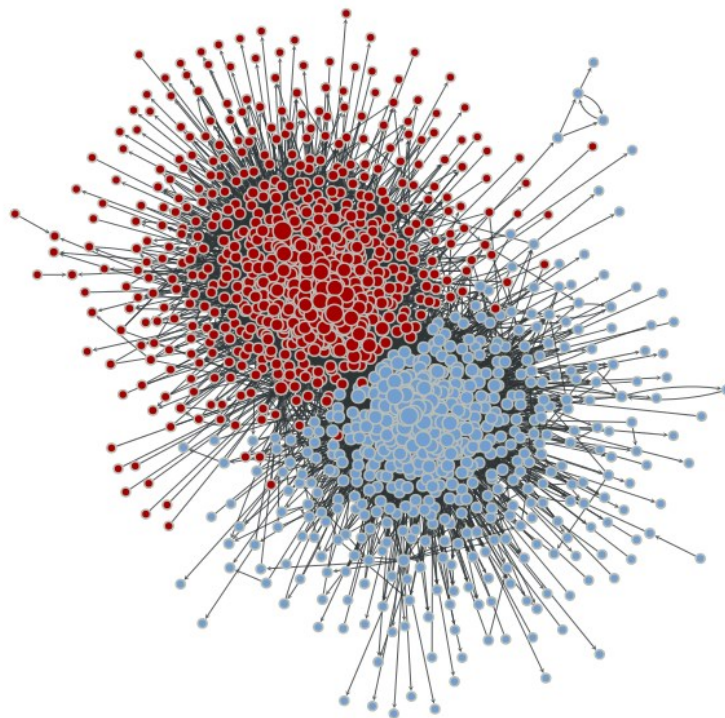
- Neuronauke modeluju veze između neurona
 - Ljudski mozak ima približno 10^{11} neurona
 - Prikazan model nematode (crva) *C. elegans*
 - Čvorovi su neuroni, veze su sinapse



Tipični primeri (3)

- World Wide Web

- Veličina veća od 10^{12} stranica
- Čvorovi su veb stranice, veze su hiperlinkovi



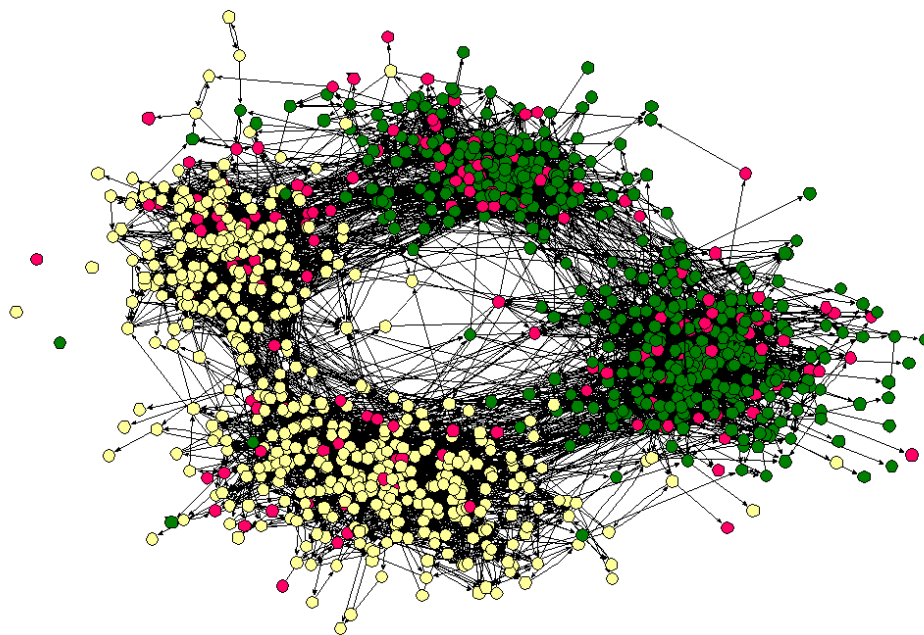
Tipični primeri (4)

- Facebook društvena mreža
 - Čvorovi su Facebook korisnici, veze prijateljstva



Tipični primeri (5)

