

УНИВЕРЗИТЕТ У БЕОГРАДУ
ЕЛЕКТРОТЕХНИЧКИ ФАКУЛТЕТ



АНАЛИЗА СОЦИЈАЛНИХ МРЕЖА

Други пројектни задатак

Ментори:

доц. др Марко Мишић
проф. др Јелица Протић
ас. мс Предраг Обрадовић

Кандидати:

Марина Вучковић 2020/3261
Божовић Александра 2020/3406

Београд, Фебруар 2022.

САДРЖАЈ

САДРЖАЈ	2
1. УВОД.....	3
2. АНАЛИЗА ПРИМАРНОГ СКУПА ПОДАТАКА И СТАТИСТИЧКА ОБРАДА.....	4
2.1. АНАЛИЗА ПРИМАРНОГ СКУПА ПОДАТАКА	4
2.2. СТАТИСТИЧКА ОБРАДА ПОДАТАКА	4
3. МОДЕЛОВАЊЕ <i>SNET</i> МРЕЖЕ.....	10
3.1. ОСНОВНА КАРАКТЕРИЗАЦИЈА МРЕЖЕ <i>SNET</i>	10
3.2. АНАЛИЗА МЕРА ЦЕНТРАЛНОСТИ	15
3.3. ДЕТЕКЦИЈА КОМУНА МРЕЖЕ <i>SNET</i>	18
4. МОДЕЛОВАЊЕ <i>SNETF</i> МРЕЖЕ	22
4.1. ОСНОВНА КАРАКТЕРИЗАЦИЈА <i>SNETF</i> МРЕЖЕ	22
4.2. МЕРЕ ЦЕНТРАЛНОСТИ	26
4.3. ДЕТЕКЦИЈА КОМУНА МРЕЖЕ <i>SNETF</i>	30
5. МОДЕЛОВАЊЕ <i>SNETT</i> МРЕЖЕ.....	34
5.1. ОСНОВНЕ ОСОБИНЕ МРЕЖЕ	34
5.2. АНАЛИЗА МЕРА ЦЕНТРАЛНОСТИ	37
5.3. ДЕТЕКЦИЈА КОМУНА МРЕЖЕ <i>SNETT</i>	40
6. МОДЕЛОВАЊЕ <i>USERNET</i> МРЕЖЕ	43
6.1. ОСНОВНА КАРАКТЕРИЗАЦИЈА <i>USERNET</i> МРЕЖЕ	43
6.2. АНАЛИЗА МЕРА ЦЕНТРАЛНОСТИ	48
6.3. ДЕТЕКЦИЈА КОМУНА МРЕЖЕ <i>USERNET</i>	52
ЛИТЕРАТУРА.....	54

1. Увод

Циљ пројектног задатка је да се практично примени стечено теоријско знање из предмета Анализа социјалних мрежа. Извршена је анализа и обрада скупа података социјалне мреже *Reddit*, издвајањем одређених скупова података моделоване су различите мреже, потом је извршена је анализа и визуелизација мрежа.

Reddit је веб страница која је огранизована у виду форума, где корисници постављају, коментаришу и оцењују веб садржај. Садржај је подељен у сабредите (*subreddit*), који су у највећем броју случајева тематски или географски организовани. Форум постоји од 2005. године и тренутно биљежи око 52 милиона дневно активних корисника. С обзиром на величину форума и велики број активних корисника, за анализу је узета 2008 година, због релативно малог обима података и догађаја који су се у њој десили, као што је светска економска криза. Анализиране су објаве (*submssion*) и коментари (*comment*).

У другом поглављу задатка дат је опис извршних обрада и прелиминарних анализа сировог скупа података, при чему су искључени непотребни и непотпуни подаци, а извдојили су се неопходни подаци. У другом делу поглавља је дат скуп одговора на прву групу истраживачких питања везаних за статистичку обраду података.

Треће поглавље обрађује мрежу SNET (*subreddit network*) која моделује интеракције сабредита на основу активности корисника. Грана између два сабредита постоји, уколико постоји корисник који је био активан на оба сабредита у виду коментара или објаве.

У поглављу 4. је обрађена SNETF (*filtered subreddit network*) у којој су филтриране гране мале тежине, и одбачене су све гране чије су тежине испод изабраног прага.

У петом поглављу је анализирана SNETT (*targeted subreddit network*) мрежа, коју чине сабредити који су везани за тему економске кризе. На крају овог потпоглавља је дато поређење SNet и SNetT мрежа.

Мрежа корисника, UserNet, моделује интеракције између корисника, у виду коментара на објаву или на коментар, и њена анализа је дата у шестом поглављу.

2. АНАЛИЗА ПРИМАРНОГ СКУПА ПОДАТАКА И СТАТИСТИЧКА ОБРАДА

У првом потпоглављу је описана анализа примарног скупа података, која је довела до секундарног скупа података и издвојене су уклоњене колоне, које нису од значаја за даљу анализу. Потом, у другом потпоглављу је одговорено на питања из статистичке анализе за секундарни скуп података.

2.1. Анализа примарног скупа података

Прво је извршена провера да ли су сви идентификатори објава и коментара јединствени и да нису празни стрингови. Закључак је да су сви различити и да нису празни, а провјере да ли су вредности „*nan*“ или „*null*“ нису вршене јер су идентификатори типа стринг, па би и те вредности биле валидне.

У следећем кораку су уклоњене колоне које нису од значаја. Тако су за табелу објава уклоњене следеће колоне: *score*, *distinguished*, *domain*, *stickied*, *locked* и *hide_score*. У табели коментара су уклоњене колоне: *score*, *distinguished*, и *controversiality*.

Коментари који немају свог родитеља у табели објава нису искључени из секундарног скупа података јер су направљени у години која се посматра, и самим тим треба да буду дио статистичке анализе.

2.2. Статистичка обрада података

1. У посматраном периоду постоји **5032** различитих сабредита.

1.a) Сабредити који су најважнији по броју корисника су дати у наредној табели. У разматрање нису узети корисници који су избрисани.

Табела 2.2.1. Најважнији сабредити по броју корисника

Subreddit	Subreddit_id	Број корисника
reddit.com	t5_6	163779
politics	t5_2cneq	38374
pics	t5_2qh0u	29753
technology	t5_2qh16	28337
funny	t5_2qh33	28186
entertainment	t5_2qh0f	26360
science	t5_mouw	25854

programming	t5_2fwo	25819
business	t5_2qgzg	25253
worldnews	t5_2qh13	24937

1.6) Који сабредити су најважнији по броју коментара?

Табела 2.2.2. Најважнији сабредити по броју коментара

Subreddit	Subreddit_id	Број коментара
reddit.com	t5_6	1884629
politics	t5_2cneq	1136188
programming	t5_2fwo	449840
pics	t5_2qh0u	414966
science	t5_mouw	333795
worldnews	t5_2qh13	333123
WTF	t5_2qh61	279226
funny	t5_2qh33	257333
technology	t5_2qh16	222590
AskReddit	t5_2qh1i	214145

2. Просечан број забележених корисника активних у посматраном периоду по сабредиту је $226641 / 5032 = 45.04$.
3. Корисници са највећим бројем објава и коментара су дати у наредним табелама.

Табела 2.2.3. Корисници са највећим бројем објава

Корисник	Број објава
gst	18870
qgyh2	12238
rmuser	9822

twolf1	8597
IAmperfectlyCalm	8308
qazamisan	6927
charlatan	5998
igeldard	5373
noname99	5334
democracy101	5332

Табела 2.2.4. Корисници са највећим бројем коментара

Корисник	Број коментара
NoMoreNicksLeft	13480
malcontent	12159
matts	11672
mexicodoug	9169
7oby	9161
aletoledo	8085
mutatron	7771
otakucode	7759
redditcensoredme	7468
h0dg3s	7439

4. Корисници који су активни на највећом броју сабредита су дати у наредној табели.

