

ANALIZA DRUŠTVENIH MREŽA

JELENA JOVANOVIĆ

EMAIL: JELJOV@GMAIL.COM

WEB: [HTTP://JELENAJOVANOVIĆ.NET](http://JELENAJOVANOVIĆ.NET)



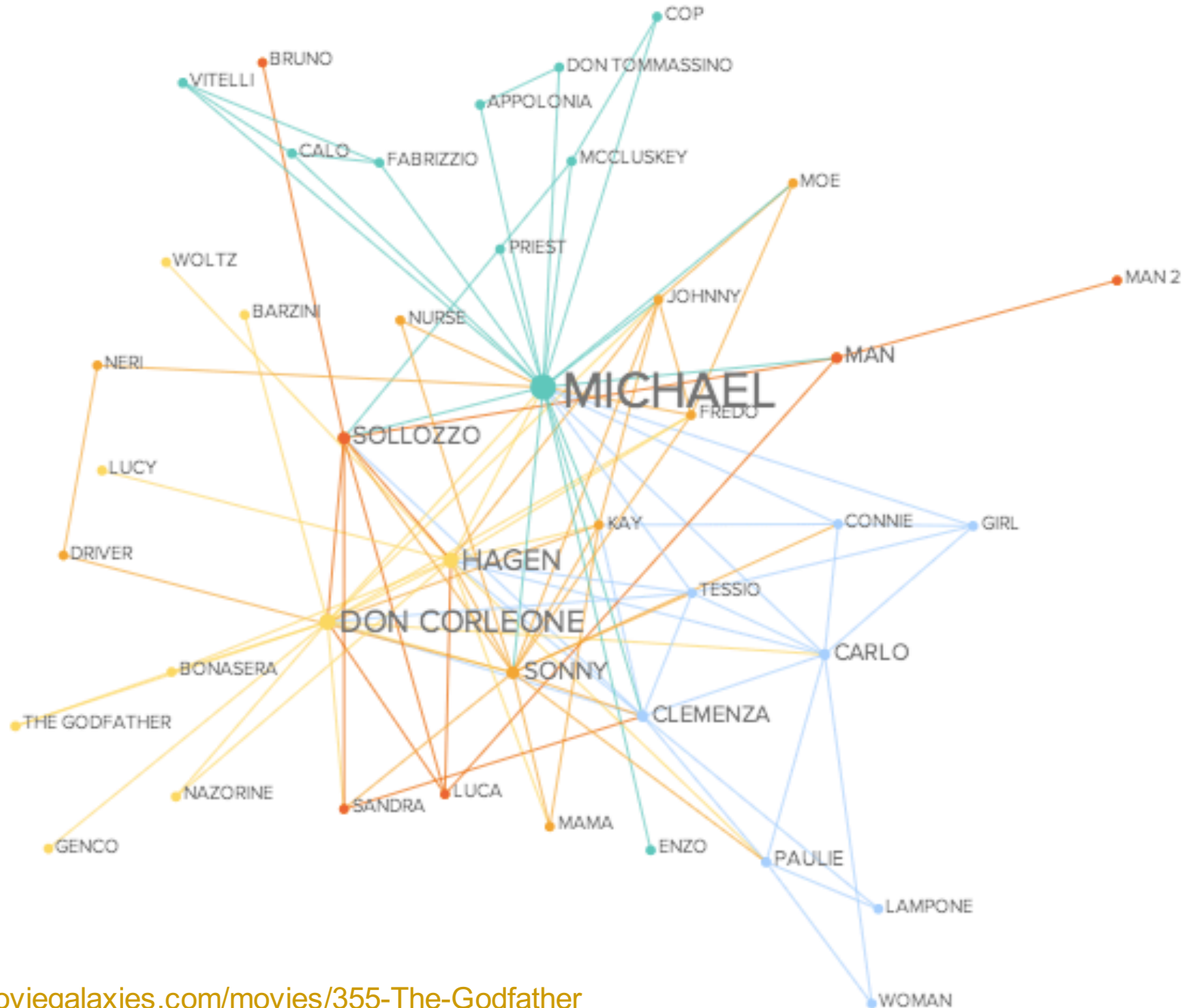
PREGLED PREDAVANJA

- Social Network Analysis (SNA)
 - Šta je SNA?
 - Zašto/kako može biti korisna?
 - Gde se primenjuje?
- Društveni graf kao model društene mreže
- Bipartitne mreže
- Ego mreže
- Osnovne mrežne metrike
- Softverska podrška za analizu društvenih mreža
- Literatura i dodatni korisni sadržaji na temu SNA

ANALIZA DRUŠTVENIH MREŽA

- Social Network Analysis (SNA)
- Podaci o članovima društvene mreže i njihovim međusobnim relacijama se predstavljaju u formi (*društvenog*) *grafa*
 - svaki član mreže predstavlja jedan čvor grafa
 - relacije među članovima mreže su ivice grafa
- Analiza se zasniva na primeni odgovarajućeg skupa SNA metrika i/ili algoritama nad formiranim društvenim grafom

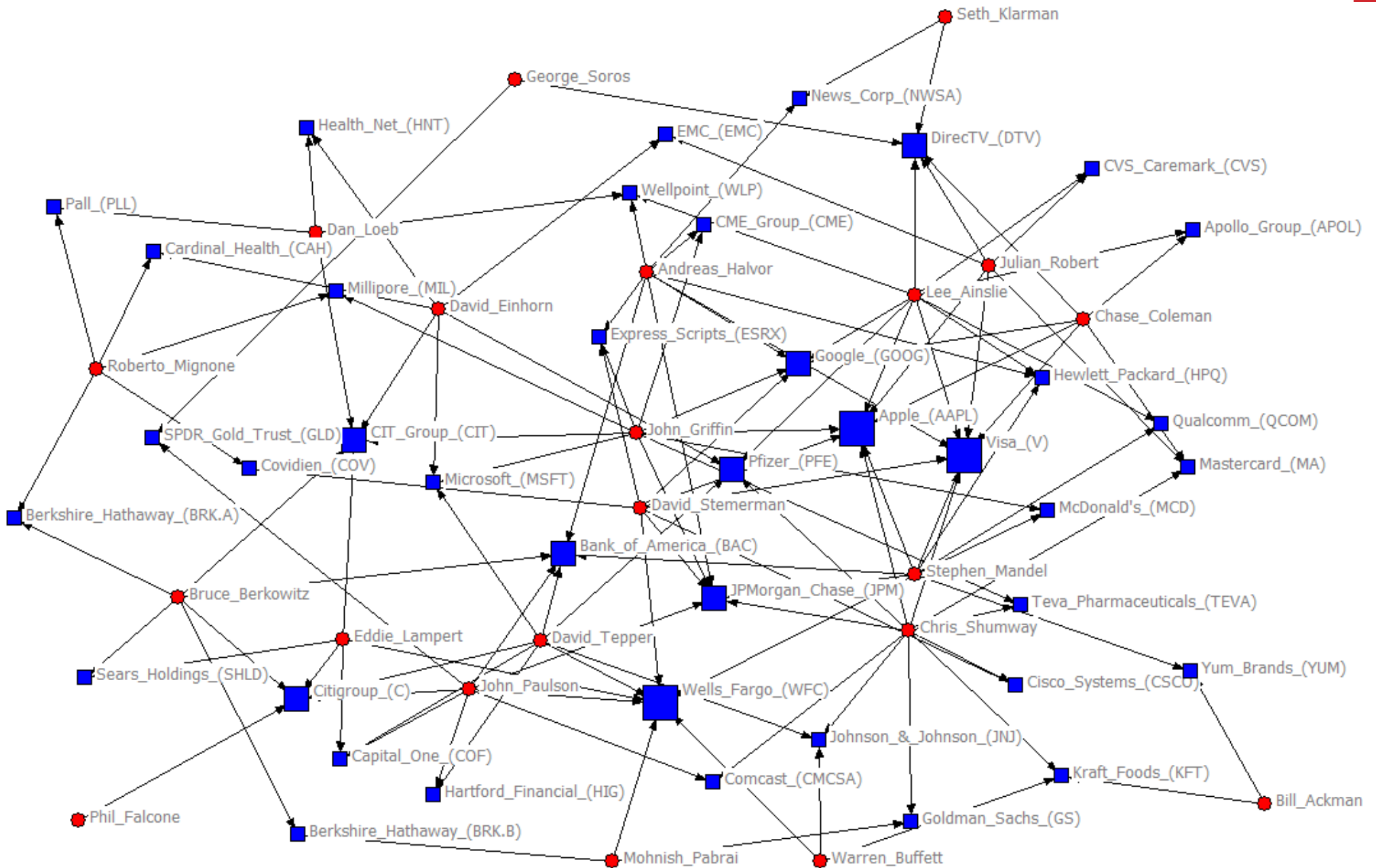
Društvena mreža koju formiraju likovi u filmu The Godfather (1972)