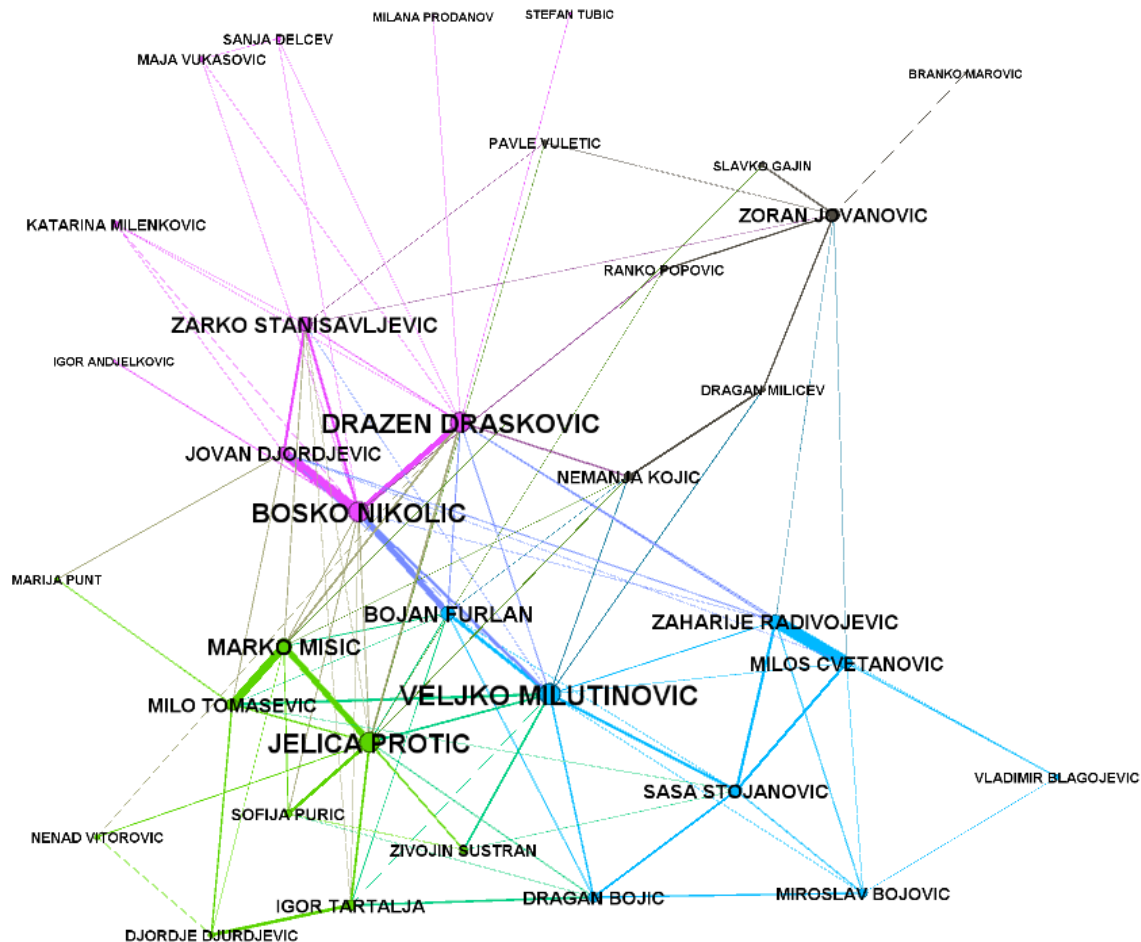
- Društvena mreža iz realnog sveta
 - Na osnovu realnih anketa
 - Čvorovi su školska deca, veze prijateljstva



Tipični primeri (6)

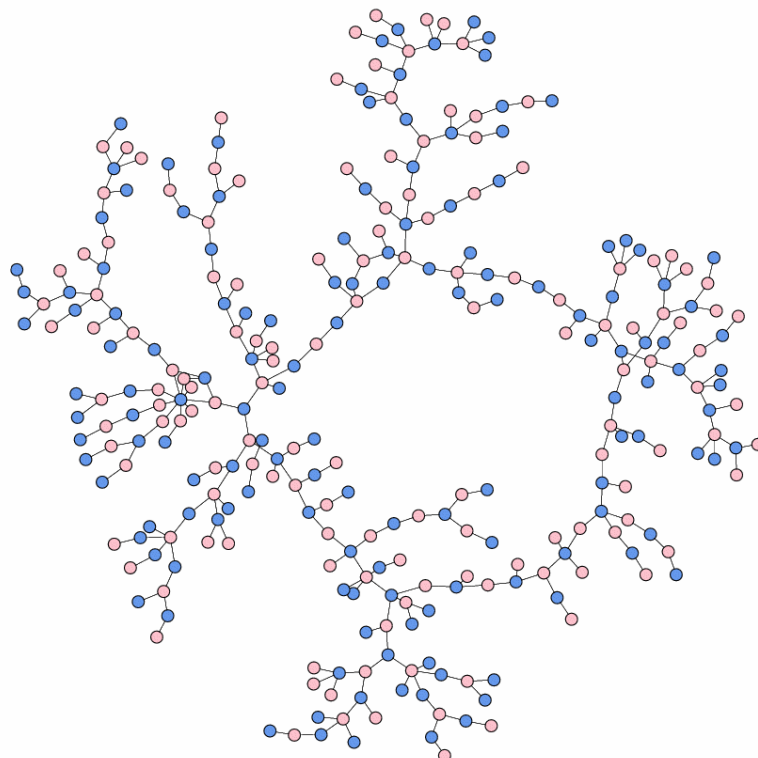
○ Kolaboraciona mreža

- Naučna produkcija zaposlenih na Katedri RTI
- Čvorovi su zaposleni
- Veze postoje ukoliko su pisali rad zajedno