Табела 2.2.5. Корисници са највећим бројем коментара

Корисник	Број сабредита
MrKlaatu	181

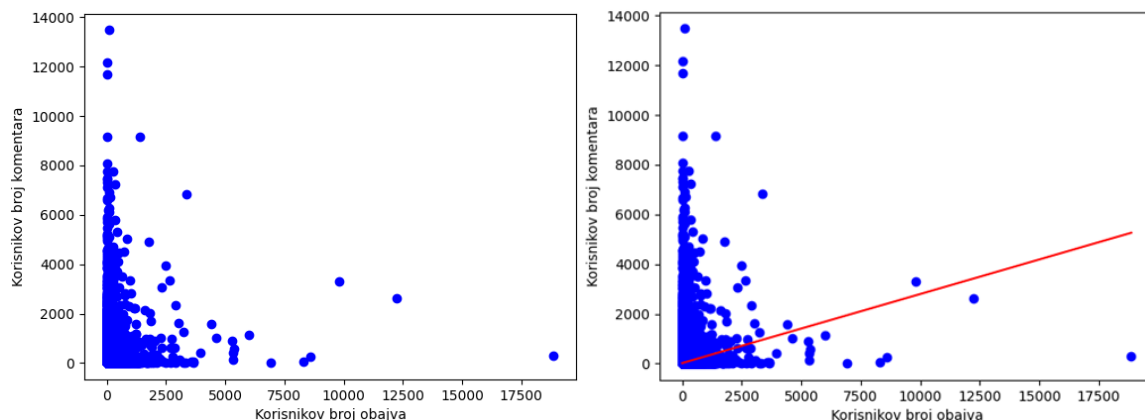
Escafane	154
omfgninja	122
codepoet	111
scientologist2	111
turkourjurs	110
b34nz	107
krugerlive	106
tuoder	103

5. Како су корелисани бројеви објава и коментара? Одредити Пирсонов коефицијент корелације и извршити визуелизацију.

Пирсонов коефицијент корелације је $(r, p) = (0.14608001194152223, 0.0)$, а матрица корелације је дата у наставку:

	Број објава	Број коментара
Број објава	1.00000	0.14608
Број коментара	0.14608	1.00000

Пирсонов коефицијент корелације r је позитиван, те показује да постоји позитивна корелација између бројева корисникових објава и коментара. Вредност коефицијента r је у опсегу $(0, 0.3)$. Стога је корелација поменутих вредности такође занемарљива. На наредној слици лево је приказан за све кориснике број објава и број коментара, одн. на слици десно је уз то приказана и права за коју би функција грешке линеарне регресије била најмања.



6. Које објаве поседују највећи број коментара и на којим су сабрeдитима постављене? У наставку је приказано првих 10 објава са садржајем, уколико садржај није означен да је за старије од 18.

	Submission_id	Subreddit	Num_comments	Over_18
1.	6nz1k	science	33329	false
2.	78n1v	WTF	3657	false
3.	7kpe5	AskReddit	3467	false
4.	675oj	reddit.com	2039	false
5.	7beo2	politics	1934	false
6.	6jbc0	reddit.com	1808	false
7.	6w7e6	reddit.com	1724	false
8.	6yn6t	politics	1668	false
9.	7bj3f	reddit.com,	1605	false
10.	7d7jm	reddit.com	1505	false

1. ONE HUNDRED PUSHUPS

If you're serious about increasing your strength, follow this six week training program and you'll soon be on your way to completing 100 consecutive pushups!

Think there's no way you could do this? I think you can. All you need is a good plan, plenty of discipline and about 30 minutes a week to achieve this goal!

No doubt some of you can already do 50 consecutive pushups, but let's face it, you're in a big minority. Most of you reading this won't even be able to manage 20 pushups. Actually, I'm sure many of you can't even do 10.

However, it really doesn't matter which group you fall into. If you follow the progressive pushups training program, I'm positive you'll soon be able to do 100 pushups!

Are you ready to take the challenge? Read on for more details...

As a symbol of health and wellness, nothing surpasses the simple push-up. The push-up is the ultimate barometer of fitness. It tests the whole body, engaging muscle groups in the arms, chest, abdomen, hips and legs. It requires the body to be taut like a plank with toes and palms on the floor. The act of lifting and lowering one's entire weight is taxing even for the very fit.

2. A black community in OH goes 50 years without running water...until one day, a white family moves in. Now, guess who has the only household on the street with running water?
3. I am constantly wondering what everyone on here looks like. Post your pictures!
4. Post the Funniest Joke You Know... Upmod the Best One
5. Obama wins the Presidency!
6. Voice a true yet unpopular fact in the comments section. Use a shill account if you feel it is

necessary.

7. Just out of curiosity: Do you have an opinion that goes against that prevailing on Reddit? State it here!
8. McCain Shakes Up Race By Picking Sarah Palin for VP

Republican John McCain shook up the presidential race with his surprise choice of little-known Alaska Gov. Sarah Palin as his running mate.

McCain, who turned 72 on Friday, worked to grab the spotlight with his selection of Palin, 44, the first woman to be a Republican vice presidential nominee. His choice was first reported by CNBC. Slideshow: The Many Sides of Sarah Palin

Palin, a self-described "hockey mom," is a conservative first-term governor of Alaska with strong anti-abortion views, a record of reform and fiscal conservatism and an outsider's perspective on Washington.

"She's exactly who I need. She's exactly who this country needs to help me fight the same old Washington politics of me first and country second," McCain told a roaring crowd of 15,000 supporters in Dayton, Ohio.

9. Nation watches as state weighs ban

A measure to once again ban gay marriage in California led Tuesday, throwing into doubt the unions of an estimated 18,000 same-sex couples who wed during the last 4 1/2 months.

As the measure, the most divisive and emotionally fraught on the state ballot this year, took a lead in early returns, supporters gathered at a hotel ballroom in Sacramento and cheered.

"We caused Californians to rethink this issue," Proposition 8 strategist Jeff Flint said.

Early in the campaign, he noted, polls showed the measure trailing by 17 points.

"I think the voters were thinking, well, if it makes them happy, why shouldn't we let gay couples get married. And I think we made them realize that there are broader implications to society and particularly the children when you make that fundamental change that's at the core of how society is organized, which is marriage," he said.

But in San Francisco at the packed headquarters of the No on 8 campaign party in the Westin St. Francis Hotel, supporters of same-sex marriage refused to despair, saying that they were holding out hope for victory.

10. How many redditors don't play WoW?

3. МОДЕЛОВАЊЕ *SNET* МРЕЖЕ

Мрежа је моделована на основу комплетног скупа података, свих сабредита и коментара, при чему нису узети подаци гдје је аутор избрисан. Тежина грана је одређена агрегацијом бројањем.

У првом потпоглављу је извршена основна карактеризација мреже, док је у дугом делу дата анализа мера централности. Треће потпоглавље обрађује тему детекција комуна.

3.1. Основна карактеризација мреже *Snet*

7. Густина мреже је 016518.
8. Мрежа није повезана, стога су просечна дужина и дијаметар мреже бесконачни. Густина гигантске компоненте мреже је 0.025268. Просечна дистанца у гигантској компоненти мреже је 2.099, док је дијаметар 5. Просечна дистанца у Erdos-Reny мрежи истих димензија је 2.077103, а дијаметар је 3. Дијаметар гигантске компоненте је нешто већи од дијаметра Erdos-Reny мреже, па можемо рећи да је дијаметар гигантске компоненте просечан у односу на дијаметар Erdos-Reny мреже.
9. Број повезаних компоненти је 814.

