Isidora Rapajić

Multipleks mreža naučnih kolaboracija

U ovom radu predložena je analiza kolaboracionih mreža iz ugla multipleks reprezentacije. Multipleks mreže predstavljaju novi način reprezentovanja i analiziranja kompleksnih sistema. Formirana je multipleks mreža naučnih kolaboracija na osnovu naučnih radova preuzetih sa arXiv.org. Mreža ima 30 slojeva, a svaki sloj odgovara mreži kolaboracija jedne od matematičkih disciplina. Izvršena je pojedinačna i kolektivna analiza slojeva multipleks mreže. Na osnovu ispitivanja distribucije stepena čvorova i najkraće prosečne putanje, pokazali smo da slojevi multipleks mreže imaju osobine scale-free i malog sveta (eng. small-world). Takođe, utvrdili smo da algebarska topologija i kvantna algebra, spadaju u jedne od „najsocijalnijih" matematičkih disciplina na osnovu koeficijenta asortativnosti. U ovom radu, uvedene su mere strukturne similarnosti među slojevima i među čvorovima multipleks kolaboracione mreže. Na osnovu njih je izvršena klasifikacija i grupisanje matematičkih disciplina i matematičara. Za pojedine grupe disciplina, dobijene isključivo na osovu strukturnih sličnosti između slojeva multipleks mreže, pokazano je da se slažu sa već postojećim klasifikacijama. Ovakava kolektivna analiza slojeva multipleks mreže je pokazala veliku prednost u odnosu na već postojeće standardne načine analiza mreža.

Uvod

Teorija mreža (ili grafova) se u poslednje vreme pokazala kao veoma korisna u analizi i modelovanju kompleksnih sistema, tj. sistema koji se sastoje od velikog broja komponenata koji međusobno interaguju (Barabasi i Bonabeau 2003). Formalno, graf (ili mreža) je uređeni par, $G = (V, E)$, skupa čvorova V, i skupa veza (grana), tj. skupa međusobno povezanih parova čvorova, E. Grafovi se mogu podeliti na usmerene i neusmerene, zavisno da li linkovi imaju smer; i na otežnjene i neotežnjene, zavisno od toga da li veze imaju neku vrednost koja karakteriše jačinu interakcije između čvorova. Neke od vrlo dobro proučenih mreža su socijalne mreže, gde čvorovi predstavljaju korisnike, a veze njihove socijalne interakcije; mreže transporta, gde čvorovi predstavljaju gradove a veze predstavljaju puteve između gradova; internet mreže, gde čvorovi predstavljaju računare a veze predstavljaju elektronske konekcije između računara.

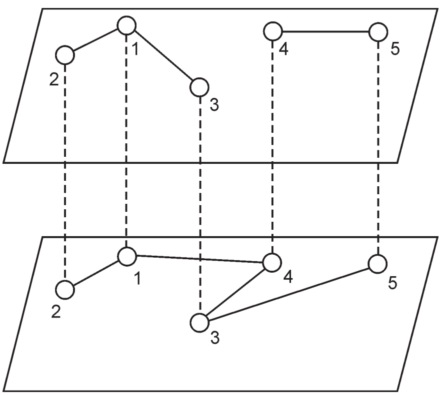
U ovom radu ćemo se baviti kolaboracionim mrežama, koje ćemo predstavljati kao neusmerene i neotežnjene grafove (Newman 2003). Kolaboracione mreže su mreže u kojima su čvorovi naučnici, a veza između njih postoji ako su zajedno publikovali naučni rad. Do sad su ispitane kolaboracione mreže naučnika iz oblasti biomedicine, fizike i računarskih nauka. Pokazano je da u svim slučajevima, naučne zajednice formiraju mreže „malog sveta" (eng. small world) koje se odlikuju uglavnom kratkim najmanjim rastojanjima između bilo koja dva nasumično odabrana čvora u mreži. Takođe, u mrežama se javlja visok stepen grupisanja (eng. clustering) što znači da će dva nasumično izabrana naučnika imati veću šansu da budu povezani ukoliko imaju trećeg naučnika u preseku (Newman 2003). Kao rezultat, ove mreže se odlikuju i velikim brojem zajednica, tj. klastera, u kojima su naučnici koji

Isidora Rapajić (1997), Aleksandrovo, Petra Drapšina 37a, učenica 4. razreda Gimnazije „Jovan Jovanović Zmaj" u Novom Sadu

MENTOR: Vladimir Gligorijević, Imperial Koledž London, Ujedinjeno Kraljevstvo

često dele zajednička interesovanja i publikuju zajedno mnogo gušće povezani nego ostatak mreže.