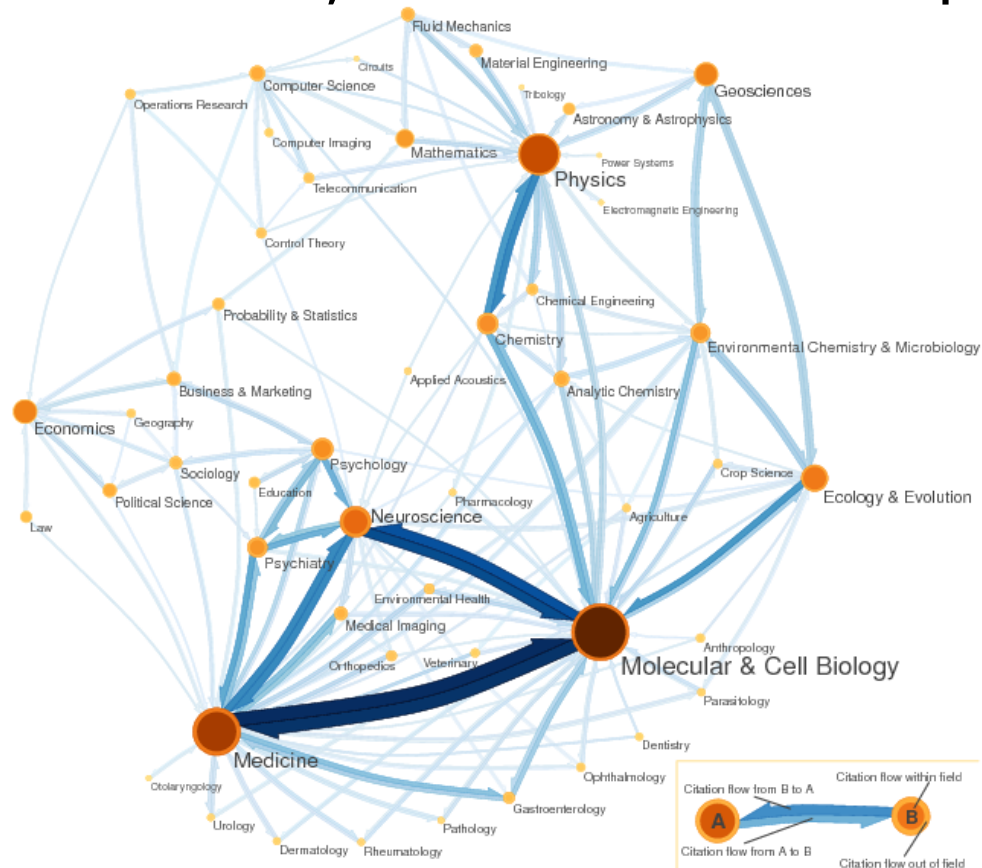
Tipični primeri (7)

- Mreža seksualnih kontakata
 - Čvorovi su muškarci i žene
 - Veze predstavljaju seksualne kontakte



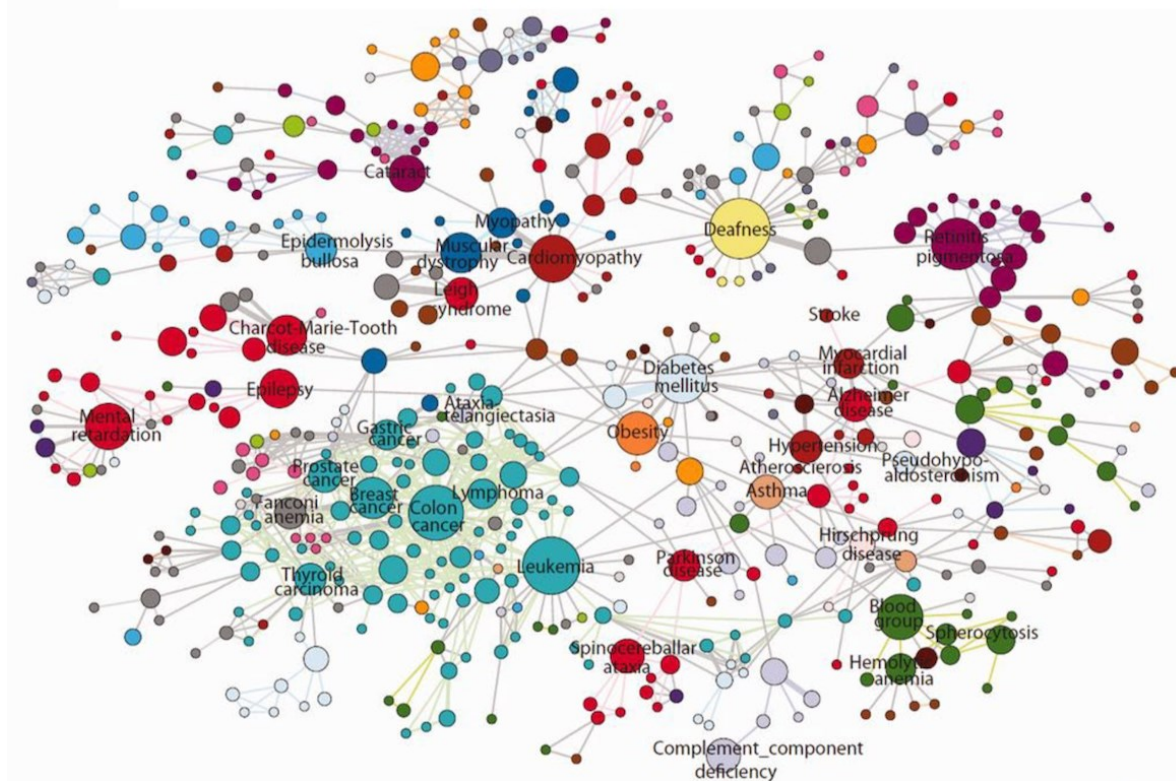
Tipični primeri (8)