Величина компоненте	Број компоненти
3525	1
3	3
2	16
1	794

Постоји гигантска компонента, јер је језгро величине 3525 чворова и чини највећи дио мреже, тј. 80.85%.

10. Просечни коефицијент кластеризације је 0.7706, док је просечни коефицијент кластеризације узимајући у обзир тежине грана 0.0003. Глобални коефицијент кластеризације је 0.0011369691.

На слици 3.1.1 је приказана расподела чворова по локалном коефицијенту кластеризације узимајући у обзир тежине грана. Она се рачуна по формули

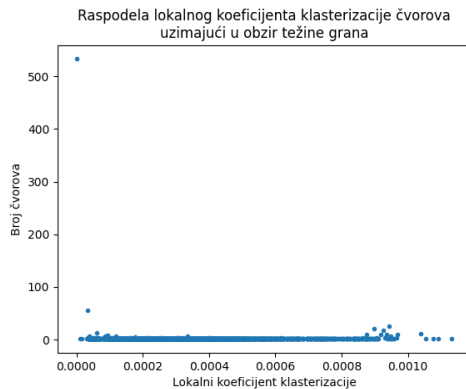
$$c_u = \frac{1}{\deg(u)(\deg(u) - 1)} \sum_{vw} (\hat{w}_{uv} \hat{w}_{uw} \hat{w}_{vw})^{1/3}$$

где је \hat{w}_{uv} нормализовано максималном тежином гране у мрежи тј. $\hat{w}_{uv} = w_{uv}/\max(w)$. За $\deg(u) < 2$ вредност c_u биће једна нули.[1] Са графика се уочава да има преко 500 чворова чији је локални коефицијент кластеризације једнак нули, што из претходног закључка значи да ти чворови имају степен мањи од 2. Пошто њихов степен није једнак нули јер се проучава повезана компонента, закључак је да има преко 500 чворова чији

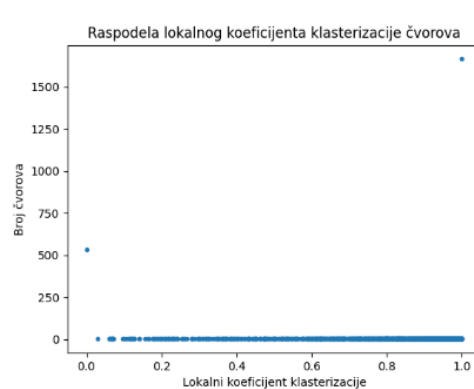
је степен једнак 1. На слици 3.1.2 је приказана расподела чворова по локалном коефицијенту кластеризације не узимајући у обзир тежине грана. Она се рачуна по формули

$$c_u = \frac{2T(u)}{\deg(u)(\deg(u) - 1)}$$

, где је $T(u)$ број троуглова који садрже чвор u , а $\deg(u)$ представља степен чвора. На графику са слике 3.1.2 се може уочити да има нешто више од 500 чворова чији је локални коефицијент кластеризације 0. На основу претходне формуле следи да ти чворови не припадају ниједном троуглу. Са графика такође закључујемо да има преко 1500 чворова са локалним коефицијентом вредности 1. Они припадају кликама, тј. максимално комплетном подграфу.



Слика 3.1.1 Лок.коефицијент узимајући у обзир тежине



Слика 3.1.2 Локални коефицијент не узимајући тежине грана у обзир

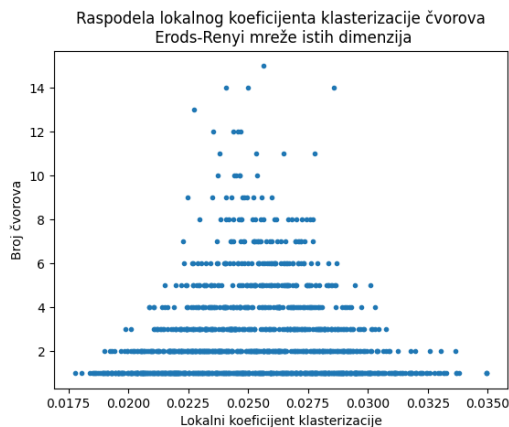
Што се тиче испитивања да ли су чворови са високим локалним коефицијентом кластеризације међусобно повезани, програмским путем је испитан подграф гигантске компоненте који се састоји само од чворова чији је локални степен кластеризације једнак 1, и њихових грана. Закључује се да тај подграф има 1666 чворова и 811 грана, и да је његова густина 0.00058, што значи да ти чворови нису изразито међусобно повезани. Исти поступак је примењен и за релаксирани критеријум, испитан је подграф којег чине чворови чији је локални коефицијент кластеризације већи од 0.5. У том случају добијамо подграф од 2827 чворова, 35502 грана, и густине 0.000004486. Дакле чворови високог локалног коефицијента кластеризације нису изразито повезани међусобно. Такође, када чворове категорички групишемо по локалном коефицијету кластеризације у скупове c0-c1 једнаких величина и прерачунамо матрицу асортативног мешања која је приказана у наредној табели, добијамо да је вероватноћа да су чворови високог локалног коефицијента кластеризације међусобно повезани 0.0059194.

	c0 (0.989247311827957, 1.0)	c1 (0, 0.989010989010989)
c0	0.0059194	0.1253337
c1	0.1253337	0.7434131

Даље следи поређење гигантске компоненте SNet мреже са Erdos-Reny мрежеом истих димензија. Добија се да је просечни коефицијент кластеризације Erdos-Reny мреже 0.025, док је глобални коефицијент кластеризације $1.6e-05$. На наредним сликама је приказана дистрибуција чворова по локалном коефицијенту кластеризације Ердос-Рену мреже. Опсег броја чворова који имају неки локални коефицијент кластеризације иде од 0 до 16, што је и приказано на слици 3.1.3. На слици 3.1.4 је

приказана иста дистрибуција чворова по локалном коефицијенту кластеризације, с тим што се посматра у односу на у осу максималне вредности 1600, због лакшег поређења графика. Поређењем графика и коефицијената кластеризације гигантске компоненте SNet мреже и Erdos-Reny мреже, добијамо да је кластерисање више изражено у гигантској компоненти SNet мреже.

Слика 3.1.3



Слика 3.1.3



Слика 3.1.4

11. Особине малог света су [2]:

1. Мала просечна удаљеност било која два чвора – важи за гигантску компоненту SNet мреже јер је просечна дистанца у мрежи је 2.099, а у Erdos-Reny мрежи димензија као гигантска компонента је 2.077103.
2. Ограничен дијаметар у односу на величину мреже – важи за гигантску компоненту SNet мреже јер је дијаметар 5, а величина мреже 3525, док је у Erdos-Reny мрежи истих димензија као гигантска компонента дијаметар 3.
3. Висок степен кластеризације – према резултатима питања 10) закључујемо да важи за гигантску компоненту SNet мреже.

Из наведеног важи да гигантска компонента SNet мреже испољава особине малог света.

12. Коефицијент асортативности мреже на основу тежинског степена чвора је -0.0323.