Izvor: <http://moviegalaxies.com/movies/355-The-Godfather>

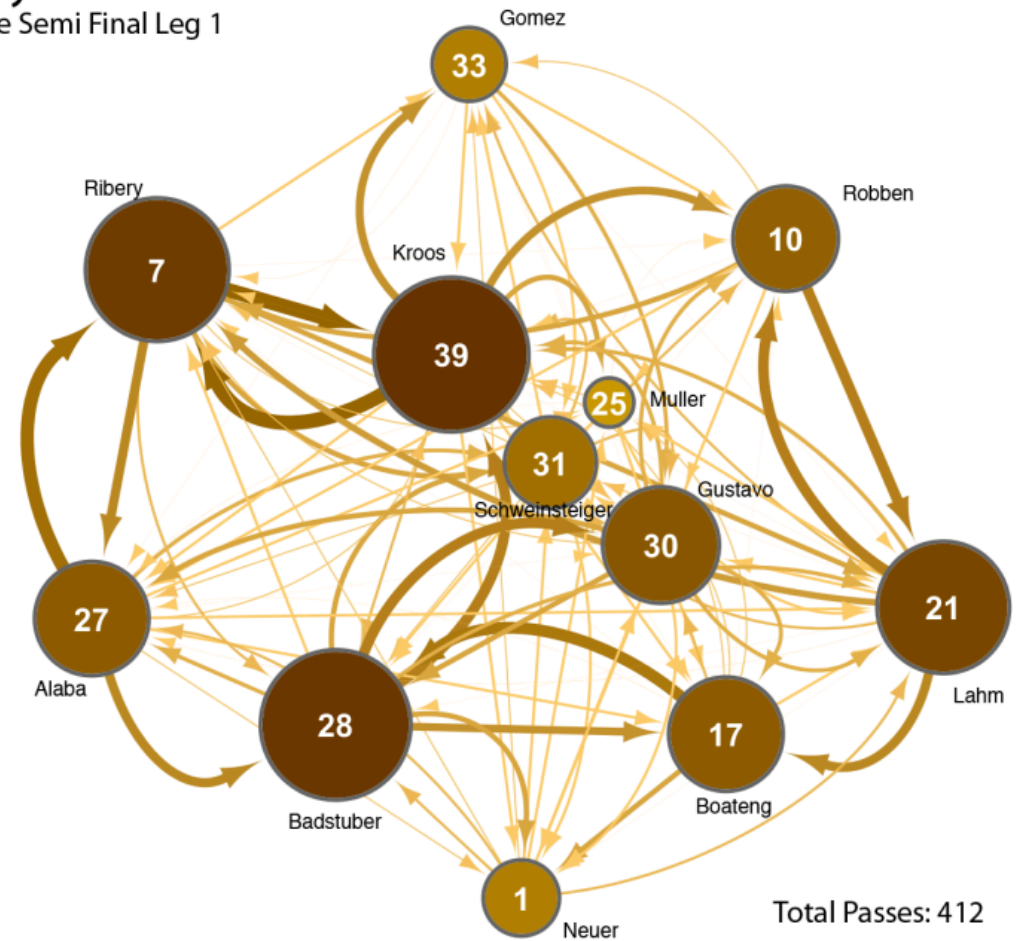
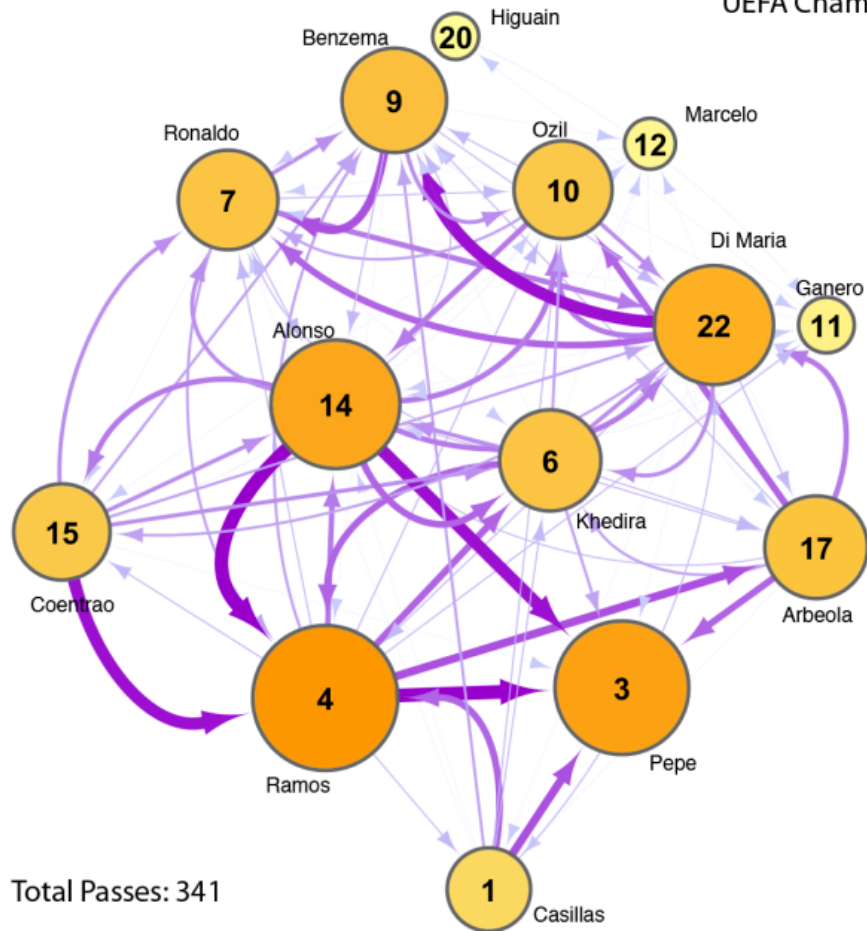
Vise o pristupu koji primenjuje MovieGalaxies: <http://goo.gl/NTdg1D>

Mreža hedž fondova i kompanija u koje su ovi fondovi investirali



Mreža dodavanja (passes) među fudbalerima tokom FIFA 2010 prvenstva

Real Madrid 1 : 2 Bayern Munchen
UEFA Champions League Semi Final Leg 1



ANALIZA DRUŠTVENIH MREŽA

SNA omogućuje da se za datu društvenu mrežu utvrdi:

- Ko su najuticajniji/centralni članovi mreže?
- Koje se grupe mogu uočiti u mreži? U kojoj meri je mreža podeljena na manje, slabo povezane grupe?
- Kako se mreža razvija? Hoće li se mreža održati?
- Kako se ideje/informacije/virusi/... šire kroz mrežu?
- ...

PRIMERI PRIMENE U PRAKSI

Kompanije: unapređenje komunikacionih tokova u okviru organizacije, kao i u široj mreži koja obuhvata i poslovne partnere i klijente

Online društvene mreže: identifikacija i preporuka potencijalnih prijatelja, saradnika, eksperata, ...

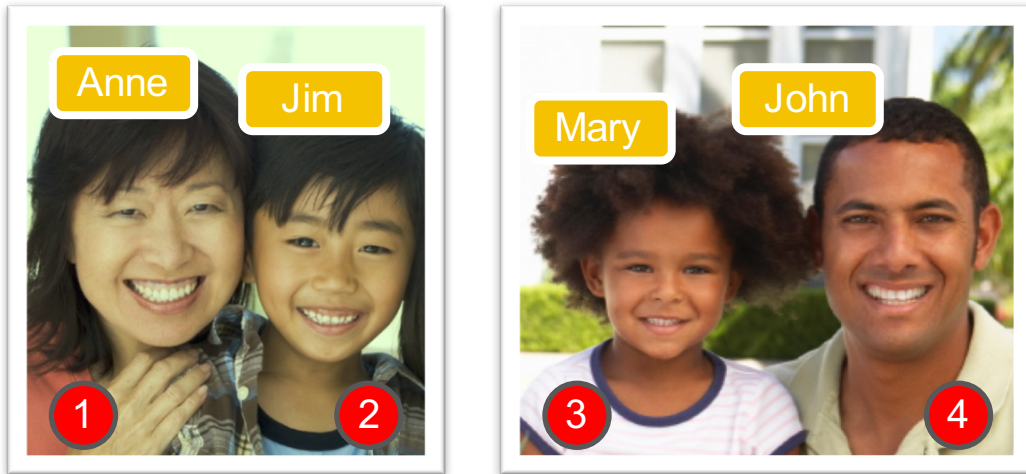
Marketing: identifikacija centralnih članova društvenih mreža u cilju boljeg promovisanja novih proizvoda/usluga/akcija/...

Telekomunikacije: optimizacija strukture i kapaciteta telekomunikacionih mreža

Policija i istražne agencije: identifikacija centralnih ličnosti u različitim kriminalnim grupama i mrežama

DRUŠTVENI GRAF KAO MODEL DRUŠTVENE MREŽE

DRUŠTVENA MREŽA / GRAF



Anne: Jim, tell the Murrays they're invited

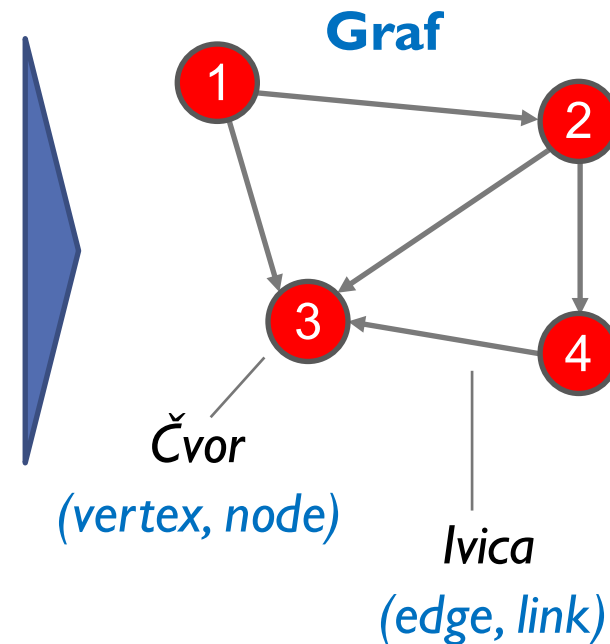
Jim: Mary, you and your dad should come for dinner!

Jim: Mr. Murray, you should both come for dinner

Anne: Mary, did Jim tell you about the dinner?

John: Mary, are you hungry?

...



DRUŠTVENA MREŽA

Ivice (društvene) mreže tipično odražavaju:

- Neki oblik društvene relacije (npr., prijateljstvo, rodbinske veze, poslovne konekcije, ...), ili
- Neki oblik društvene interakcije (npr., razmena poruka na chat-u, komunikacija putem email-a, ...), ili
- Posedovanje neke zajedničke osobine članova mreže

DRUŠTVENA MREŽA

Zavisno od toga da li je relacija koju predstavlja eksplicitno ili implicitno pristutna među akterima (čvorovima), mreža može biti

- *Direktna* - predstavlja eksplicitno prisutne relacije između aktera
 - Npr., dve osobe su prijatelji, ili jedna osoba sledi drugu na nekoj društvenoj mreži

- *Indirektna* - opisuje implicitno prisutne relacije među akterima
 - Npr., dve osobe su editovale istu Wiki stranicu, ili su prisustvovalе istom događaju, ...

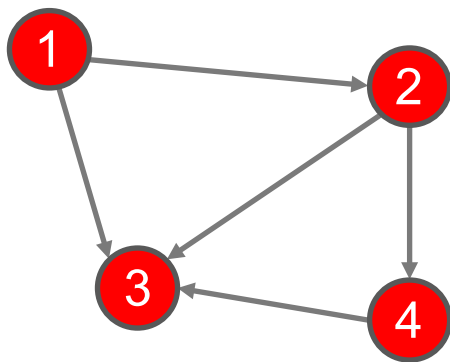
DRUŠTVENA MREŽA

Zavisno od toga da li je relacija između dva čvora mreže prisutna u oba smera ili samo u jednom, mreža može biti

- *Usmerena (directed)* – smer u kome je relacija prisutna je eksplicitno zadat usmerenom strelicom u grafu
 - Npr. mreža koja povezuje pošiljaoce i primaoce poruka
- *Neusmerena (undirected)* – relacija je prisutna u oba smera; smer relacije nije naznačen usmerenim strelicama jer se podrazumeva da relacija važi u oba smera
 - Npr. mreža prijatelja

PREDSTAVLJANJE USMERENOG GRAFA

Usmereni graf



Lista ivica grafa

Čvor	Čvor
1	2
1	3
2	3
2	4
4	3

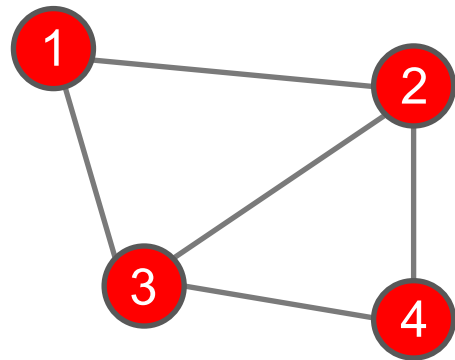
Matrica susedstva

Čvor	1	2	3	4
1	-	1	1	0
2	0	-	1	1
3	0	0	-	0
4	0	0	1	-

PREDSTAVLJANJE NEUSMERENOG GRAFA

Lista ivica ostaje nepromenjena
Ali je njena interpretacija sad drugačija

Neusmereni graf



Čvor	Čvor
1	2
1	3
2	3
2	4
3	4

Matrica susedstva postaje simetrična

Čvor	1	2	3	4
1	-	1	1	0
2	1	-	1	1
3	1	1	-	1
4	0	1	1	-

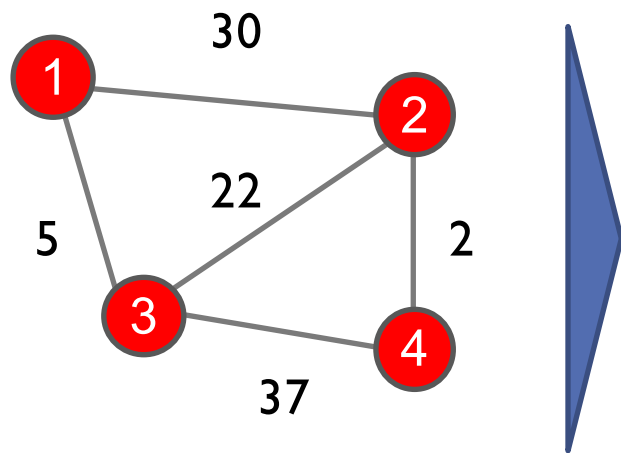
DODAVANJE TEŽINA IVICAMA GRAFA

Težine ivica mogu predstavljati:

- Učestanost interakcije između članova neke grupe
- Broj poruka razmenjenih u nekom vremenskom periodu
- Subjektivnu ocenu stepena povezanosti dve osobe
- Fizičku blizinu ili udaljenost dve osobe
- Neku kombinaciju prethodnih elemenata
- ...