- Citatna mreža u nauci
 - Čvorovi su naučne oblasti, a veze citati iz časopisa
 - Makro model



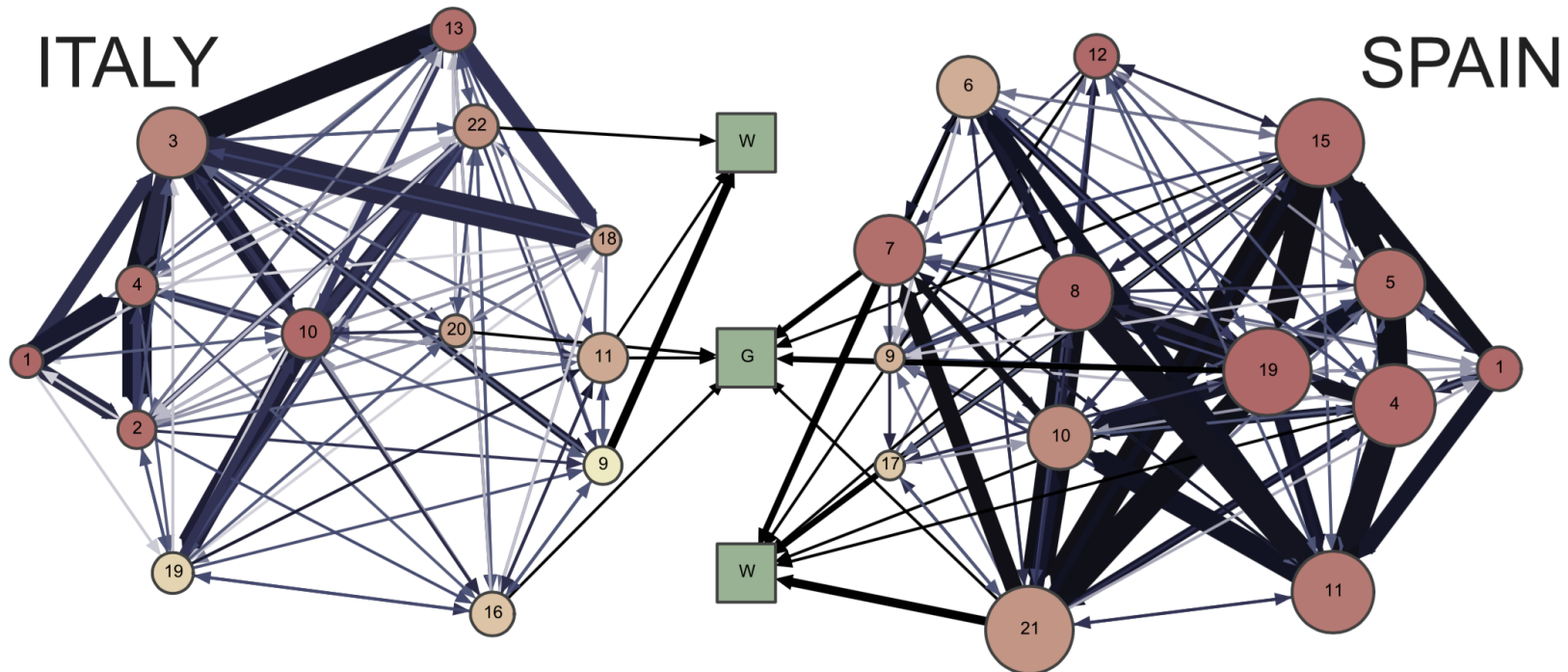
Tipični primeri (9)