Коефицијент асортативности мреже на основу степена чвора је -0.43665. Чворови графа су категорички груписани у скупове c0-c5 једнаких величина. У табели су поред имена скупова приказани и опсежи степена чворова који припадају тој групи.

	c0 (22, 3144)	c1 (1, 22)
c0	0.9353897	0.0306771
c1	0.0306771	0.003256

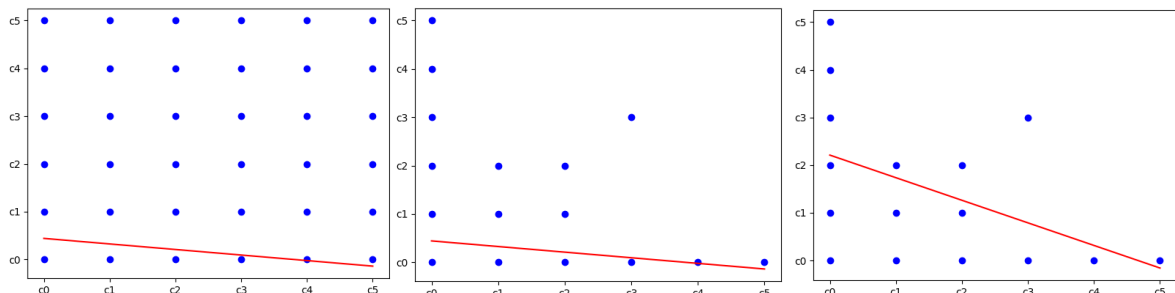
Коефицијент асортативности за категорички подељене чворова према њиховим степенима чвора је 0.0642. За Havel-Nakimi мрежу са истом расподелом чворова по степену као код посматране мреже добијамо коефицијент асортативности -0.4146. Сабирањем елемента главе дијагонале матрице добијамо да је вероватноћа 0.9386 да су чворови истих група међусобно повезани.

	c0(125,3144)	c1(51,125)	c2(22,51)	c3(6,22)	c4(2,6)	c5(1,2)
--	--------------	------------	-----------	----------	---------	---------

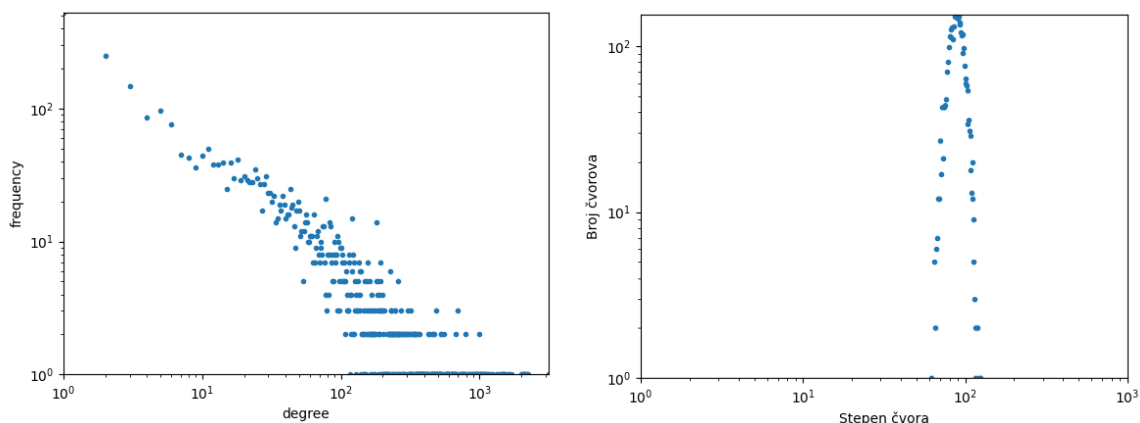
c0	0.5195503	0.1404058	0.0590539	0.0221261	0.0051835	0.0018638
c1	0.1404058	0.0096851	0.0023002	0.0005034	0.0001211	3.5e-05
c2	0.0590539	0.0023002	0.0029119	0.0005416	0.0001211	4.14e-05
c3	0.0221261	0.0005034	0.0005416	0.0018064	0.000274	4.14e-05
c4	0.0051835	0.0001211	0.0001211	0.000274	0.0007264	4.46e-05
c5	0.0018638	3.5e-05	4.14e-05	4.14e-05	4.46e-05	6.4e-06

Коефицијент асортативности за категоријски подељене чворова према њиховим степенима чвора је -0.1297. Сабирањем елемента главе дијагонале матрице добијамо да је вероватноћа 0.5347 да су чворови истих група међусобно повезани.

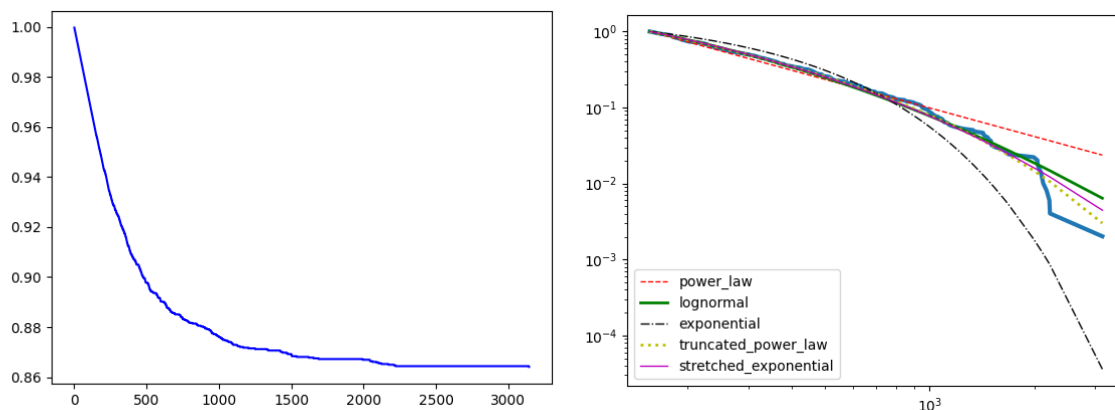
На наредним сликама је прво представљена корелација између поменутих група чворова, затим су визуелно занемарене оне комбинације група код којих је вероватноћа мешања мања од 0.001. На трећој слици је приказана корелација када би се вероватноћа мешања мања од 0.001 заменила са 0, па би се добио и нешто другачији Пирсонов коефицијент одн. коефицијент праве.



13. Проучавањем степена чворова закључујемо да сабрeдит reddit.com има виши степен чвора од осталих чворова (3144). Наредних 10 највећих сабрeдита по степену чвора имају степен у опсегу (1994, 2221) – { technology, politics, pics, funny, science, worldnews, entertainment, WTF, Programming, business}, па је наредни највећи степен 1688. То што сабрeдит reddit.com има висок степен у односу на остале доводи до закључка да су корисници који су активни на reddit.com били активни и на другим сабрeдитима у неком тренутку. Значи нови корисници који долазе а немају налог највероватније ће прво коментарисати и објављивати на главном сабрeдиту reddit.com, или на неком од популарних тема, а потом и на неким од осталих сабрeдита. Стога се може закључити да ће степени чворова који већ имају висок степен расти брже од преосталих. Тиме је закључено да мрежа испољава феномен клуба богатих.
14. На наредним сликама су дате дистрибуције чворова по степену за SNet мрежу, и за Erdos-Reny мрежу истих димензија. Закључак је да се чворови у SNet мрежи не повезују на случајан начин, него да постоје правилности.



На наредној слици приказано је исцртавање комплементарне кумулативне расподеле степена чвора $p(X \geq x)$. Добијамо да функција степена чвора не изгледа као да би могла да се фитује power law функцијом јер на слици не добијамо нешто налик правој. На слици десно је приказана иста функција у $\log\text{-}\log$ скали, и могући модели фитовања.



Пакетом power-law је прерачунао да се најбоље фитује са power law, међутим тек од вредности $x_{\min}=156$.

Alfa = 2.2464827254754156

Xmin = 156.0

Sigma = 0.05602523213855549

Поређење *power law* и *truncated power law* модела

Loglikelihoo ratio: -15.610393791125452

Statistical significance: 2.3028685403048144e-08

Вероватнији је *truncated power law*. Потом је извршено поређење и са осталим моделима, и ту се дошло до истог закључка, а то је да је *truncated power law* изабран за модел.

truncated_power_law exponential: R: 38.20759016516083 p: 2.1843097673389458e-05	truncated_power_law lognormal: R: 2.9684140699707697 p: 0.0014993251280441475	truncated_power_law stretched_exponential: R: 1.6219657292478011 p: 0.009687082968891817
--	--	---

R - Loglikelihoo ratio, *p* - Statistical significance

Атрибути за truncated power law су Alfa= 1.5799608716389861, Xmin=156.0.