Međutim, ovo je ispitivano u mrežama u kojima su naučnici povezani jednim tipom veza. Ovakava reprezentacija je često limitirana na jednu vrstu naučne podoblasti i ne obuhvata celokupnu kompleksnost sistema naučnih kolaboracija. Iz toga razloga, u ovom radu, uvodimo multipleks mreže i multipleks reprezentaciju naučnih kolaboracija. Multipleks mreže se matematički mogu opisati kao, $G = (V, \{E_1, E_2, \ldots, E_M\})$, gde V predstavlja skup čvorova, a $\{E_1, E_2, \ldots, E_M\}$ predstavlja skup od M podskupova veza koji predstavljaju M različitih tipova interakcija između čvorova. Drugim rečima, u multipleks mrežama dva naučnika mogu biti povezani sa više različitih tipova veza i često se takve mreže mogu podeliti na slojeve gde je u svakom sloju jedan tip veze (slika 1)



Slika 1. Multipleks mreža koji se sastoji od 5 čvorova i 2 sloja (tj. dva tipa veza)

Figure 1. Multiplex network with 5 nodes and 2 layers

U ovom radu bavićemo se multipleks kolaboracionim mrežama matematičara koji sarađuju u $M = 30$ različitih oblasti. Mreže su konstruisane iz radova matematičara skinutih sa arXiv.org.

Cilj projekta je da se, koristeći ovu reprezentaciju, prikažu osobine matematičara i matematičkih oblasti koji se inače ne bi mogle prikazati analizom prostih, jednoslojnih mreža. Prvo, analizom pojedinačnih slojeva multipleks mreže pokazaćemo kako se matematičke discipline međusobno razlikuju. Dalje, kolektivnom analizom svih slojeva i interakcija između slojeva pokazaćemo kako se matematičke discipline međusobno grupišu u srodne kategorije. Na kraju, kolektivnom analizom svih slojeva mreže, ali ovog puta iz ugla matematičara, pokazaćemo kako se, koristeći formalizam multipleks mreže, može formulisati svestranost matematičara i kako se matematičari grupišu po svestranosti.

Materijal i metode

Analizirane su mreže kolaboracija matematičara u 30 različitih matematičkih disciplina. Radovi publikovani u periodu 2010-2014. pokrivaju 30 matematičkih disciplina (tabela 1) a preuzeti su sa arXiv.org preko arXiv API (Application programming interface). Baza arXiv.org je baza podataka naučnih radova koji su dostupni u elektronskoj formi. Naučnici sami objavljuju i arhiviraju svoje radove na arXiv.org pre zvaničnog objavljivanja u naučnim časopisima.

Na osnovu ovih radova, za svaku matematičku disciplinu, konstruisali smo mrežu kolaboracija između matematičara. Ukupno 30 mreža je konstruisano po principu koji je objašnjen u uvodu. Njihove osnovne osobine su prikazane u tabeli 1.

Za konstrukciju mreža i obradu podataka koristio se programski jezik Python i biblioteka Networkx.

Analiza pojedinačnih mreža

Mreže u realnosti nisu ni potpuno pravilne ni potpuno nasumične i zbog toga se uvode različite mere za opisivanje njihove strukture. Najvažnije mere su najkraća prosečna putanja, distribucija stepena čvora i koeficijent klasterisanja. Ovde ćemo matematički formulisati ove mere, koje ćemo koristiti da okarakterišemo strukturne osobine pojedinačnih mreža kolaboracija. Koristićemo ove strukturne osobine da poredimo matematičke discipline.

Najkraća prosečna putanja. Najkraća prosečna putanja između parova čvorova računa se po formuli:

$$a = \sum_{s,t \in V} \frac{d(s,t)}{n(n-1)}$$

Tabela 1. Matematičke discipline i njihove osnovne osobine. Skraćenice su preuzete sa arXiv.org

Mreža	Skraćenica	Broj čvorova	Broj veza
Komutativna algebra	AC	1638	2279
Algebarska geometrija	AG	4463	7000
Analiza parcijalnih diferencijalnih jednačina	AP	5846	8911
Algebarska topologija	AT	1674	3684
Parcijalne diferencijalne jednačine	CA	3113	3873
Kombinatorika	CO	7160	10840
Teorija kategorija	CT	716	669
Kompleksne promenljive	CV	2129	2342
Diferencijalna geometrija	DG	5892	7494
Dinamički sistemi	DS	4921	5849
Funkcionalna analiza	FA	4092	4844
Opšta topologija	GN	783	813
Teorija grupa	GR	3236	3763
Geometrijska topologija	GT	3510	6075
Teorija informacija	IT	6570	14126
K-teorija i homologija	KT	868	816
Logika	LO	1386	1611
Metrička geometija	MG	1772	1793
Matematička fizika	MP	10261	13214
Numerička analiza	NA	3889	5080
Teorija brojeva	NT	4550	5613
Algebra operatora	OA	2524	2893
Optimizacija i upravljanje	OC	3997	5148
Verovatnoća	PR	8605	12060
Kvantna algebra	QA	3188	5392
Prsteni i algebra	RA	2738	2945
Teorija reprezentacija	RT	3617	4020
Simplektička geometrija	SG	1692	1567
Spektralna teorija	SP	1881	1992
Teorija statistike	ST	4048	5406

gde je V skup čvorova u mreži, $d(s, t)$ najkraća putanja od čvora s do čvora t, a n broj čvorova u mreži. Fenomen mreža „malog sveta” se javlja ukoliko je vrednost najkraće prosečne putanje (l) mala u odnosu na ukupan broj čvorova u mreži (npr. pokazano je da u socijalnim mrežama $l < 6$, Newman 2003). Formalno, ako je a ~ log$n$, gde je n broj čvorova u mreži, za mrežu se kaže da je „small-world”. Najkraća udaljenost između dva čvora u mreži računata je koristeći Dijkstra algoritam.