- Medicina – mreža bolesti na osnovu deljenih gena
 - Čvorovi su bolesti ljudi
 - Veze predstavljaju deljene, zajedničke gene



Tipični primeri (10)

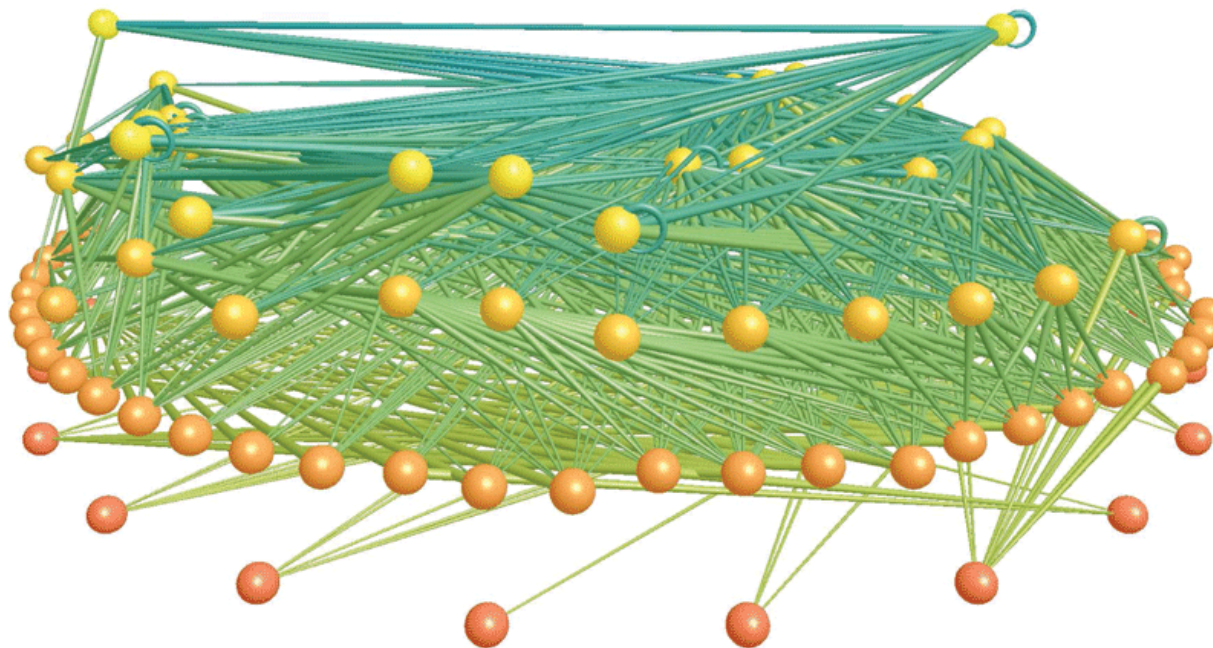
- Mreža fudbalskih igrača
 - Čvorovi su fudbaleri, veze pasovi na utakmicima



Tipični primeri (11)