DODAVANJE TEŽINA IVICAMA GRAFA

Lista ivica: dodajemo kolonu za težinu ivica



Čvor	Čvor	Težina
1	2	30
1	3	5
2	3	22
2	4	2
3	4	37

Matrica susedstva: polja sadrže težinu ivica

Vertex	1	2	3	4
1	-	30	5	0
2	30	-	22	2
3	5	22	-	37
4	0	2	37	-

PREPORUKA

Video “Building a Network” (by J. Golnbeck; [YouTube link](#))

lepo objašnjava šta treba uzeti u obzir, koje korake treba slediti i koje odluke je potrebno doneti pri kreiranju društvene mreže

BIPARTITNE MREŽE / GRAFOVI

BIPARTITNI GRAFOVI

Bipartitne (bi-modalne) grafove (bipartite, two-mode) karakteriše sledeće:

- postoje 2 vrste čvorova tj. čvorovi predstavljaju 2 različite vrste entiteta; npr. osobe i dokumenti; osobe i projekti, ...
- konekcije su moguće samo između entiteta (čvorova) različitog tipa

Generalno, *multi-modalni grafovi*, su grafovi u kojima postoji K različitih vrsta čvorova ($K > 1$), a konekcije su moguće samo između čvorova različitog tipa

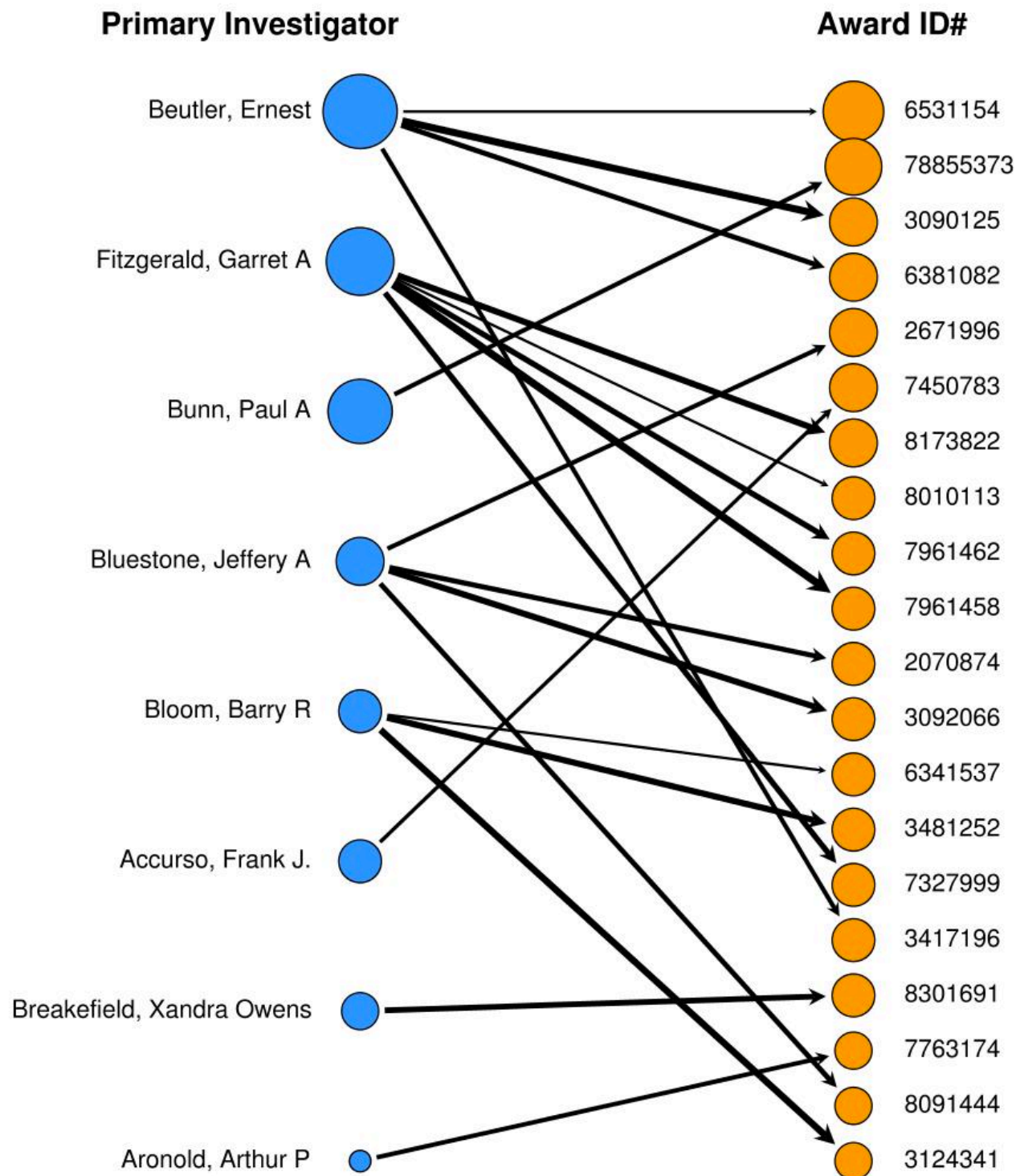
PRIMER:

BIPARTITNI

GRAF ISTRAŽIVAČA

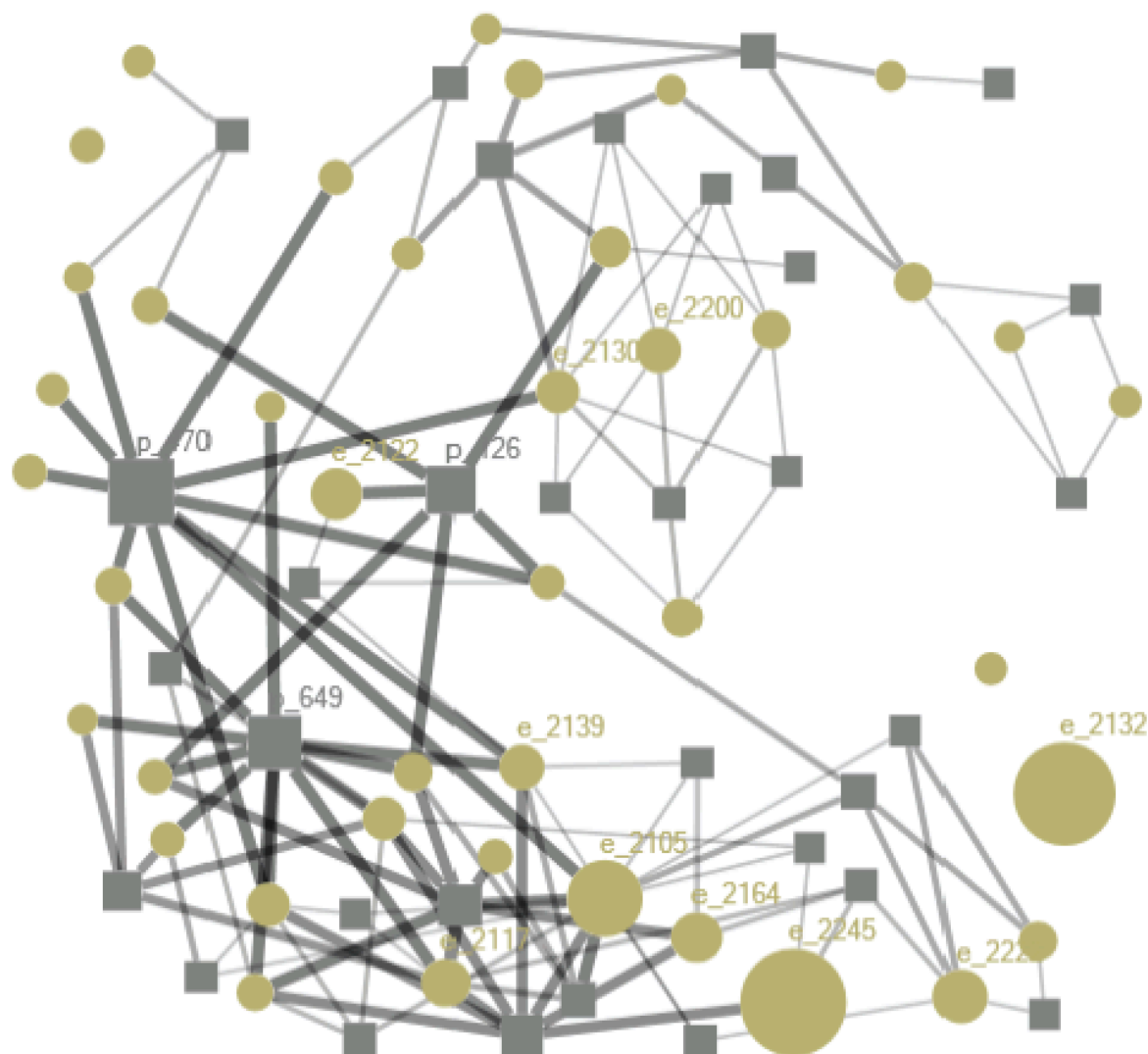
I PROJEKATA KOJE

VODE



PRIMER:

BIPARTITNI GRAF WIKI STRANICA I NJIHOVIH EDITORA



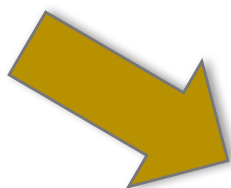
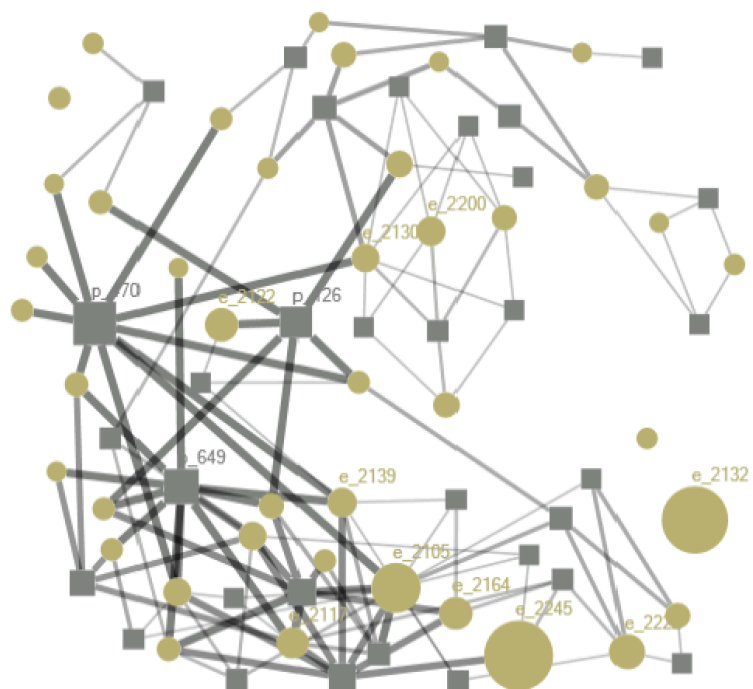
Napomena:
čvorovi koji predstavljaju Wiki
stranice editovane od strane samo
jedne osobe nisu prikazane
(radi preglednosti prikaza)

BIPARTITNI GRAFOVI

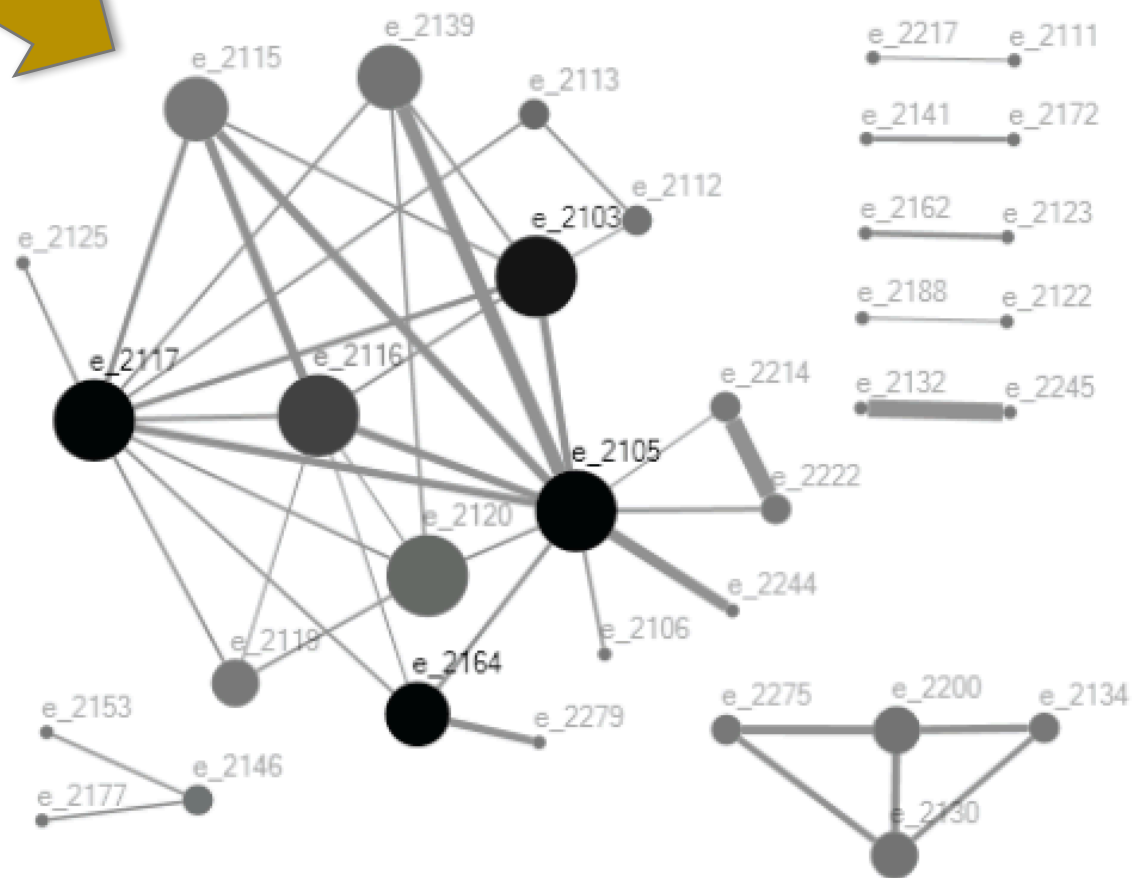
Bipartitni grafovi se, za potrebe analize, često transformišu u unipartitne (uni-modalne), projekcijom grafa na čvorove samo jednog tipa

Konekcije se u unipartitnom grafu uspostavljaju primenom nekog oblika agregacije konekcija u polaznom bipartitnom grafu

Primer projekcije bipartitnog u unipartitni graf



Unipartitni graf editora wiki stranica



Oblik agregacije primenjen u ovom primeru:
osobe koje su koeditovale bar 5 istih Wiki stranica su povezane u unipartitnom grafu

EGO MREŽA

EGO MREŽA (EGO NETWORK)

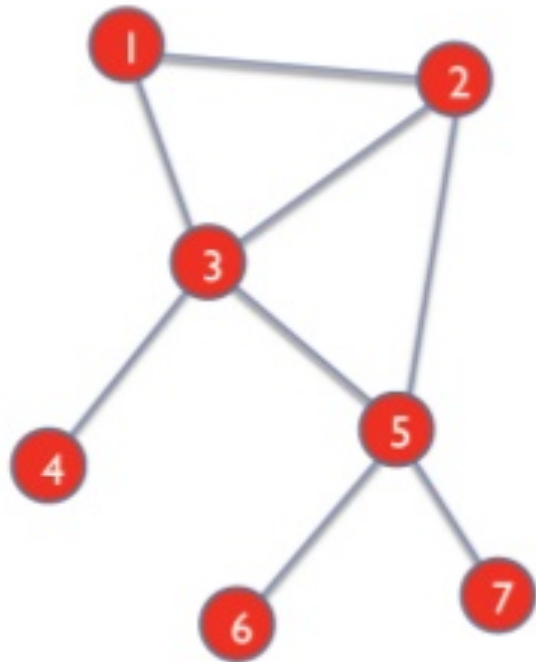
Deo kompletne društvene mreže koji obuhvata izabranog pojedinca (ego) i njegovo neposredno okruženje (alter)

Preciznije, ego mrežu čine:

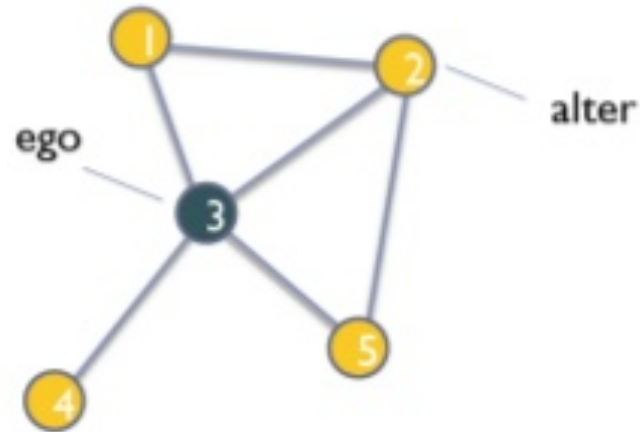
- izabrani akter (ego),
- ostali akteri (čvorovi) mreže koji su direktno povezani sa egom
- konekcije između ega i njegovih suseda, kao i konekcije koje međusobno povezuju susede ega

EGO MREŽA

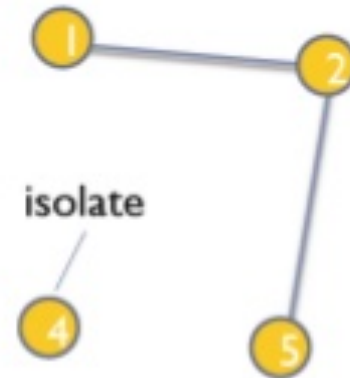
'whole' network*



3's ego network



3's ego network without ego**



* no studied network is 'whole' in practice; it's usually a partial picture of one's real life networks (boundary specification problem)

** ego not needed for analysis as all alters are by definition connected to ego

EGO MREŽE

Ego mreža može obuhvatiti ne samo susede koji su direktno povezani sa egom, već i one koji su 2 ili više koraka udaljeni; tada govorimo o N-nivojskoj ego mreži (*N-levels ego network*)

- Npr., 2-nivojska ego mreža obuhvata susede koji su 2 koraka udaljeni od ega, kao i međusobne relacije ovih suseda

Napomena: ego mreža opisana na prethodnim slajdovima se nekad naziva i 1.5-level ego mreža

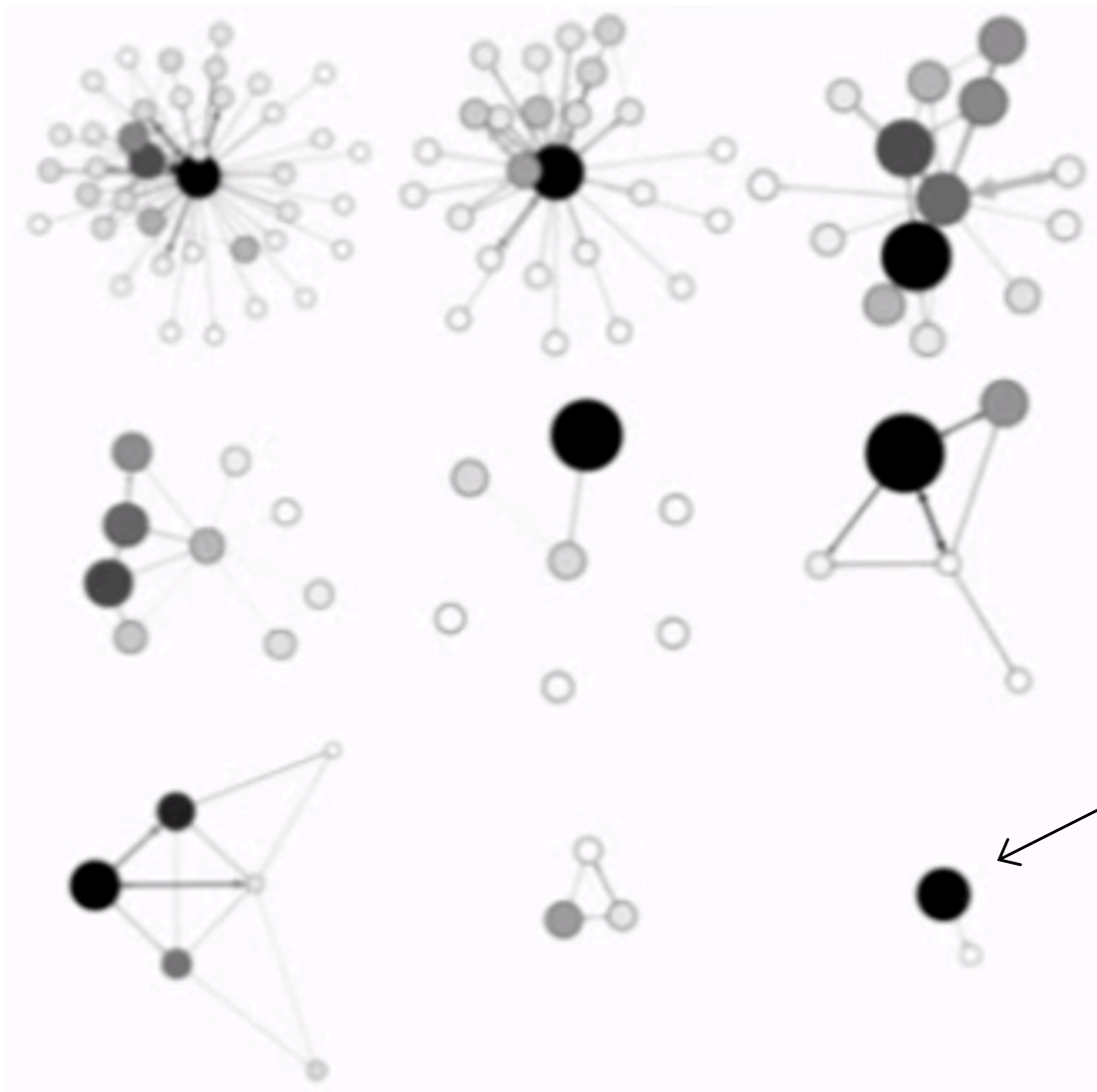
- Razlog: mreža obuhvata ne samo relacije ega i suseda, već i relacije između suseda
- Problem: različiti autori koriste različitu terminologiju