Distribucija stepena čvora. Stepen čvora, k_i, je broj grana povezanih sa čvorom. Verovatnoća da izabran čvor ima stepen k je P_k. Distribucija stepena je:

$$P_k = \sum_{k'=k}^{\infty} p_{k'}$$

što je verovatnoća da je stepen veći ili jednak k. (Newman 2003)

U analizi naših mreža koristićemo srednju vrednost stepena, koju računamo po formuli:

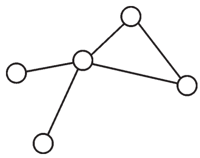
$$k_{\mathrm{sr}} = \sum_{i=1}^{n} k_i$$

gde je n broj čvorova a k_i stepen čvora i (Newman 2003).

Koeficijent klasterisanja. Koeficijent klasterisanja u kontekstu naučnih kolaboracija znači da postoji velika verovatnoća da će dva naučnika u mreži biti povezana ako imaju treću osobu u preseku. Računa se po formuli:

$$C = \frac{3 \times \text{broj trouglova u mreži}}{\text{broj povezanih trojki čvorova}}$$

gde povezana trojka predstavlja strukturu u kojoj je jedan čvor povezan sa druga dva čvora a ta dva čvora nisu međusobno povezana (slika 2) (Newman 2003).



Slika 2. Ilustracija definicije koeficijenta klasterisanja za srednji čvor. Mreža ima jedan trougao i osam povezanih trojki čvorova, dakle koeficijent grupisanja je 3×1/8 = 3/8.

Figure 2. Illustration of the definition of clustering coefficient. This network has one triangle and eight connected triples and therefore has a clustering coefficient ×1/8 = 3/8.

Asortativnost. Asortativnost je mera korelacije između stepena povezanih čvorova. Računa se kao Pearsonov korelacioni koeficijent između stepena čvora k_i, srednje vrednosti stepena njegovih suseda k_i^{nn}. Vrednost asortativnosti je mera socijalnosti čvorova; ukoliko je ona pozitivna to znači da sa vrednošću stepena čvora raste i srednja vrednost stepena njegovih suseda, dakle naučnici su socijalni. Koristeći formulu za Pearsonov koeficient korelacije, koeficijent asortativnosti se može zapisati kao:

$$r = \frac{n\sum k_i k_i^{nn} - \left(\sum k_i\right)\left(\sum k_i^{nn}\right)}{\sqrt{n\left(\sum k_i^2\right) - \left(\sum k_i\right)^2}\sqrt{n\left(\sum k_i^{nn2}\right) - \left(\sum k_i^{nn}\right)^2}}$$

gde je n ukupan broj čvorova u mreži.

Analiza multipleks mreže kolaboracijâ u kontekstu disciplina

Slojevi multipleks mreže se formiraju nad zajedničkim skupom čvorova u različitim mrežama. U ovom radu, svaki sloj predstavlja jednu kolaboracionu mrežu matematičara koji sarađuju u nekoj od disciplina prikazanih u tabeli 1. Stoga naša multipleks mreža ima 30 slojeva.

Kako bi se utvrdilo koje se matematičke discipline grupišu, potrebno je uvesti mere sličnosti između slojeva. Novina ovog rada se ogleda u tome, što uvodimo mere topološke similarnosti između slojeva multipleks mreže (tj. različitih matematičkih disciplina u slučaju naše multipleks mreže kolaboracija). Naime, definisaćemo meru poklapanja veza, $J(G_u,G_v)$, između slojeva G_u i G_v i korelaciju između stepena čvorova koji se nalaze u preseku slojeva G_u i G_v, $r(G_u, G_v)$.

Poklapanje veza između slojeva. Jedna od osobina koja je od interesa jeste poklapanje veza u slojevima. Neka su veze e i e' takve da $e(i, j) \in G_u$ i $e'(i', j') \in G_v$, gde u i v predstavljaju dva različita sloja. Kažemo da se veze poklapaju ako i samo ako je $e = e'$, $i = i'$ i $j = j'$ (Loe i Jensen 2015).