- Ауторитети и хабови су карактеристични за усмерене мреже. Стога их нема потребе испитивати над гигантском компонентом SNet мреже јер је она неусмерена.

3.2. Анализа мера централност

16. Посматрали смо гигантску компоненту графа која има 3525 чворова и 156941 грана.

Табела 3.2.1 Централност по степену(лево) и централност по блискости (десно)

Чвор	Сабредит	Вредност	Чвор	Сабредит	Вредност
t5_6	eddit.com	0.892168	t5_6	eddit.com	0.901971
t5_2qh16	technology	0.63025	t5_2qh16	technology	0.72825
t5_2cneq	politics	0.627696	t5_2cneq	politics	0.726149
t5_2qh0u	pics	0.616345	t5_2qh0u	pics	0.720213
t5_2qh33	funny	0.60244	t5_2qh33	funny	0.712495
t5_mouw	science	0.597616	t5_mouw	science	0.710198
t5_2qh13	worldnews	0.59336	t5_2qh13	worldnews	0.708342
t5_2qh0f	entertainment	0.584847	t5_2qh0f	entertainment	0.703955
t5_2qh61	WTF	0.580306	t5_2qh61	WTF	0.701433
t5_2fwo	programming	0.580023	t5_2fwo	programming	0.700875

Табела 3.2Error! No text of specified style in document..2 Релациона централност

Чвор	Сабредит	Вредност
t5_6	reddit.com	0.368649
t5_2qh16	technology	0.049345
t5_2fwo	programming	0.043684
t5_2cneq	politics	0.043102
t5_2qgzg	business	0.041273
t5_2qh0f	entertainment	0.037738
t5_2qh0u	pics	0.037659
t5_2qh13	worldnews	0.033641
t5_2qh33	funny	0.032297
t5_mouw	science	0.031969

17. Централност по својственом вектору

Чвор	Сабредит	Вредност	Расподела суседа чвора по централности по степену	Централност по степену
t5_6	reddit.com	0.374684	0.6x10, 0.5x3, 0.4x12, 0.3x32, 0.2x56, 0.1x333, 0.0x2698	0.892168
t5_2cneq	politics	0.312669	0.9x1, 0.6x9, 0.5x3, 0.4x12, 0.3x32, 0.2x56, 0.1x333, 0.0x1766	0.627696
t5_2qh0u	pics	0.291857	0.9x1, 0.6x9, 0.5x3, 0.4x12, 0.3x32, 0.2x56, 0.1x333, 0.0x1726	0.616345
t5_2qh33	funny	0.279569	0.9x1, 0.6x9, 0.5x3, 0.4x12, 0.3x32, 0.2x56, 0.1x333, 0.0x1677	0.60244
t5_mouw	science	0.27305	0.9x1, 0.6x9, 0.5x3, 0.4x12, 0.3x32, 0.2x56, 0.1x333, 0.0x1660	0.597616
t5_2qh16	technology	0.263952	0.9x1, 0.6x9, 0.5x3, 0.4x12, 0.3x32, 0.2x56, 0.1x333, 0.0x1775	0.63025
t5_2qh13	worldnews	0.258876	0.9x1, 0.6x9, 0.5x3, 0.4x12, 0.3x32, 0.2x56, 0.1x333, 0.0x1645	0.59336
t5_2qh61	WTF	0.257432	0.9x1, 0.6x9, 0.5x3, 0.4x12, 0.3x32, 0.2x56, 0.1x333, 0.0x1599	0.580306
t5_2qh0f	entertainment	0.247071	0.9x1, 0.6x9, 0.5x3, 0.4x12, 0.3x32, 0.2x56, 0.1x333, 0.0x1615	0.584847
t5_2fwo	programming	0.220668	0.9x1, 0.6x9, 0.5x3, 0.4x12, 0.3x32, 0.2x56, 0.1x333, 0.0x1598	0.580023

Посматран је подграф који садржи чворове са највећим вредностима централности по својственом вектору и закључено је да су чворови повезани и гранама тежине 0.5 и тако је подграф комплетан. Посматрањем четврте колоне може се закључити да је сваки сабрeдит повезан са више сабрeдита ниске централности по степену него високе. Из тога се закључује да су сабрeдити са високом вредношћу централности по сопственом вектору моћнији, јер њихови суседи нису повезани са великим бројем суседа тј. претежно имају малу централност по степену.

18. Израчунат је $\alpha_{max} = (5.452e-06+0j)$, и за α је узимана вредност $1/3 \alpha_{max}$ и $2/3 \alpha_{max}$, а за β вредности из скупа $\{1,2,10,100\}$. Комбинација ових параметара у Katz-овој централности приказана је у наредним табелама. Повећањем параметра β за сабрeдит *reddit.com* повећава се и централност тог чвора и постаје значајније већа од централности осталих чворова

Табела 5.2.6 $\alpha=1.799\text{e-}06$, $\beta=10$ за
reddit.com, за остале 1

Чвор	Сабредит	Вредност
t5_6	reddit.com	0.034581
t5_2cneq	politics	0.03145
t5_2qh0u	pics	0.030617
t5_2qh33	funny	0.030035
t5_mouw	science	0.029791

Табела 5.2.10 $\alpha=3.598\text{e-}06$, $\beta=10$ за
reddit.com, за остале 1

Чвор	Сабредит	Вредност
t5_6	reddit.com	0.085093
t5_2cneq	politics	0.073372
t5_2qh0u	pics	0.069853
t5_2qh33	funny	0.067601
t5_mouw	science	0.066528

Табела 5.2.7 $\alpha=1.799\text{e-}06$, $\beta=100$ за
reddit.com, за остале 1

Чвор	Сабредит	Вредност
t5_6	reddit.com	0.051679
t5_2cneq	politics	0.032614
t5_2qh0u	pics	0.0316
t5_2qh33	funny	0.030962
t5_mouw	science	0.030687

Табела 5.2.11 $\alpha=3.598\text{e-}06$, $\beta=100$ за
reddit.com, за остале 1

Чвор	Сабредит	Вредност
t5_6	reddit.com	0.243993
t5_2cneq	politics	0.104128
t5_2qh0u	pics	0.096931
t5_2qh33	funny	0.093295
t5_mouw	science	0.091464

Табела 5.2.6 $\alpha=1.799\text{e-}06$, $\beta=10$ за
reddit.com, за остале 1

Чвор	Сабредит	Вредност
t5_6	reddit.com	0.185062
t5_2cneq	politics	0.041387
t5_2qh0u	pics	0.038965
t5_2qh33	funny	0.037899
t5_mouw	science	0.037385

Табела 5.2.10 $\alpha=3.598\text{e-}06$, $\beta=10$ за
reddit.com, за остале 1

Чвор	Сабредит	Вредност
t5_6	reddit.com	0.761279
t5_2cneq	politics	0.183565
t5_2qh0u	pics	0.164957
t5_2qh33	funny	0.157368
t5_mouw	science	0.153307

Табела 5.2.7 $\alpha=1.837\text{e-}06$, $\beta=100$ за
reddit.com, за остале 1

Чвор	Сабредит	Вредност
t5_6	reddit.com	0.847028
t5_2cneq	politics	0.073535
t5_2qh0u	pics	0.064308
t5_2qh33	funny	0.061301
t5_mouw	science	0.059672

Табела 5.2.11 $\alpha=3.598\text{e-}06$, $\beta=100$ за
reddit.com, за остале 1

Чвор	Сабредит	Вредност
t5_6	reddit.com	0.825465
t5_2cneq	politics	0.198943
t5_2qh0u	pics	0.178438
t5_2qh33	funny	0.170167
t5_mouw	science	0.165754