EGO MREŽE

Analiza ego mreža najčešće ima za cilj identifikovanje sličnosti i razlika u poziciji i povezanosti pojedinačnih čvorova u mreži

U slučaju analize online društvenih mreža i medija, analiza ego mreža se često kombinuje sa analizom sadržaja koje učesnici u mreži razmenjuju ili kreiraju

PRIMER: EGO MREŽE IDENTIFIKOVANE U DRUŠTVENOJ MREŽI UČESNIKA DISKUSIONOG FORUMA



Ego u svakoj mreži je središnji čvor

Najdominantniji tip ego mreže u datoj društvenoj mreži (~60% članova)

Sledeći korak je analiza poruka koje su razmenjivali ovi članovi foruma kako bi se bolje razumele specifičnosti interakcije u forumu

OSNOVNE MREŽNE METRIKE

VELIČINA MREŽE

Odnosi se na broj čvorova u mreži

U mreži koja sadrži n čvorova, mogući broj konekcija je:

- $n(n-1)$ u slučaju usmerene mreže,
- $n(n-1)/2$ ako je mreža neusmerena

VELIČINA MREŽE

Mogući broj različitih načina povezivanja n čvorova je $2^{n(n-1)/2}$

- I za male mreže, broj mogućih načina formiranja mreže je ogroman
- Primer: 30 čvorova mogu biti povezani u neusmereni graf na 2^{435} različitih načina; procena je da se broj atoma u univerzumu kreće u opsegu 2^{158} i 2^{246} !!!

GUSTINA MREŽE

Gustina mreže se definiše kao količnik broja konekcija prisutnih u datoj mreži i ukupnog broja mogućih konekcija u toj mreži

U slučaju težinskog grafa, gustina se definiše kao količnik sume težina svih konekcija i broja mogućih konekcija u grafu

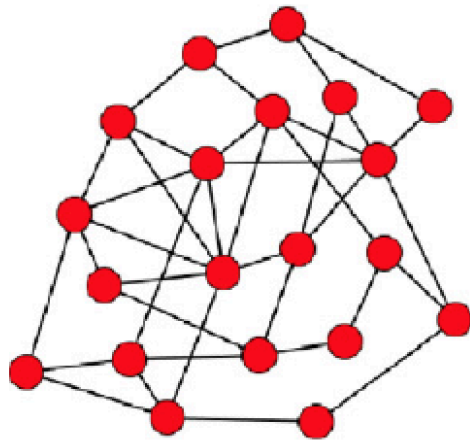
Gustina mreže

- ukazuje na stepen povezanosti čvorova te mreže
- daje uvid u pojave poput brzine širenja informacija kroz mrežu

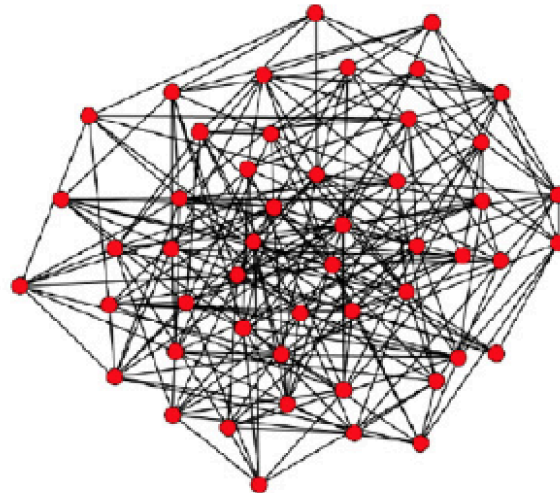
GUSTINA MREŽE VS. LINKS PER NODE

Ukoliko se koristi za poređenje grafova različite veličine, gustina može dovesti do pogrešnih zaključaka

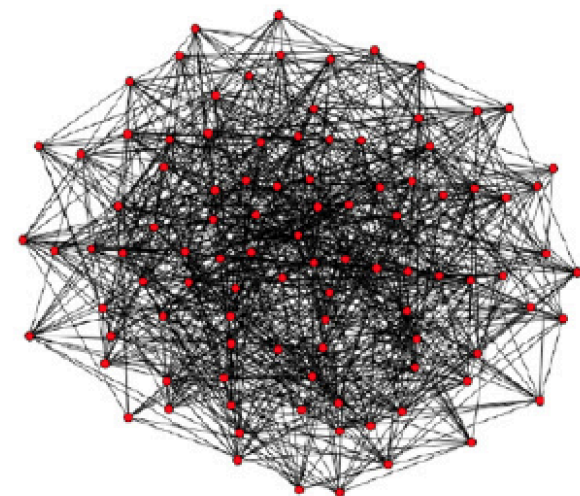
Umesto gustine, koristi se *links per node* metrika koja predstavlja racio broja konekcija i broja čvorova u grafu



20 nodes, 38 links
Density = 0.20
Links per node = 1.9



50 nodes, 245 links
Density = 0.20
Links per node = 4.9

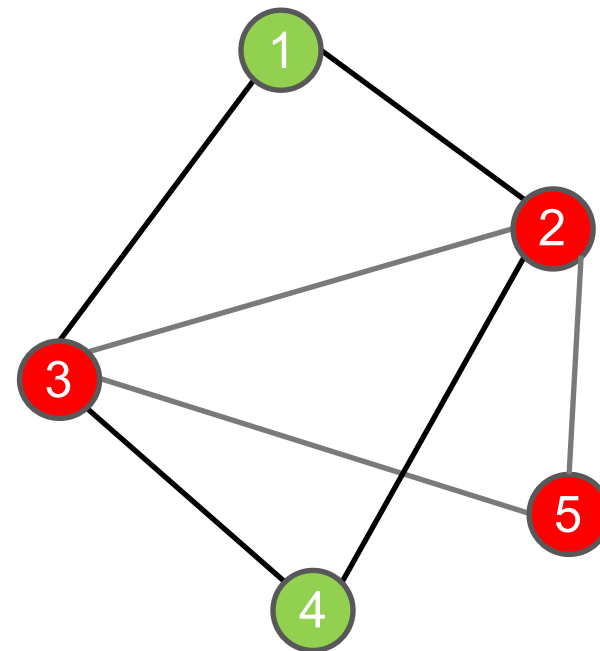


100 nodes, 990 links
Density = 0.20
Links per node = 9.9

PUTANJE KROZ MREŽU

Putanja između dva čvora A i B je bilo koja sekvenca čvorova koja povezuje čvorove A i B, u kojoj se čvorovi *ne ponavljaju*

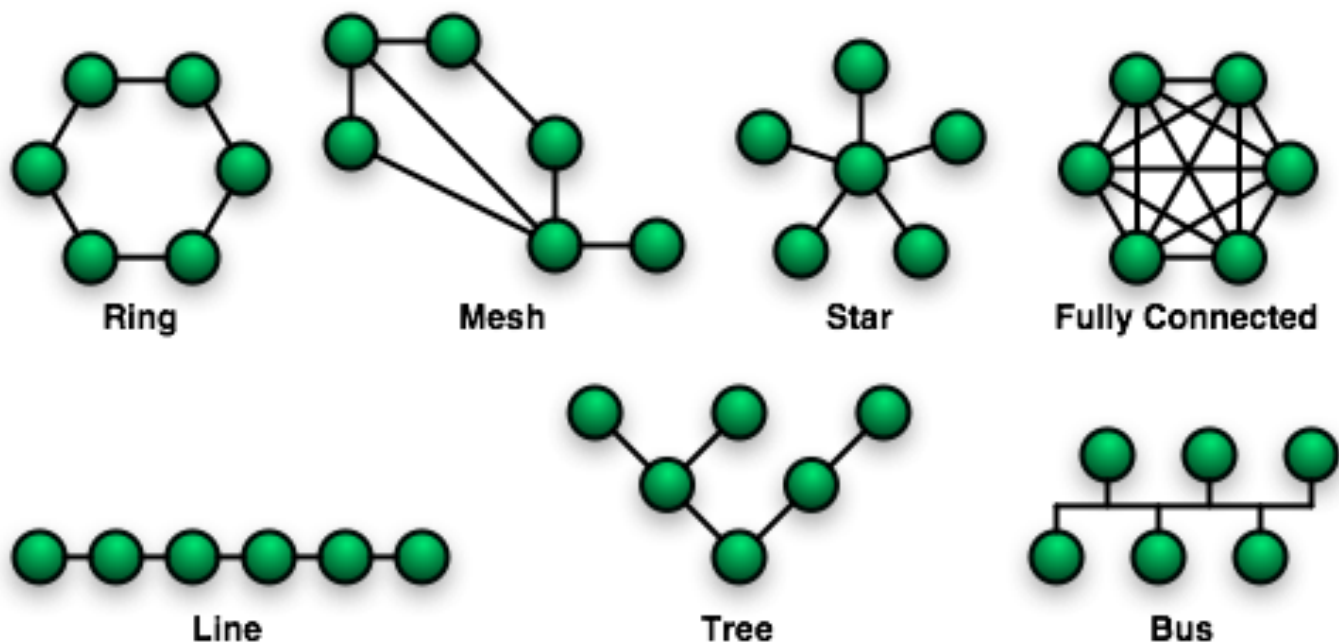
Npr., neke od putanja između čvorova 1 i 4 su: {1,3,4}, {1,2,3,4}, {1,3,2,4}, {1,2,5,3,4}, {1,3,5,2,4}



POVEZANOST MREŽE

Mreža je povezana ukoliko postoji bar jedna putanja između bilo koja dva čvora mreže

Povezana mreža može imati više različitih topologija:



POVEZANOST MREŽE

Povezanost mreže je viša/jača, odnosno mreža je stabilnija, ukoliko postoji više putanja koje povezuju parove čvorova u mreži

Redundantnost putanja omogućuje mreži da ostane povezana i kad neki akteri (čvorovi) napuste mrežu

Point connectivity predstavlja min broj čvorova koje je potrebno isključiti iz mreže da bi se prekinula povezanost dva čvora

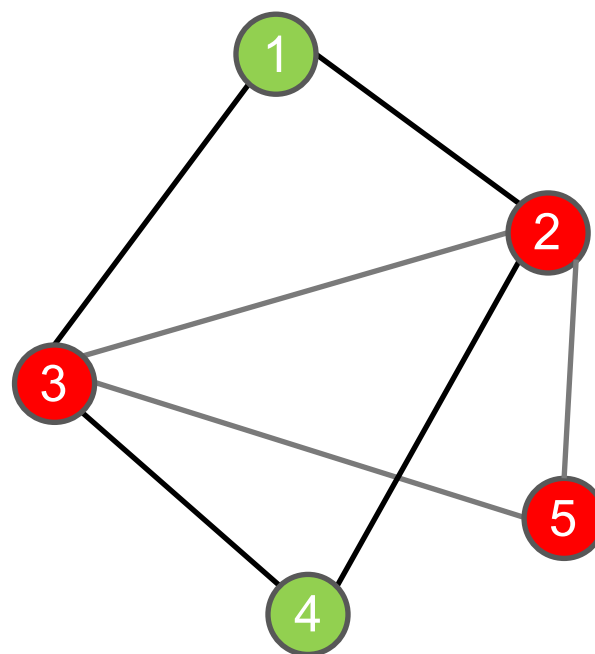
- korisna je za identifikovanje “osetljivih/ranjivih tačaka” mreže

NAJKRAĆA PUTANJA (GEODESIC)

Najkraća putanja (geodesic)

između dva čvora A i B je *putanja* koja povezuje čvorove A i B preko najmanjeg broja ivica

U grafu desno, najkraće putanje između čvorova 1 i 4 su:
 $\{1,2,4\}$ i $\{1,3,4\}$



NAJKRAĆA PUTANJA (GEODESIC)

Geodesic predstavlja optimalnu ili najefikasniju vezu između dva aktera u mreži

Kad se govori o putanji, obično se podrazumeva da je reč o najkraćoj putanji

U slučaju težinskih grafova, interpretacija najkraće putanje je drugačija i zavisi od toga šta težine konekcija predstavljaju

- Npr., ako težine predstavljaju fizičku udaljenost aktera ili neki oblik troškova vezanih za njihovu konekciju, onda je najkraća putanja ona sa najmanjim ukupnom težinom

ECCENTRICITY

Za datog aktera, *eccentricity* predstavlja najdužu najkraću putanju do ostalih čvorova u mreži, odnosno

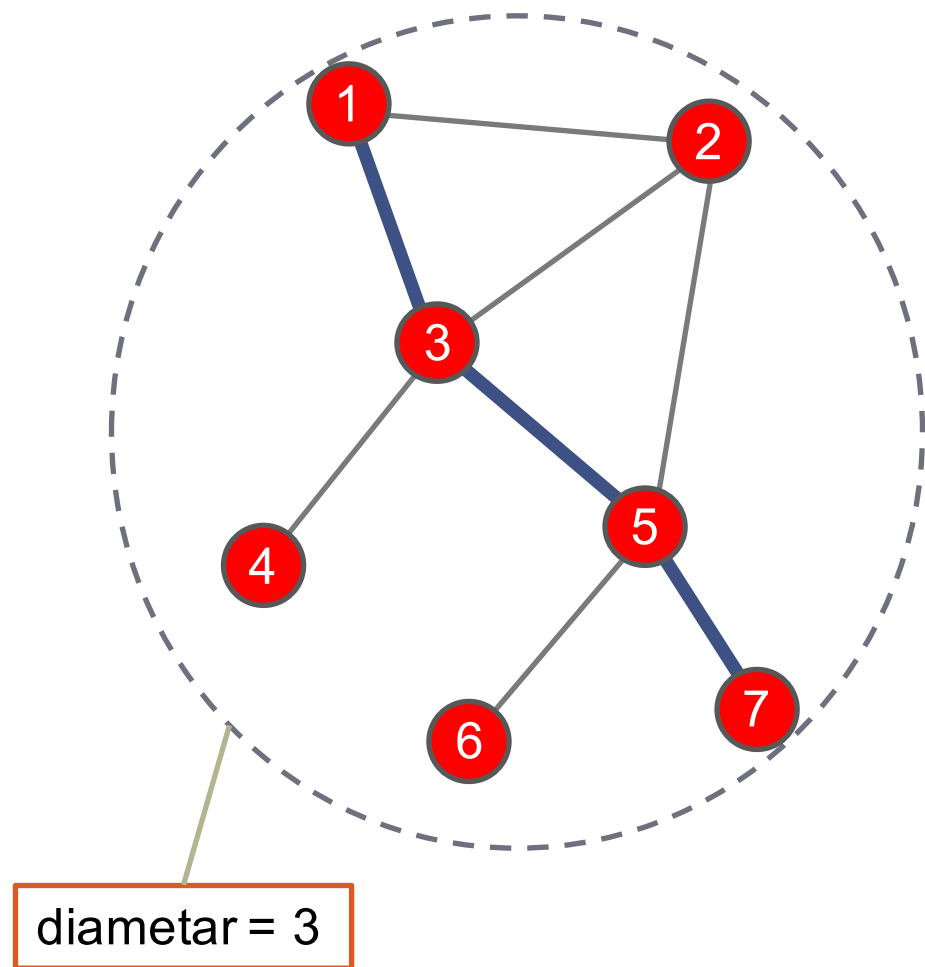
- za datog aktera (čvor), odredi se najkraća putanja prema svim ostalim čvorovima u mreži
- među izračunatim putanjima utvrđuje se najduža putanja – ta putanja predstavlja *eccentricity* za datog aktera

Ova metrika predstavlja koliko je dati akter (čvor) udaljen od njemu najudaljenijeg aktera u mreži

DIAMETAR

Najduža najkraća putanja između bilo koja dva čvora u mreži predstavlja **dijametar** mreže

Dijametar pokazuje koliko je “koraka” u mreži *najviše* potrebno napraviti da bi se polazeći od bilo kog čvora stiglo do bilo kog drugog čvora mreže



DIAMETAR I PROSEČNA DUŽINA PUTANJE

Može se desiti da mreža sadrži par čvorova koji su značajno više međusobno udaljeni nego što je to slučaj sa ostalim parovima čvorova (outliers)

U tom slučaju, umesto diametra, bolji indikator udaljenosti čvorova u mreži je prosečna dužina putanja u mreži (*Average Path Length - APL*)

APL se definiše kao prosečna dužina najkraćih putanja između svaka dva čvora u mreži

DIAMETAR I PROSEČNA DUŽINA PUTANJE

U realnim društvenim mrežama, diametar i prosečna dužina putanje (APL) u mreži su relativno male vrednosti

Na primer:

- U mreži od 721 miliona korisnika Facebook-a, Backstrom et al (2012) su pokazali da je $APL = 4.74$
- U mreži od 50M Web stranica, Adamic & Pitkow (1999) su utvrdili $APL = 3.1$

Ova pojava poznata je kao “mali svet” (*small world*) i *6 degrees of separation*

Više o tome na nekom od narednih časova

PREPORUKA

Pogledati video “Network Theory: 8. Diameter & Scale”

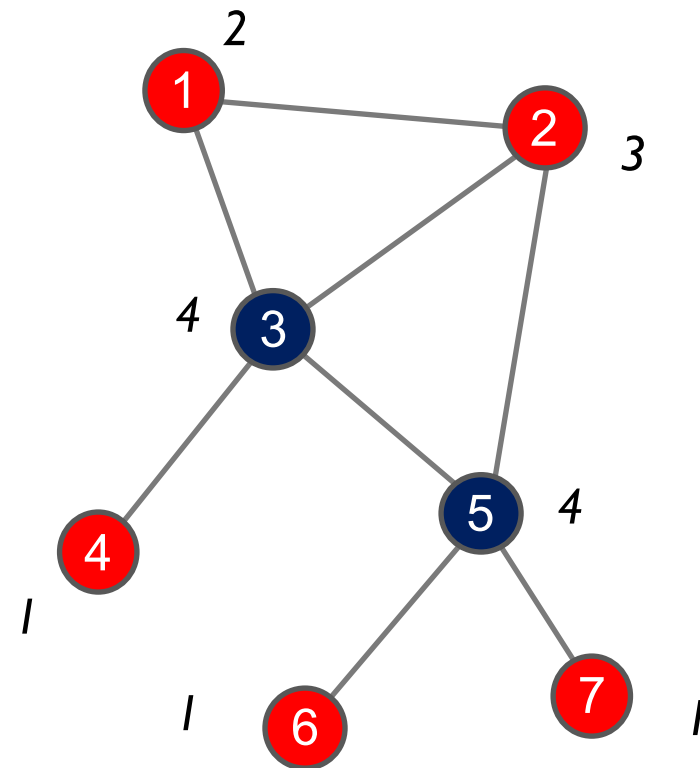
- <https://www.youtube.com/watch?v=c1CwUumWhKQ&index=9>
- lepo objašnjeni i ilustrovani koncepti povezanosti u mreži, najkraće putanje, dijametra i sl

DEGREE

Degree - broj drugih čvorova sa kojima je dati čvor povezan

Često se koristi kao mera stepena umreženosti aktera, a time i njegovog uticaja i/ili popularnosti u mreži

Korisna mera pri proceni/predikciji širenja informacija kroz mrežu i mogućnosti aktera da utiče na svoje neposredno okruženje



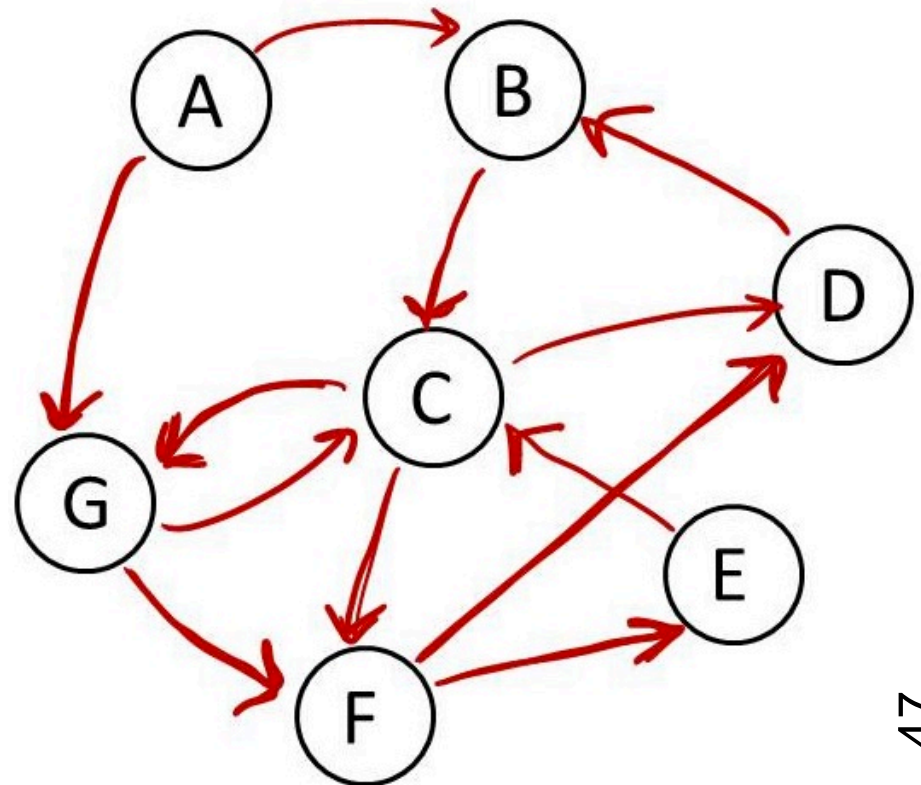
DEGREE

U slučaju usmerenog grafa, razlikujemo:

- **In-degree** – broj ivica koje uviru u dati čvor
- **Out-degree** – broj ivica koje polaze iz datog čvora

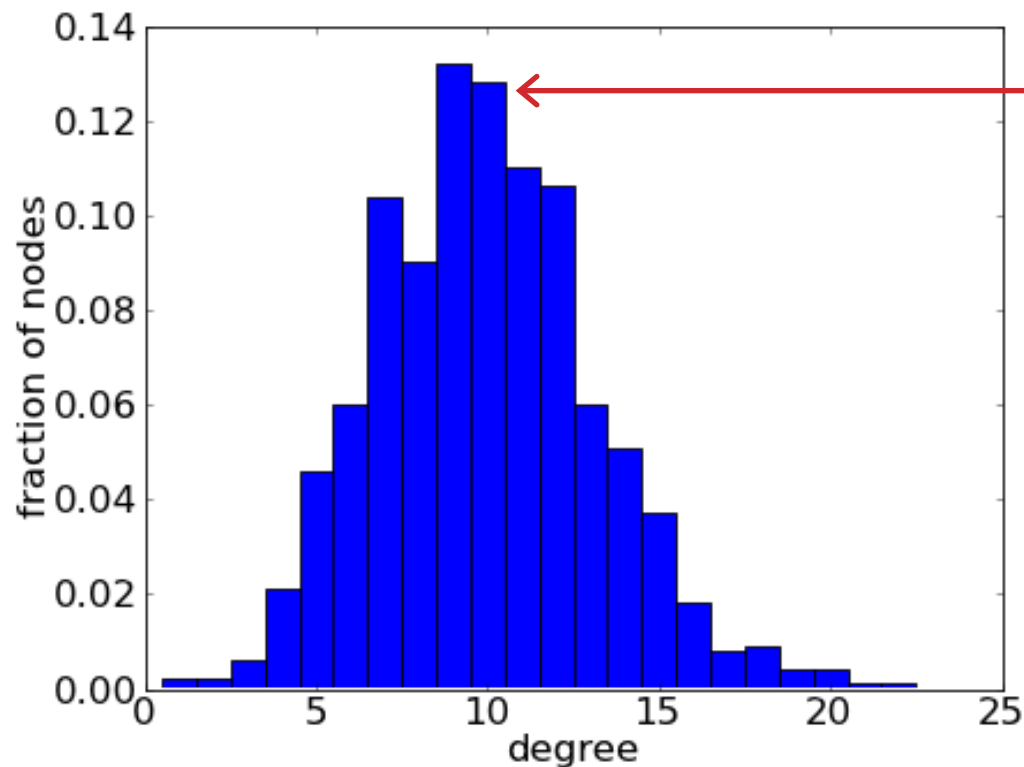
Na primer, za čvor B

- in-degree je 2
- out-degree je 1
- (ukupan) degree je 3



DEGREE DISTRIBUTION

Degree distribution, $P(k)$, predstavlja procenat čvorova mreže kod kojih degree metrika ima vrednost k



Npr. U mreži čiji je degree distribution prikazan na slici, verovatnoća da čvor ima degree 10 (tj. 10 neposrednih suseda) je 0.13

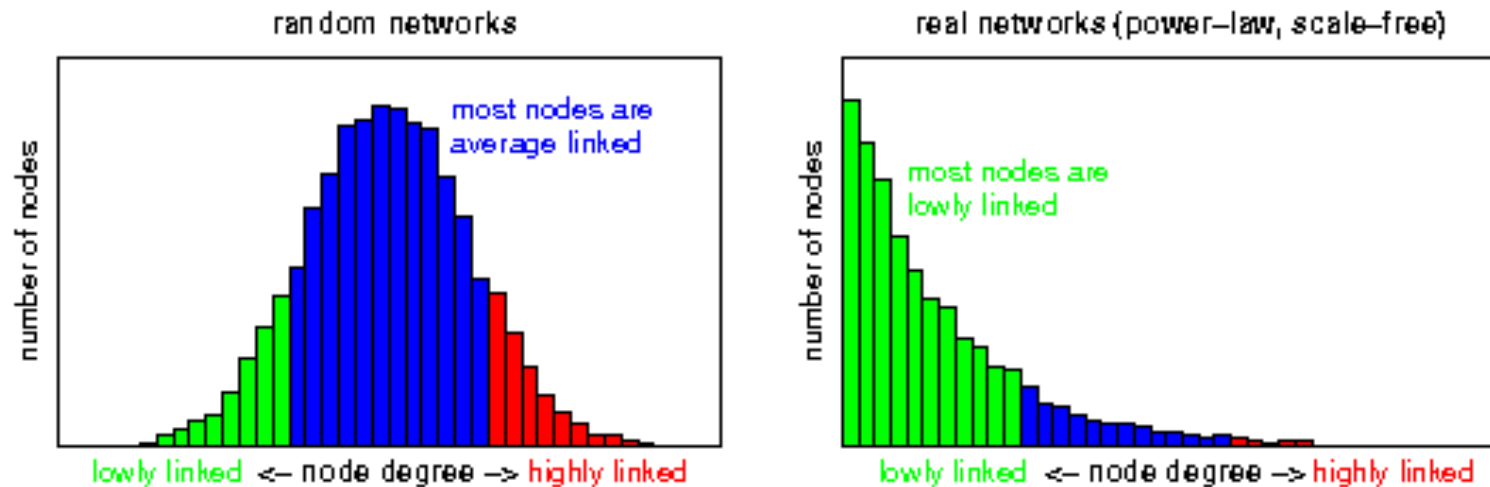
POWER LAW DISTRIBUTION

U realnim mrežama, degree metrika sledi *Power Law* distribuciju

Šta to znači?

Pretpostavimo da je u nekoj mreži prosečna vrednost za degree 3.5; ako nasumice izaberemo jedan čvor iz te mreže, šta možemo očekivati za njegovu degree vrednost?

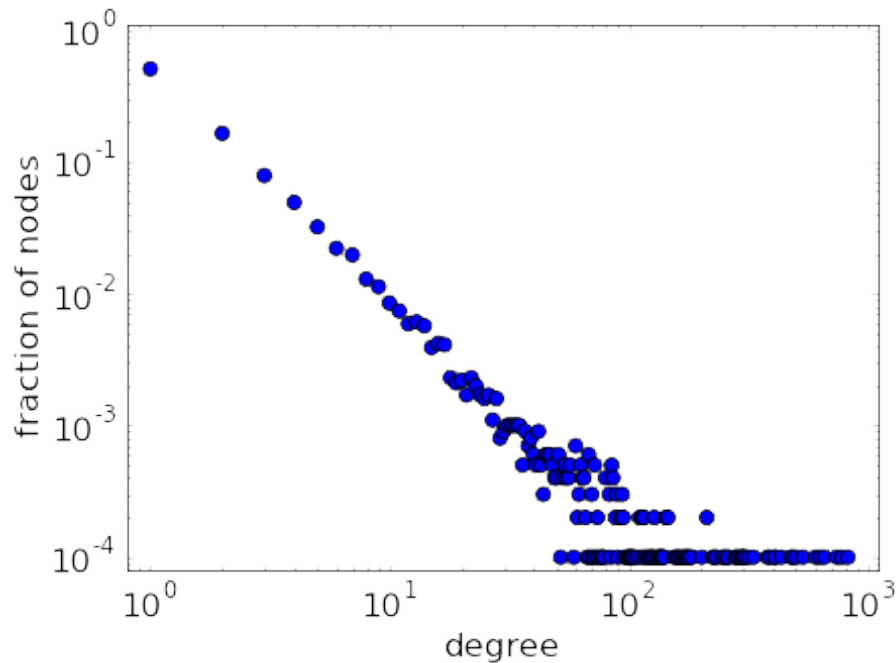
3.5 bi bio očekivani odgovor, ali to nije slučaj u realnim društvenim mrežama; već je degree značajno manji



Izvor:

<http://www.network-science.org/>

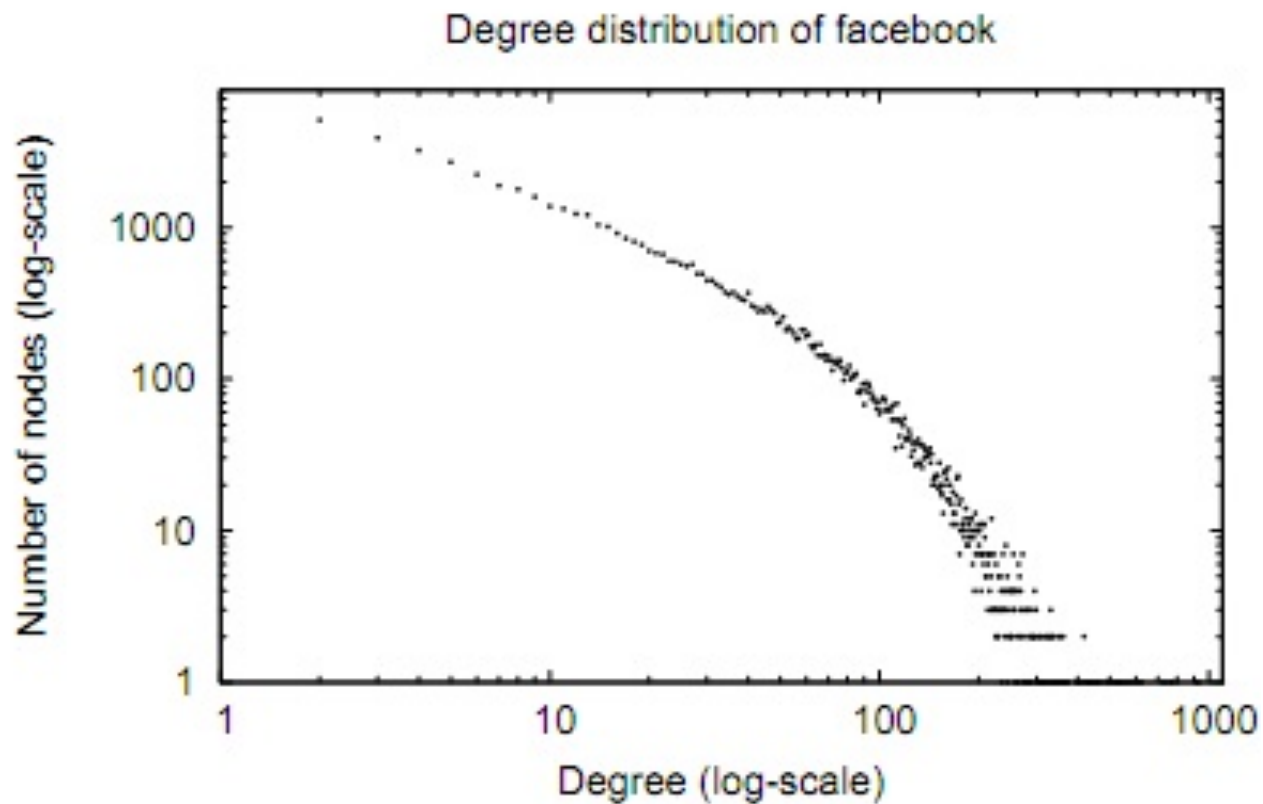
POWER LAW DISTRIBUTION



Istraživanja su pokazala da ova raspodela

- opstaje godinama, tj. ne menja se tokom vremena
- važi u brojnim realnim mrežama, posebno onim vezanim za Web