U ovom radu, kao meru poklapanja slojeva (tj. meru similarnosti među matematičkim disciplinama) uvodimo Jaccardov indeks poklapanja veza. Definiše se kao:

$$J(G_u, G_v) = \frac{E_u \cap E_v}{E_u \cup E_v}$$

gde su E_u i E_v skupovi veza u slojevima multipleks mreže u i v. Vrednost Jaccradovog indeksa je u intervalu [0, 1], gde 0 znači da nema preklapanja između slojeva, a 1 znači da su slojevi identični. Na osnovu ove mere konstruišemo matricu sličnosti između parova matematičkih disciplina, S_J. U našem slučaju, matrica S_J je simetrična matrica dimenzija 30 × 30, sa elementima $S_J(u, v) = J(G_u, G_v)$.

Korelacija između stepena čvorova u preseku slojeva. Druga mera sličnosti među matematičkim disciplinama koja će biti predstavljena u ovom radu je korelacija između stepena čvorova koji se nalaze u preseku slojeva. Ovde je računamo preko Pearsonovog koeficijenta korelacije. Naime, koeficijent se računa po formuli:

$$r(G_u, G_v) =$$

$$= \frac{n\sum k_i^u k_j^v - (\sum k_i^u)(\sum k_j^v)}{\sqrt{n(\sum k_i^{u2}) - (\sum k_i^u)^2}\sqrt{n(\sum k_j^{v2}) - (\sum k_j^v)^2}}$$

gde je n broj čvorova u preseku slojeva G_u i G_v, a k_i^u i k_j^v su stepeni čvorova i i j u mrežama v i u. Vrednost korelacije je u intervalu [–1, 1], gde 1 znači maksimalnu korelaciju, a –1 znači antikorelaciju između slojeva. Slično kao kod poklapanja veza, konstruišemo matricu sličnosti između parova disciplina, S_r. U našem slučaju, matrica S_r je simetrična matrica dimenzija 30×30, sa elementima $S_r(u,v) = r(G_u, G_v)$.

Obe navedene mere smo koristili za računanje similarnosti između parova matematičkih disciplina (tj. slojeva multipleks mreže). Na osnovu matrica sličnosti i primenjujući algoritam za hijerarijsko grupisanje (Rokach i Maimon 2005), utvrdili smo koje se oblasti međusobno grupišu. Prednost ovog algoritma je u tome što ne moramo unapred znati broj grupa.

Algoritam prvo formira „linkage” matricu u kojoj su sadržani indeksi disciplina ili grupacija disciplina i rastojanje (udaljenost) između njih. Za formiranje „linkage” matrice korišćen je Wardov metod minimalne varijanse (eng. Ward’s minimum variance method). Konkretno, u našem slučaju, rastojanje između disciplina v i u se računa na osnovu matrica sličnosti. Naime:

$$d(v, u) = 1 - S_J(v, u)$$

u slučaju preklapanja veza između slojeva ili:

$$d(v, u) = 1 - S_R(v, u)$$

u slučaju korelacija stepena čvorova u preseku slojeva.

Algoritam hijerarhijski grupiše discipline na osnovu ovog rastojanja. Zatim, konstruiše dendogram sa koga se može uočiti hijerarhija između klastera i rastojanja između njih. Presekom dendrograma na odgovarajućoj visini (eng. cut-off) možemo zaključiti koje oblasti se grupišu i koliko grupa ima. Vrednost za cut-off je zadata kao polovina maksimalnog rastojanja među disciplinama.

Analiza multipleks mreže kolaboracijâ u kontekstu matematičara

Za razliku od prethodnog pristupa u kojem je fokus bio na matematičkim disciplinama u kontekstu multipleks mreže kolaboracija, u ovom odeljku se više fokusaramo na same matematičare i njihove pozicije u multipleks mreži. Konkretno, da bismo opisali povezanost matematičara u skupu svih slojeva (tj. disciplina) multipleks mreže, uvodimo tridesetodimenzionalni vektor, za svakog matematičara i:

$$m_i = [m_i^1, m_i^2, \ldots, m_i^{30}]$$

gde, m_i^j predstavlja stepen čvora i u sloju j. Broj nenultih koordinata ovog vektora predstavlja stepen svestranosti matematičara.

Kako bismo utvrdili koji se naučnici grupišu, na osnovu ovog vektora uvodimo novu meru – kosinusnu sličnost. Kosinusna sličnost između dva naučnika, i i j, se računa po formuli:

$$S_{ij} = \frac{\sum_{i=1}^{30} m_i m_j}{\sqrt{\sum_{i=1}^{30} m_i^2}\sqrt{\sum_{i=1}^{30} m_i'^2}} \tag{1}$$

gde su m i m' tridesetodimenzionalni vektori za dva naučnika, a m_i je i-ta koordinata vektora.

Pre računanja kosinusne sličnosti potrebno je normalizovati koordinate vektora, da bismo uračunali razlike između struktura i veličina slojeva multipleks mreže. To radimo tako što formiramo matricu naučnika, M, tj. redovi matrice su vektori m_i, a kolone su matematičke discipline. Da bismo povećali efikasnost našeg računanja, ograničili smo se samo na matematičare koji su radili u više od 5 oblasti. U tom slučaju, ukupan broj matematičara koje razmatramo je 4866, stoga je dimenzionalnost matrice M je 4866 × 30.