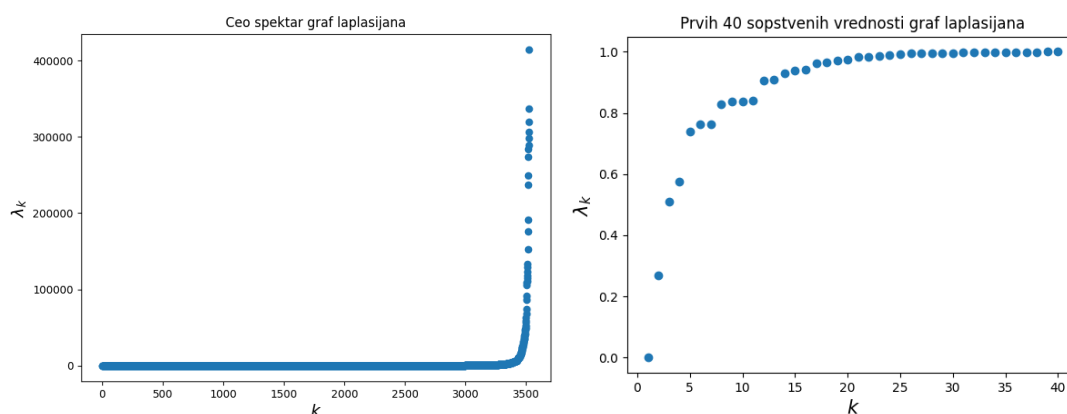
19. 10 чворова гигантске компоненте SNet мреже са највећом вредношћу композитног ранга

Чвор	Сабредит	DC ранг	CC ранг	BC ранг	EVC ранг	Композитни ранг
t5_6	reddit.com	1.0	1.0	1.0	1.0	1.0
t5_2qh16	technology	2.0	2.0	2.0	6.0	48.0
t5_2cneq	politics	3.0	3.0	4.0	2.0	72.0
t5_2qh0u	pics	4.0	4.0	7.0	3.0	336.0
t5_2qh33	funny	5.0	5.0	9.0	4.0	900.0
t5_mouw	science	6.0	6.0	10.0	5.0	1800.0
t5_2qh13	worldnews	7.0	7.0	8.0	7.0	2744.0

t5_2fwo	programming	10.0	10.0	3.0	10.0	3000.0
t5_2qh0f	entertainment	8.0	8.0	6.0	9.0	3456.0
t5_2qgzg	business	11.0	11.0	5.0	5.0	6655.0

3.3. Детекција комуна мреже *Snet*

20. Цео спектар граф лапласијана и првих 40 сопствених вредности дати су на наредним сликама. Са графика првих 40 сопствених вредности граф лапласијана видимо да скокови постоје за вредност k из скупа $\{2,3,4,5,8,12\}$. Међутим за вредности $\{4,8,12\}$ ће се добити нешто слабија подела, јер је скок на тим вредностима нешто мањи од других.



Надаље су приказане колике би биле величине компоненти када би поделе биле на поменути број компоненти.

Подела на 2: величине компонената су [3522, 3]

Подела на 3: величине компонената су [3517, 3, 5]

Подела на 4: величине компонената су [3463, 3, 5, 54]

Подела на 5: величине компонената су [3456, 3, 7, 5, 54]

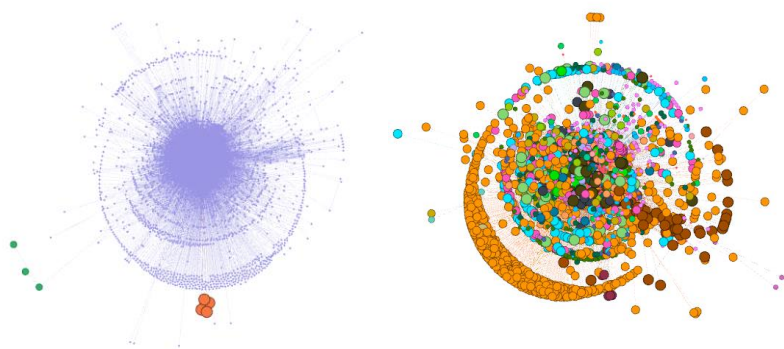
Подела на 8: величине компонената су [3439, 4, 3, 7, 7, 5, 6, 54]

Подела на 12: величине компонената су [3413, 4, 3, 3, 7, 7, 5, 6, 4, 16, 54, 3]

21. За параметар резолуције вредности 5 добија се једна комуна, што је тривијалан случај. Ако се за пареметар резолуције узме вредност 2, добија се 12 комуна величина [52, 5, 4, 12, 7, 3, 4, 3, 3, 4, 3425, 3]. Визуелизација је дата наредном сликом. Спектрална анализа са 12 комуна дала је нешто другачију поделу. У оба случаја добила се модуларност око 0.0003, са разликом на петој децимали.



За параметар резолуције вредности 3, добијају се 3 комуна величина [5, 3, 3517] и визуелизација је дата на наредној слици лево. А на десној слици је приказана визуелизација мреже уколико се као параметар резолуције узме вредност 0.4. Тада се добија 32 комуне величина [119, 5, 64, 54, 9, 63, 68, 7, 106, 32, 728, 142, 89, 31, 54, 57, 30, 3, 84, 27, 170, 66, 29, 73, 241, 291, 42, 50, 30, 688, 7, 66].



22. У наредној табели су приказана поређења коришћених метода за поделу мреже на комуне. Добија се да је модуларност за сваку мала, али да је ипак највећа за спектралну анализу која је поделила мрежу на 12 комуна, па је тај случај и разматран.

Метода	Бр. комуна	Величина комуна	Модуларност
Асинхрона пропација лабела	14	[3454, 30, 7, 6x2, 3x4, 2x5]	0.000160138773
Синхрона пропација лабела	20	[3476, 6, 5, 3x4, 2x13]	2.98609270e-05
Спектрална анализа	2	[3522, 3]	1.08587386e-06
Спектрална анализа	3	[3517, 3, 5]	6.51522011e-06
Спектрална анализа	4	[3463, 3, 5, 54]	0.000258916716
Спектрална анализа	5	[3456, 3, 7, 5, 54]	0.000290398320
Спектрална анализа	8	[3439, 4, 3, 7, 7, 5, 6, 54]	0.000305052338
Спектрална анализа	12	[3413, 4x2, 3x3, 7x2, 5, 6, 16, 54]	0.000340327985
Лувенска метода	3	[5, 3, 3517]	6.51522011e-06
Лувенска метода	12	[3425, 52, 12, 7, 5, 4x3, 3x4]	0.000322968793
Лувенска метода	32	[728.. 3]	-0.00202238635

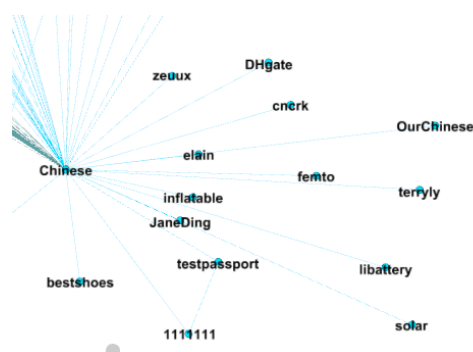
Највећа комуна од 3425 чворова није приказана због обимности. Она представља централни део мреже, и у њој се налази сви чворови мреже. У наставку су приказани сабредити који припадају комунама, и претпоставке зашто су тако гурписани.