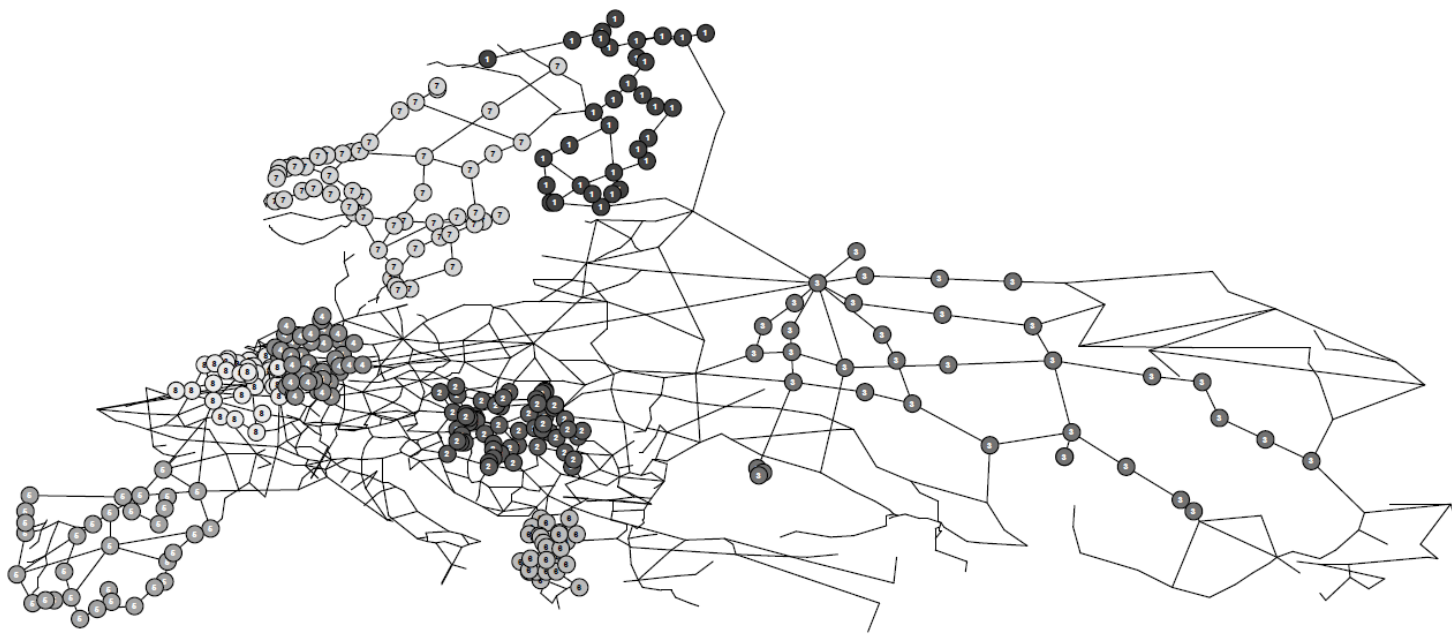
- Ekologija

- Čvorovi su vrste koje žive na nekom području
- Veze predstavljaju lanac ishrane



Tipični primeri (12)

- Evropska putna mreža
 - Čvorovi su gradovi
 - Veze su autoputevi kojima su gradovi povezani



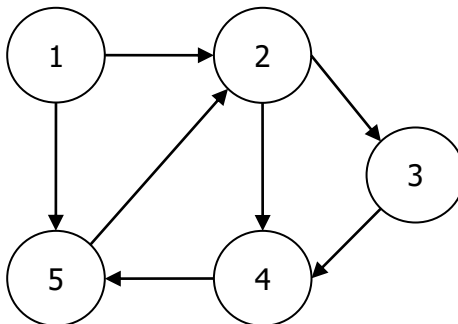
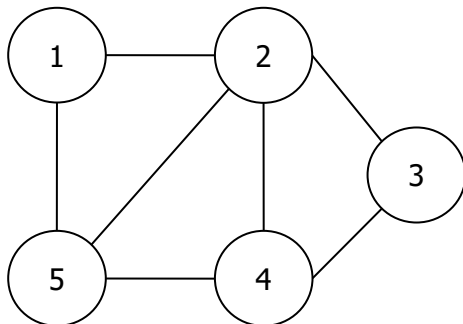
Predstavljanje mreža (1)

- Matrična reprezentacija
 - Matrica susednosti
 - Matrica incidentnosti
- Reprezentacija listom
 - Neka forma liste susednosti
 - Najčešće linearizovana lista susednosti
- Različiti formati zaglavlja
kako bi se očuvale dodatne informacije
 - Podaci o težinama grana, labele čvorova, različiti atributi čvorova i sl.
- Postoje standardizovani formati
 - GML, GEXF i sl.

<https://gephi.org/users/supported-graph-formats/>

Predstavljanje mreža (2)

- Sadržaj matrice susednosti definiše se na sledeći način:
 - $a[i, j] = 1$ ako postoji grana između čvorova i i j
 - $a[i, j] = 0$ ako ne postoji grana između čvorova i i j
 - $a[i, j] = w(i, j)$ za težinske grafove
- Simetrična za neusmerene mreže



	1	2	3	4	5
1	0	1	0	0	1
2	1	0	1	1	1
3	0	1	0	1	0
4	0	1	1	0	1
5	1	1	0	1	0

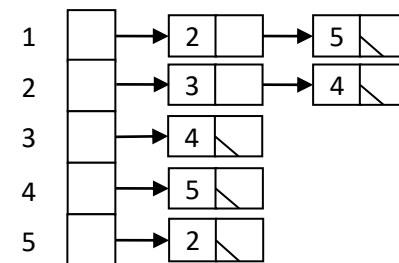
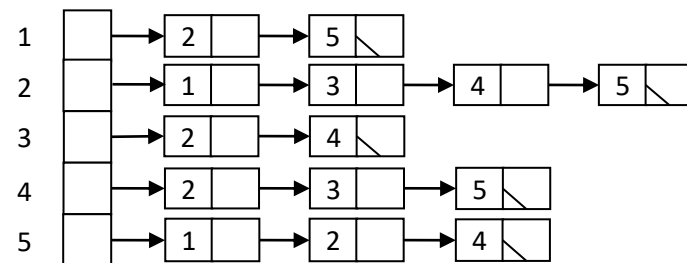
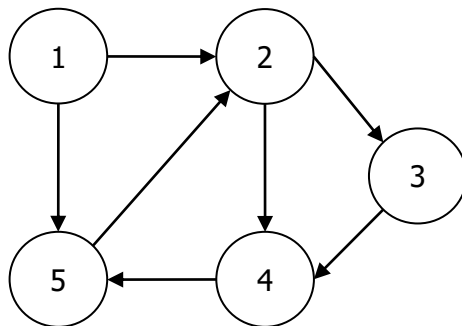
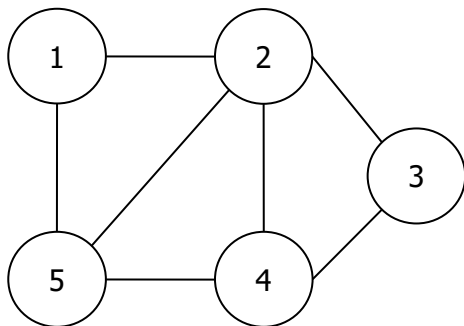
	1	2	3	4	5
1	0	1	0	0	1
2	0	0	1	1	0
3	0	0	0	1	0
4	0	0	0	0	1
5	0	1	0	0	0