U ovom radu, vektori su normalizovani koristeći formulu za z-skor:

$$C_i = \frac{C_i - \langle C_i \rangle}{\sigma(C_i)}$$

gde je $i = 1, 2, \ldots, 30$, C_i je i-ta kolona matrice M, $\langle C_i \rangle$, $\sigma(C_i)$ standardna devijacija. Zatim formiramo simetričnu matricu similarnosti S (na osnovu formule 1), gde elementi matrice predstavljaju kosinusnu sličnost za svaka dva naučnika.

Koristeći istu proceduru za hijerarhijsko grupisanje matematičara na osnovu rastojanja $D_{ij} = 1 - S_{ij}$, konstriušemo dendrogram i na osnovu njega dobijamo grupe matematičara.

Rezultati i diskusija

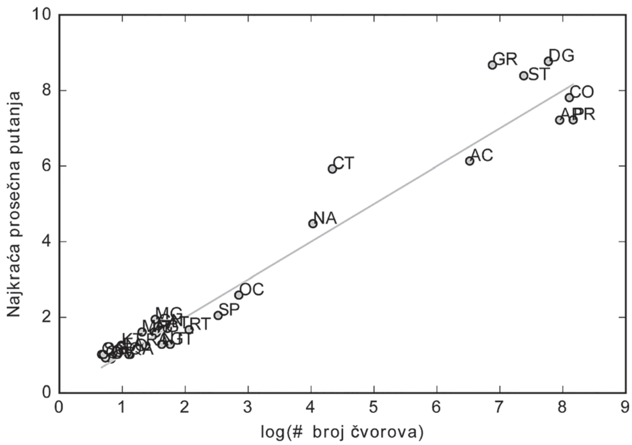
Pojedinačne mreže kolaboracijâ matematičkih disciplina

U tabeli 2 prikazane su strukturne osobine pojedinačnih mreža (slojeva multipleks mreže) kolaboracija. U kolonama su prosečna vrednost stepena, maksimalna vrednost stepena, prosečan koeficijent grupisanja, prosečna najkraća putanja i asortativnost.

Srednja vrednost stepena čvora je slična za većinu kolaboracionih mreža (i kreće se između 2 i 3), osim za kolaboracione mreže algebarske topologije (AT) i teorije informacija (IT), gde je srednja vrednost stepena čvorova veća od 4. Ovo ukazuje na veću dinamiku kolaboracije između matematičara ovih disciplina, a time i na veći broj prosečnih veza po čvoru. Takođe, proverom distribucija stepena čvorova ovih mreža utvrđeno je da su sve mreže ovde analizirane „scale-free”, tj. da se distribucije odlikuju stepenim zakonom. Pored toga što se IT odlikuje velikim

Tabela 2. Strukturne osobine kolaboracionih mreža matematičara u različitim disciplinama

Mreža	Pros. stepen	Maks. stepen	Pros. koeficijent klasterisanja	Najkraća pros. putanja	Asort.
AC	2.783	26	0.461	6.131	0.225
AG	3.137	52	0.374	6.593	0.167
AP	3.048	32	0.462	7.217	0.173
AT	4.401	62	0.418	9.465	0.984
CA	2.488	34	0.487	9.104	0.156
CO	3.028	63	0.497	7.813	0.121
CT	1.868	12	0.386	5.926	0.173
CV	2.201	27	0.399	11.154	0.165
DG	2.544	32	0.412	8.772	0.202
DS	2.378	42	0.476	9.252	0.162
FA	2.367	43	0.461	11.371	0.067
GN	2.076	43	0.442	3.818	0.019
GR	2.326	22	0.443	8.679	0.226
GT	3.462	81	0.425	7.733	0.912
IT	4.301	191	0.622	5.770	0.119
KT	1.881	14	0.395	5.986	0.338
LO	2.325	85	0.477	5.631	-0.019
MG	2.023	12	0.471	4.196	0.308
MP	2.576	29	0.501	12.424	0.272
NA	2.613	26	0.583	10.174	0.326
NT	2.467	27	0.415	8.522	0.189
OA	2.292	26	0.446	7.609	0.213
OC	2.576	34	0.582	11.570	0.149
PR	2.803	76	0.464	7.223	0.201
QA	3.383	67	0.446	8.489	0.956
RA	2.151	19	0.443	10.346	0.218
RT	2.223	27	0.408	8.338	0.563
SG	1.852	11	0.391	5.372	0.397
SP	2.118	11	0.535	5.072	0.508
ST	2.671	28	0.512	8.393	0.138



Slika 3. Najkraća prosečna putanja, *a*, u zavisnosti od log*n* prikazana za svaki sloj mreže (discipline). Mreže koje su bliže liniji ispoljavaju efekat malog sveta.