- [hotel, supermarkets, travelers] – Путници се често распитују за хотеле и маркете.
- [Concerts, Event_Launch, Las_Vegas, Launch_Strategies, SportsEvent, Theaters]– корисници заинтересовани за друштвена дешавања посећују ове сабредите да би били у току или ишли на њих
- [Peiman, bestactorofiran, PersianDigg, Sobhaneh] - *Payman Maadi* је иранско-амерички глумац, те су прва два сабредита вероватно посвећена њему. Друга два сабредита представљају иранску храну. Ови сабредити заједно окупљају кориснике које привлачи оно што је карактеристично за Иран.
- [Cetatenie, dezvoltarepersonala, positivethinker] – Назив првог сабредита на румунском значи држављанство, а други назив значи лични развој. Дакле ови сабредити окупљају румунске кориснике које интересује позитивно размишљање, лични развој, напредак и вероватно начин да добију држављанство.
- [bestoutdoortowns, bestbarsboulder, bestentrepbooks]– *Boulder* је место у Колораду које је на планини. Назив првог сабредита представља тему најбољих градова који су на надморској висини, међу њима је и поменуто место. Други сабредит представља најбоље барове у том месту, а трећи најбоље књиге за предузетништво. Дакле, корисници који су заинтересовани за боравак у природи, планинама су вероватно заинтересовани за предузетништво, или има корисника који би желили да буду предузетници и отворе свој локал на неком месту тог типа.
- [Hepsi, Marketallica, bloggerdal, fx15, kultur_sanat, lida, tala] – сабредити који представљају ентитете везане за Турску. *Hepsi* је био турски женски бенд, *Marketallica* је консултантска фирма за трендове настала у Истанбулу, *bloggerdal* је турски блог за музику. *Kultur sanat* је компанија која организује музичке и комерцијалне догађаје широм Турске, *Lida* је компанија за трговину драгим камењем, а *Tala* компанија за некретнине у Турској.
- [algeria, 7moode, arabic, arabicblogs, d1g, testa, wazayf] – *Algeria* је Алжир, где је службени језик арапски. На сабредиту 7moode има доста објава и коментара на арапском. *D1g* је веб портал који омогућава Арапима да граде заједнице и мреже. *Wazayf* на арапском значи послови.
- [1111111, 3ren, 5edwin, AlertsCar, AlvinHe, BilaBila, Chinese, DHgate, FunnyOlympic, JACKgg, JaneDing, Josie, Kaede_Rukawa, Liang, OurChinese, PGD, Rachel61, SinoBeta, XYON, Zharster, adown, bestshoes, chentao807, cimigo, cnbeta, cnck, colorreddit, dior_lei, elain, femto, inflatable, lamengao, libattery, llin, lullaby, mygame, ourex, reddit7752, sdfsd, shoe, simon, slashdotxx, solar, startfs, terryly, testpassport, testtest, tiantiankankan, tosoft, try, wagi, wenyude, yuen999, zeuux] –

доста појмова везаних за Кину: Chinese, OurChinese, Liang – кинеска династија, FunnyOlympic – Олимпијада која се 2008 одржала у Кини, SinoBeta, cnbeta – кинеске компаније, JaneDing – доста Кинескиња тог имена на LinkedIn-у, Kaede_Rukawa – аниме.

- [AnnikRubens, BBMeals, ClothesBracelets, Danielsbestsites, Fotobookmarks, Jeffrey, Mister, Neoogilvy, Preisvergleich, Programminginterests, SensitiveHeadlines, Woerterbuecher, basis, bestsitesever, sevenload, yoomoo] – сабредити имају везе са Немачком, рецимо *Larisu Vasilian* је полу-Немица која има подкаст где наступа под именом *AnnikRubens*, *Sevenload* је био сајт у Немчкој за дељење видео записа, Preisvergleich, Woerterbuecher – речи на немачком.

23. Брокери су они чворови у мрежи који имају релативно ниска мрежа ограничења, а високу релациону централност. За проналажење брокера коришћења је функција bridges из библиотеке networkx.algorithms.bridge, која проналази ламбда скуп тј. оне гране чијим би се уклањањем мрежа поделила на више неповезаних компоненти. Резултат је дат као низ парова чворова који представљају поменути ламбда скуп грана и има их 531. Види се да се сабредит *reddit.com* појављује у највећем броју грана (357), и он има највећу релациону централност, тако да он дефинитивно представља мост у мрежи. Сабредит *Chinese* се налази у 14 грана, и визуелно се види да доста чворова зависе само од њега, тако да и он представља мост мреже.

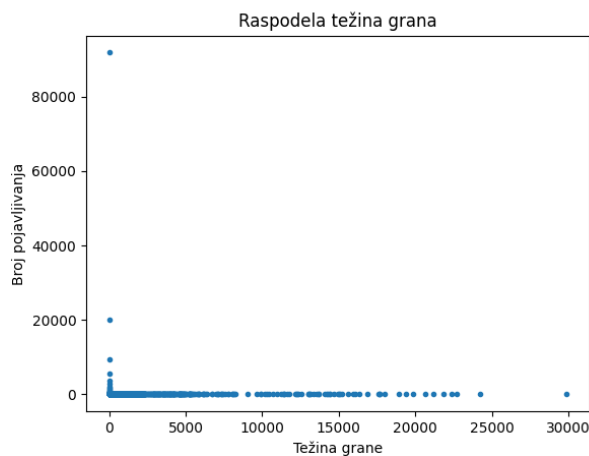


У наредној табели су дати сабредити који се налазе у највећем броју грана за уклањање. Може се уочити да чворови постоје и међу највишим релационим централностима. У мостове ће се сврстати и чворови Business, Programming, Technology и Politics, јер бар 8 чворова зависе директно од њих, а и они имају највишу релациону централност после чвора reddit.com.

Сабредит	Број грана
Business	21
Programming	15
Technology	9
Politics	9
Entertainment	8
Pics	7
WorldsNews	7
Science	6
WTF	5
Nsfw	4
Funny	4
Economics	3
Gossip	3

4. МОДЕЛОВАЊЕ SNETF МРЕЖЕ

Мрежа SnetF је генерисана као подграф мреже Snet, филтрирањем грана мале тежине. За $w_threshold$ је узета вредност 5, и одбачене су све гране чија је тежина испод те вредности. Вредност $w_threshold$ је одабрана тако да гране које преостану чине до 19.26% укупног броја грана Snet мреже. Кад се од Snet мреже која има 156941 грана одбаце гране тежина 4 којих има 5590, гране тежина 3 којих има 9292, гране тежина 2 којих има 20000, и тежина 1 којих има 91856, добијамо мрежу од 768 чворова и 30288 грана ($30288/156941=0.1926$). Тежине грана су приказане на следећем графику.



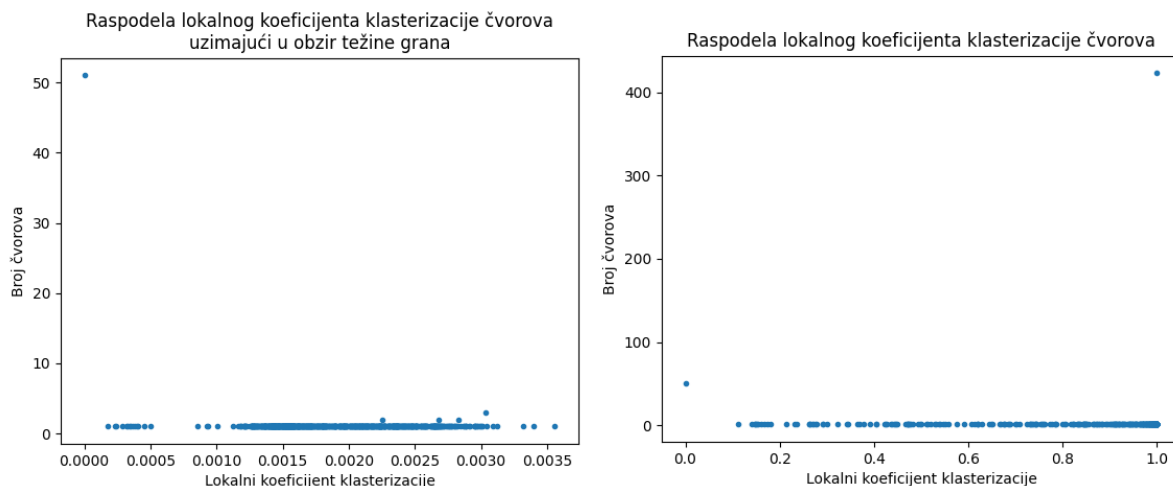
У првом потпоглављу је извршена основна карактеризација мреже, док је у дугом делу дата анализа мера централности. Треће потпоглавље обрађује тему детекција комуна.