Predstavljanje mreža (3)

- Graf se predstavlja listom pomoću:

- Vektora zaglavlja za svaki čvor
- Ulančane liste suseda za svaki čvor
- Element liste odgovara grani

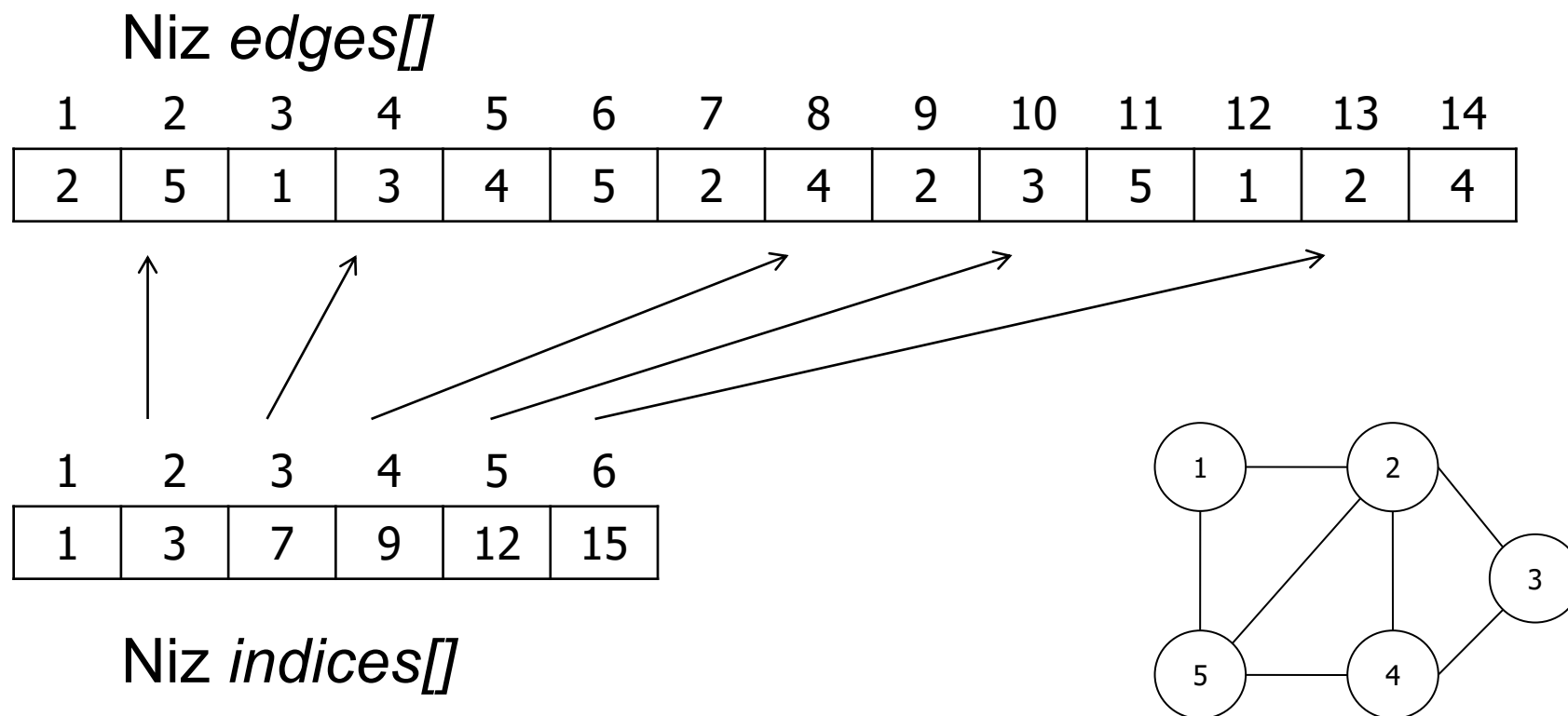
- Ulančana reprezentacija je nepraktična!