Figure 3. Average shortest path (*a*) vs. log*n*, where *n* is number of nodes. Networks closer to the line are “small-world”.

prosečnim stepenom čvorova, odlikuje se takođe i najvećim koeficijentom grupisanja. Vratimo se ponovo na interpretaciju ovog koeficijenta: ako matematičar A sarađuje sa matematičarom B i nezavisno sarađuje i sa matematičarom C, koeficijent grupisanja ukazuje na to kolika je verovatnoća da i matematičari B i C međusobno sarađuju. U slučaju IT mreže, verovatnoća za formiranje ovog trougla, A-B-C je znatno veća nego kod ostalih mreža, što ukazuje da međusobna poznanstva i prethodne saradnje između matematičara često indukuju nove saradnje (tj. IT mreža ima mnogo veći broj trouglova nego ostale mreže). Ovaj fenomen je najviše izražen u slučaju IT kolaboracione mreže, a najmanje u slučaju AG.

Da bismo ispitali efekat „small-world” (mali svet) u mrežama matematičkih kolaboracija, za svaku mrežu računali smo najkraću prosečnu putanju. Najkraća prosečna putanja u zavisnosti od logaritma broja čvorova te mreže je prikazana na slici 3.

Za mreže koje su blizu prave $y = x$ (tj. $a \sim \log n$), se može reći da su „small-world”, što je u našem slucaju karakteristika svih mreža. Manja odstupanja su uočena kod teorije grupa (GR) i teorije kategorija (CT). Veliki koeficijent klasterisanja kod gotovo svih mreža je još jedan argument više u prilog ovoj tvrdnji (videti tabelu 2).

U kontekstu asortativnosti, većina kolaboracionih mreža ispoljava „socijalnost”, okarakterisanu pozitivnom vrednošću koeficijenta asortativnosti. Ekstremno „socijalni” karakter je uočen u kolaboracionim mrežama algebarska topologija (AT) i kvantna algebra (QA), gde je asortativnost je skoro 1, što znači da u tim mrežama matematičari sa velikim brojem kolaboratora teže da se povezuju sa matematičarima koji takođe imaju veliki broj kolaboratora, i obrnuto.

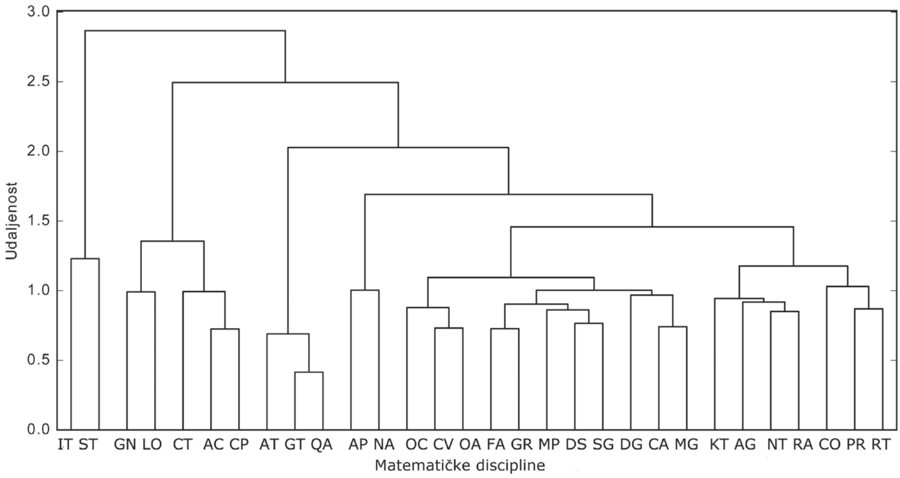
S druge strane, mreža logika (LO) ima negativnu asortativnost što znači da u toj oblasti naučnici nisu socijalni. U ostalim mrežama asortativnost je pozitivna.

Multipleks mreža naučnih kolaboracijâ u kontekstu matematičkih disciplina

Na osnovu matrica sličnosti koje su dobijene za Jaccardov indeks i Pearsonov koeficijent, konstruisani su dendogrami sa kojih se vidi koje discipline se grupišu (slike 4 i 5).