POWER LAW DISTRIBUTION



Degree distribution u
Facebook mreži

SOFTVERSKA PODRŠKA ZA ANALIZU DRUŠTVENIH MREŽA

SOFTVERSKI ALATI ZA ANALIZU DRUŠTVENIH MREŽA

- [Gephi](#) (podrška za sve platforme, besplatan, open source)
- [SocNetV](#) (podrška za sve platforme, besplatan, open source)
- R paketi za SNA (podrška za sve platforme, besplatan, open source): [igraph](#), [statnet](#)
- [Pajek](#) (Windows, besplatan)
- [NodeXL](#) (Windows; Basic verzija besplatna, Pro verzija se plaća; open source; integriše se u Excel)
- [UCInet](#) (Windows, komercijalni)
- [NetLytic](#) (Web-based, besplatan)

FORMATI ZA ČUVANJE PODATAKA O DRUŠTVENIM MREŽAMA

Široko korišćeni formati:

- .net (Pajek format)
- .dat (UCINet format)
- GML (Graph Modelling Language)
- GraphML i GEXF – popularni XML formati za zapis grafa
- csv (Comma Separated Value) – široko korišćen format za predstavljanje najrazličitih vrsta strukturiranih podataka

Većina softverskih alata omogućuje import grafa predstavljenog u bilo kom od navedenih formata

JAVNO DOSTUPNI SNA DATASETS

Network Data Repository:

<http://networkrepository.com/index.php>

UCINET datasets:

<https://sites.google.com/site/ucinetsoftware/datasets>

Gephi sample datasets:

<https://github.com/gephi/gephi/wiki/Datasets>

Stanford Large Network Dataset Collection:

<http://snap.stanford.edu/data/index.html>

Koblenz Network Data Repository:

<http://konect.uni-koblenz.de/networks/>

PRISTUP PODACIMA ONLINE DRUŠTVENIH MREŽA

Brojne online društvene mreže omogućuju pristup (javno dostupnim) podacima o članovima mreže i njihovim međusobnim relacijama posredstvom programskih interfejsa

Primeri:

- Twitter Search API* ([link](#))
- Diigo API ([link](#))
- LastFM API ([link](#))
- Flickr API ([link](#))
- ...

*API = Application Programming Interface

PRISTUP PODACIMA ONLINE DRUŠTVENIH MREŽA

Neki SNA alati omogućuju direktni import podataka sa online društvenih mreža; na primer:

- NodeXL omogućuje import podataka sa Twitter-a, YouTube-a, i Flickr-a
- [Dodaci \(plug-ins\) za NodeXL](#) omogućuju import podataka i još nekih popularnih društvenih mreža i medija
- NetLytics omogućuje import podataka sa Twitter-a, Instagram-a, RSS feeds, YouTube-a

PRISTUP PODACIMA ONLINE DRUŠTVENIH MREŽA

Postoje i javno dostupni repozitorijumi podataka sa kojih se podaci nekih online društvenih mreža mogu preuzeti u formi tzv. data dumps (fajlovi sa podacima)

Ovi podaci tipično nisu u nekom od standardnih SNA formata, već u nekom od otvorenih formata (npr. csv, json) koji se mogu transformisati u SNA format

Primeri:

- [MyPersonality project](#) – Facebook user profiles, status updates, likes, ...
- [StackExchange](#) – podaci sa Q&A foruma tipa StackOverflow
- [SNAP](#) – podaci sa različitih online društvenih mreža

PREPORUKA

Video “Converting Various Formats into Adjacency Lists for Gephi”

- <https://www.youtube.com/watch?v=cjsxFr6RIG0>
- odlični praktični saveti / instrukcije za transformaciju podataka u format pogodan za Gephi alat

NAREDNI KORAK: SENZORI KAO IZVOR PODATAKA

Detekcija društvenih mreža na osnovu podataka prikupljenih posredstvom senzora u mobilnim uređajima

*“**Sandy Pentland** and the SenseNetworks company are building tools that use cell phone towers’ ability to **collect data from cell phones** to map the locations of groups of people over time. The resulting patterns are used **to group people into “tribes” based on common, overlapping habits**. Even if two people have never met, their common use of certain kinds of spaces and transit systems soon build a link based on their shared visits to the same kinds of restaurants, theaters, office buildings, and highways or rail lines...”*

(Smith et al., 2009)

SENZORI KAO IZVOR PODATAKA

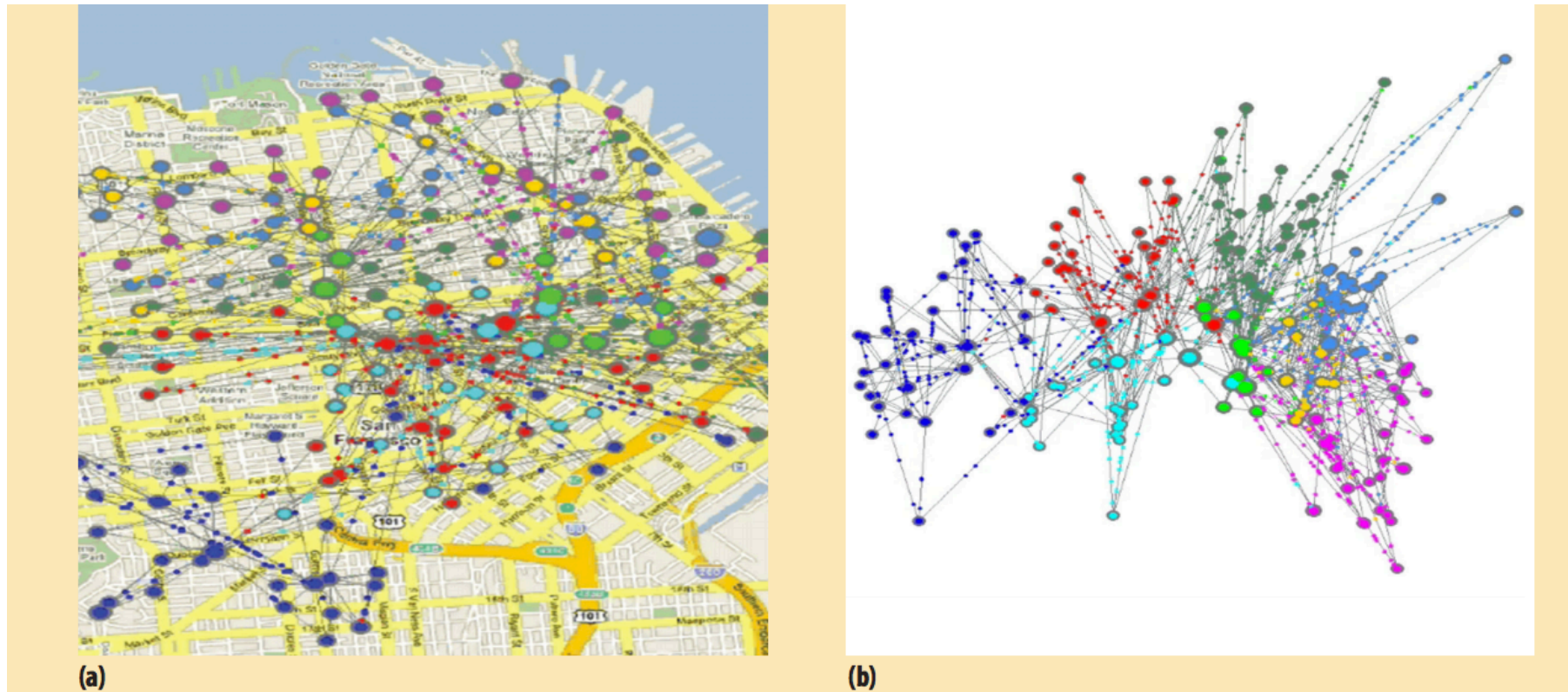


Figure 1. “Reality mining” of GPS data from mobile phones. (a) Human mobility patterns within a city, color coded by common patterns of movement. (b) The patterns reveal distinct population subgroups. Images courtesy of Sense Networks (www.sensenetworks.com).

Preporuka: “Social Physics: from Ideas to Actions” by MIT Professor Sandy Pentland ([YouTube link](#))

LITERATURA I DODATNI KORISNI SADRŽAJI NA TEMU SNA

ODLIČNA KNJIGA



Robins, G. (2015). Doing Social Networks Research: Network Research Design for Social Scientists. Sage.

<https://uk.sagepub.com/en-gb/eur/doing-social-network-research/book241817>

GEPHI KROZ PRAKTIČNE PRIMERE

- Hirst, T. Getting Started With The Gephi Network Visualisation App – My Facebook Network, Part I ([link](#))
- Hirst, T. Getting Started With Gephi Network Visualisation App – My Facebook Network, Part II: Basic Filters I ([link](#))
- Gephi Tutorial on Network Visualization and Analysis by Jane Golbeck ([YouTube link](#))
- (official) Gephi Visualization tutorial ([link](#))
- (official) Gephi Layout tutorial ([link](#))

KORISNI/INTERESANTNI SADRŽAJI NA NETU

- Short lectures by [Prof. Jane Golbeck](#) (available on YouTube)
 - Network Structure ([video](#))
 - Network Content and Structure ([video](#))
 - Building Networks ([video](#))
- [website] Network of Thrones ([link](#)) – an entire web site devoted to SNA of characters from the Game of Thrones series
- [news article] Degrees of separation: After all, it is who you know ([link](#))
- [blog post] Preparing Data for Analysis is (more than) Half the Battle ([link](#))

ZAHVALNICE I PREPORUKE

Social Network Analysis (SNA)

Dr. Giorgos Cheliotis (gcheliotis@nus.edu.sg)

Communications and New Media, National University of Singapore

Slides URL: <http://www.slideshare.net/gcheliotis/social-network-analysis-3273045>

Graph Mining Techniques for Social Media Analysis

Mary McGlohon and Christos Faloutsos, Carnegie Mellon University

Tutorial @ International Conference on Weblogs and Social Media 2008

Slides URL: <http://www.icwsm.org/2008/tutorials.shtml>