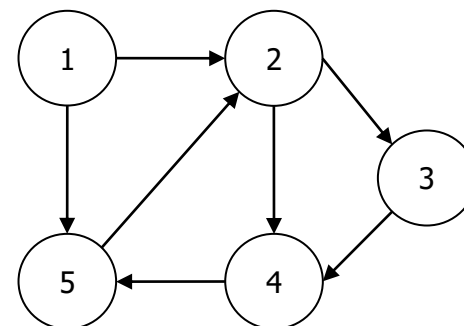
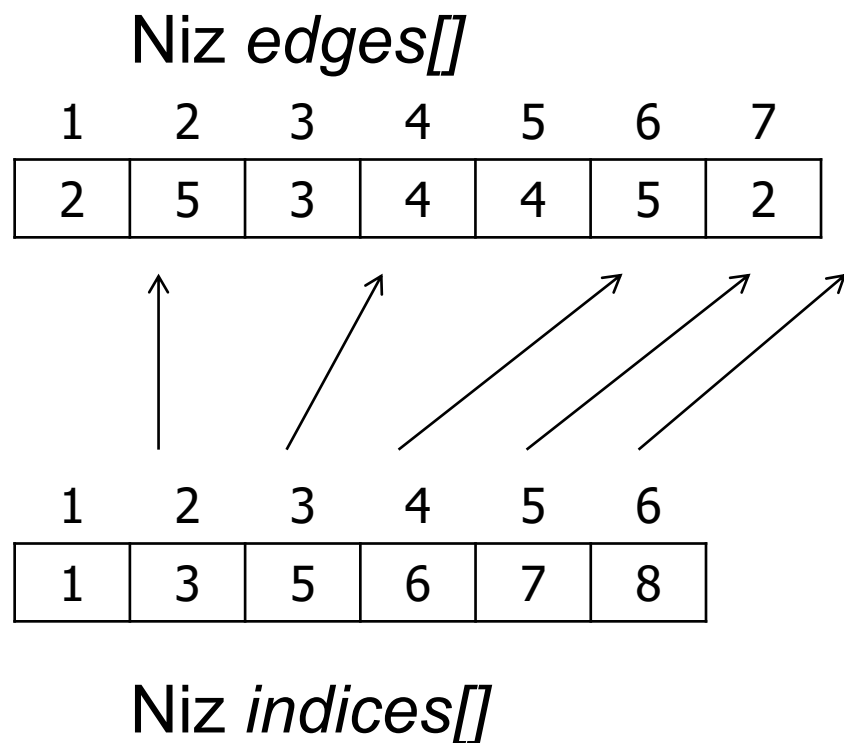
Predstavljanje mreža (4)

- Rešenje kroz linearizovanu listu susednosti
- Graf se predstavlja listom pomoću dva niza:
 - Niz grana (*edges*) predstavlja listu susednosti linearizovanu u niz
 - Ima onoliko elemenata koliko ima grana u grafu
 - Niz čvorova (*indices*) predstavlja indeks prvog elementa j -tog čvora u nizu grana
 - Uobičajeno ima $n+1$ elemenata, gde n predstavlja broj čvorova u grafu
 - Broj suseda i -tog čvora je određen razlikom dva polja u nizu čvorova na pozicijama j i $j+1$

Predstavljanje mreža (5)



Predstavljanje mreža (6)



Prikupljanje podataka (1)

- Način prikupljanja informacija o relacijama
 - *Full network* metodi
 - Sakupljaju se potpune informacije o vezama svih aktera sa drugim akterima
 - *Snowball* metodi
 - Polazi se od određenog skupa aktera u fokusu, a zatim se podaci o vezama prikupljaju na osnovu njihovih veza i novoidentifikovanih aktera
 - *Ego*-centrični metodi
 - Prikupljaju se podaci o pojedinačnom akteru (egu)
 - Sa ili bez veza između njegovih *alter*-a

Prikupljanje podataka (2)

- Tehnike prikupljanja podataka
 - Intervjuisanje aktera
 - Upitnici (papirni ili elektronski)
 - Posredstvom veb API-a *online* društvenih mreža
 - Implementacijom veb tragača (*web crawler*) za pristup strukturiranim stranicama
- Postoje namenski radni okviri (*frameworks*) za implementaciju veb tragača
 - *Apache Nutch* (Java), *Scrapy* (Python)

Literatura

- Hanneman, Robert A. and Mark Riddle, Introduction to social network methods, University of California, Riverside, 2005.
- L. Šubelj, Introduction to Network Analysis, FRI Ljubljana, 2016.
- J. Jovanović, Softverska analiza društvenih mreža, FON, 2017.
- M. Mišić, Unapređenja sistema za detekciju plagijarizma u izvornom programskom kodu – Primena metoda za analizu socijalnih mreža u detekciji plagijarizma, PhD disertacija, ETF 2017.