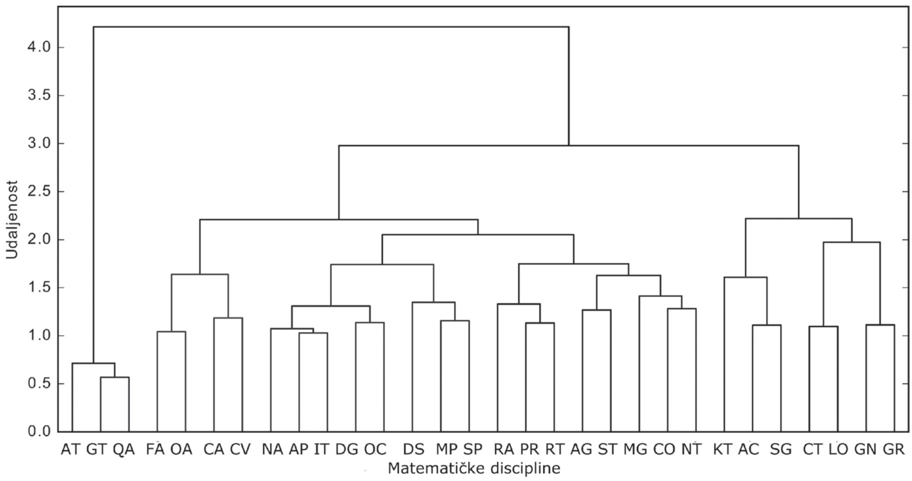
Poredeći dobijene rezultate sa *Mathematics subject classification* (MSC) Američkog matematičkog društva, zaključak je sledeći:

- Obe metode su pokazale da se grupišu algebarska topologija (AT), geometrijska topologija (GT) i kvantna algebra (QA), što je i očekivano, budući da su ove tri oblasti srodne.
- Obe metode su pokazale da se grupišu discipline opšta topologija (GN), logika (LO) i teorija kategorija (CT), što nije očekivano jer po MSC pripadaju različitim matematičkim oblastima.
- Obe metode su pokazale da se grupišu discipline prsteni i algebra (RA), verovatnoća (PR), teorija reprezentacija (RT), algebarska geometrija (AG), kombinatorika (CO) i teorija brojeva (NT). Isto važi i za discip-



Slika 4. Dendogram konstruisan za Jaccardov index. Na horizontalnoj osi su matematičke discipline (pogledati tabelu 1 za skraćenica), a na vertikalnoj udaljenost.

Figure 4. Jaccard index dendogram. Mathematical disciplines are on the horizontal and distances are on the vertical axis.

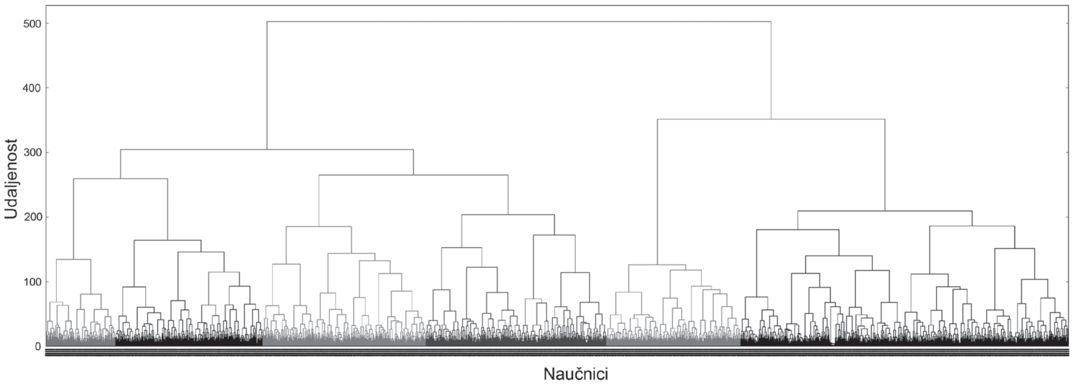


Slika 5. Dendogram konstruisan za Pearsonov koeficijent. Na horizontalnoj osi su matematičke discipline (pogledati tabelu 1 za značenja skraćenica), a na vertikalnoj udaljenost.

Figure 5. Pearson coefficient dendogram. Mathematical disciplines are on the horizontal and distances are on the vertical axis

line funkcionalna analiza (FA), algebra operatora (OA), klasična analiza parcijalnih diferencijalnih jednačina (CA) i kompleksne promenljive, zatim diferencijalna geometrija (DG), dinamički sistemi (DS), optimizacija i upravljanje (OC) i matematička fizika (MP), i numerička analiza (NA) i analiza parcijalnih diferencijalnih jednačina (AC).

– Prema Pearsonovom koeficijentu korelacije, tri discipline: K-teorija i homologija (KT), komutativna algebra (AC) i simp-



Slika 6. Dendogram konstruisan za kosinusnu sličnost. Na horizontalnoj osi su naučnici, a na vertikalnoj udaljenost.

Figure 6. Cosine similarity dendogram. Scientists are on the horizontal and distances on the vertical axis.

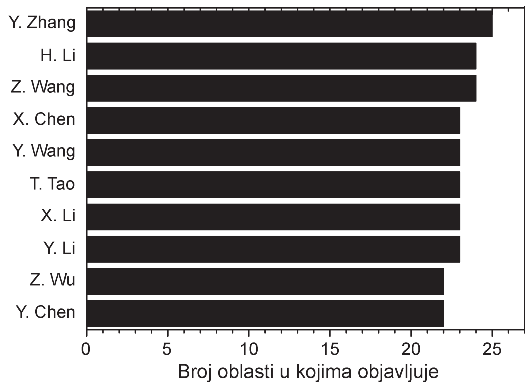
lektička geometrija (SG) formiraju grupu, a prema Jaccardovom indeksu sličnosti ove discipline se nalaze u tri različite grupe, što u ovom slučaju pokazuje da je Pearsonov koeficijent više u skladu sa postojećim klasifikacijama, jer se zna da su ove discipline srodne.

– Prema Jaccardovom indeksu sličnosti, teorija informacija (IT) i teorija statistike (ST) formiraju grupu, a prema Pearsonovom koeficijentu nalaze se u dve različite grupe, što u ovom slučaju pokazuje da je Jaccardov indeks u većoj meri saglasan sa postojećim klasifikacijama.

Multipleks mreža naučnih kolaboracija u kontekstu naučnika

Koristeći vektor *m* rangirali smo matematičare na osnovu njihove svestranosti. Prvih 10 najsvestranijih matematičara (odnosno oni koji su radili u najviše oblasti) prikazani su na slici 7.

Vidimo da ne postoji ni jedan naučnik koji se pojavljuje u svih 30 matematičkih disciplina, maksimalan broj je 25. Detaljnim istraživanjem na Google Scholaru, uočeno je da su ovi matematičari takođe jako citirani (h-index > 40).



Slika 7. Deset najsvestranijih naučnika

Figure 7. The most versatile scientists

Koristeći metodologiju za hijerarhijsko klasterisanje, primenjenu na matricu similarosti *S*, pokazali smo da se naučnici takođe grupišu na osnovu njihove aktivnosti u različitim slojevima multipleks mreže. Na slici 6 prikazan je dendogram sa koga se jasno izdvaja 6 grupa naučnika. Detaljna analiza i validacija ovih klastera može biti predmet budućeg istraživanja.

Zaključak

U ovom radu predložena je analiza multipleks mreža naučnih kolaboracija matematičara. Svaka matematička disciplina je predstavljena kao jedan sloj multipleks mreže. Na osnovu pojedinačne analize strukture slojeva mreže pokazali smo veliku raznovrsnost matematičkih disciplina, kao i zajedničke karakteristike, kao sto je karakteristika „malog sveta" kojom se odlikuju kolaboracione mreže u svim disciplinama. Još jedna od zajedničkih karakteristika svih mreža je veliki koeficijent klasterisanja, koji implicira da su reputacija i poznanstvo važni preduslovi za uspostavljanje kolaboracija. Kolektivnom analizom slojeva multipleks mreže smo otkrili grupacije disciplina kao i matematičara koje se inače ne bi mogle otkriti pojedinačnom analizom slojeva (tj. disciplina).

Detaljna analiza grupacija matematičara na osnovu njihove svestranosti i aktivnosti u svim slojevima multipleks mreže može biti predmet budućeg istraživanja. Formalizam poređenja matematičara u multipleks mrežama i generalizacija stepena čvorova u multipleks mreži, koji smo uveli u ovom radu, može poslužiti kao startna tačka daljim istraživanjima kolaboracionih mreža i multipleks mreža uopšte.

Zahvalnost. Zahvaljujem se mentoru Vladimiru Gligorijeviću, rukovodiocu seminara fizike Vladanu Pavloviću na korisnim savetima i nesebičnoj pomoći pri izradi ovog rada.

Literatura

Barabasi A. L., Bonabeau E. 2003. Scale-free networks. *Scientific American*, **288** (5): 60.

Loe C. W., Jensen H. J. 2015. Comparison of Communities Detection Algorithms for Multiplex. *Physica*, **A 431**: 29.

Newman M. E. J. 2003. The Structure and Function of Complex Networks. *Society of industrial and applied mathematics review,* **45** (2): 167.

Rokach L., Maimon O. (ur.) 2005. Clustering methods. *Data mining and knowledge discovery handbook*. Springer, str. 321-325.

Isidora Rapajić

Multiplex Network of Scientific Collaboration

In this paper, a new way of analyzing collaboration networks by using multiplex network representation has been proposed. Multiplex networks represent a new way of modeling and analyzing complex systems. In this paper, a multiplex collaboration network is formed based on papers published on arXiv.org. Our network has 30 layers, where each layer corresponds to a mathematical discipline. Single and collective analysis of layers was carried out. By examining the degree distribution of nodes and shortest paths between them, we have shown that layers of multiplex network are "scale-free" and "small-world". Furthermore, we found that scientists from Algebraic Topology and Quantum Algebra networks are the most social according to the assortativity coefficient.

Similarity measures between layers and nodes of the multiplex networks are introduced in this paper for the first time. Based on these measures, we classified mathematical disciplines and mathematicians. Some of the clusters obtained solely from these structural similarity measures are showed to be in an agreement with already existing classifications. This type of collective analysis of network layers has a great advantage over the standard ways of network analysis.