4.1. Основна карактеризација SnetF мреже

7. Густина SNetF мреже је 0.102632.
8. Мрежа није повезана, стога су просечна дужина и дијаметар мреже бесконачни. Густина гигантске компоненте мреже је 0.103166. Просечна дистанца у гигантској компоненти мреже је 1.973, док је дијаметар 5. Просечна дистанца у Erdos-Reny мрежи истих димензија је 1.897285, а дијаметар је 3. Можемо рећи да је дијаметар гигантске компоненте SNetF мреже просечан у односу на дијаметар Erdos-Renymреже.
9. Постоје две повезане компоненте у SNetF мрежи. Величина компоненти је дата као низ где први члан представља величину компоненте, а други број компоненти које су те величине: 766x1, 2x1. Постоји гигантска компонента мреже јер постоји језгро величине 766 чворова и представља највећи део мреже тј. 99.74% мреже.
10. Просечни коефицијент кластеризације SNetF мреже је 0.852, док је просечни коефицијент кластеризације узимајући у обзир тежине грана 0.0018. Глобални коефицијент кластеризације је 0.0149167351.

На наредној слици лево је приказана дистрибуција чворова по локалном коефицијенту кластеризације узимајући у обзир тежине грана. Са графика се уочава да има преко 50 чворова чији је локални коефицијент кластеризације једнак нули, што значи да ти чворови имају степен мањи од 2. Пошто њихов степен није једнак нули јер се проучава

повезана компонента, закључак је да има преко 50 чворова чији је степен једнак 1. На следећој слици десно је приказана дистрибуција чворова по локланом коефицијенту кластеризације не узимајући у обзир тежине грана.



Проучавањем десног графика уочавамо да има нешто више од 50 чворова чији је локални коефицијент кластеризације 0. Закључујемо да су чворови који имају локални коефицијент кластеризације једнак 0 заправо они који не припадају ни једном троуглу. Са графика такође закључујемо да има преко 400 чворова чији је локални коефицијент кластеризације једнак 1. То су чворови који припадају кликама, тј. припадају максималном комплетном подграфу (он и његови суседи представљају максималан комплетан подграф).

Што се тиче испитивања да ли су чворови са високим локалним коефицијентом кластеризације међусобно повезани, програмским путем је испитиван подграф гигантске компоненте који се састоји само од чворова чији је локални степен кластеризације једнак 1, и њихових грана. Закључује се да тај подграф има 423 чвора и једну грану, што значи да ти чворови нису изразито међусобно повезани. Исти поступак је примењен и за релаксирани критеријум да се испита подграф којег чине чворови чији је локални коефицијент кластеризације већи од 0.5. У том случају добијамо подграф од 662 чворова, 9859 грана, и густине 0.04506. Дакле чворови високог локалног коефицијента кластеризације нису изразито повезани међусобно. Такође, када чворове категорички групишемо по локалном коефицијенту кластеризације у скупове c_0 - c_1 једнаких величина и прерачунамо матрицу асортативног мешања која је приказана у наредној табели, добијамо да је вероватноћа да су чворови високог локалног коефицијента кластеризације међусобно повезани 0.

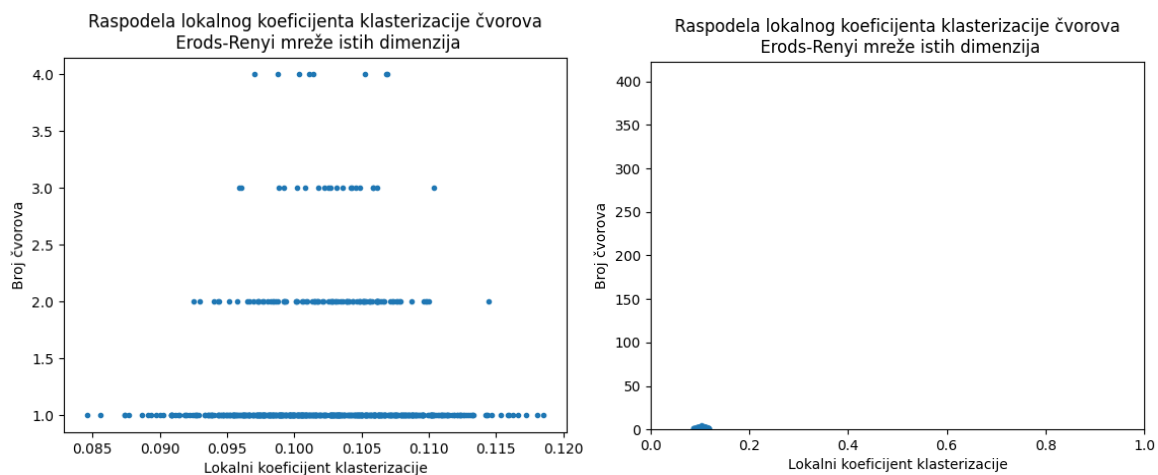
	c_0 (1.0, 1.0)	c_1 (0, 1.0)
c_0	0.0	0.14229
c_1	0.14229	0.71542

Такође, када чворове категорички групишемо по локалном коефицијенту кластеризације у скупове c_0 - c_8 једнаких величина (пошто чворова са локалним коефицијентом од 1 има 423, а укупно чворова 766, ако поделимо на 9 класа свака класа ће бити величине 85 чворова, па ће првих 5 класа одговарати чворовима са

локалним коефицијентом од 1) и прерачунамо матрицу асортативног мешања која је приказана у наредној табели, добијамо да је вероватноћа да су чворови високог локалног коефицијента кластеризације међусобно повезани 0.

	c0 (1.0, 1.0)	c1 (1.0, 1.0)	c2 (1.0, 1.0)	c3 (1.0, 1.0)	c4 (0.9996297667530544, 1.0)
c0	0.0	0.0	0.0	0.0	0.0
c1	0.0	0.0	0.0	0.0	0.0
c2	0.0	0.0	0.0	0.0	0.0
c3	0.0	0.0	0.0	0.0	0.0
c4	0.0	0.0	0.0	0.0	3.31e-05

Добија се да је просечни коефицијент кластеризације Erdos-Reny мреже 0.102, док је глобални коефицијент кластеризације 0.001073. На наредним сликама је приказана расподела локалног коефицијента кластеризације Erdos-Reny мреже. Опсег броја чворова који имају дати локални коефицијент кластеризације иде од 0 до 4, што је и приказано на левој слици. На десној слици је приказана иста дистрибуција чворова по локалном коефицијенту кластеризације, с тим што се посматра у односу на у осу максималне вредности 400, због лакшег поређења тог графика и графика расподеле локалног коефицијента кластеризације чворова гигантске компоненте SnetF мреже.



Поређењем графика и коефицијената кластеризације гигантске компоненте SNetF мреже и Erdos-Reny мреже, добијамо да је кластерисање више изражено у гигантској компоненти SNetF мреже.

11. Феномен малог света је карактеристика реалних мрежа за које важи:

1. Мала просечна удаљеност било која два чвора – важи за гигантску компоненту SNetF мреже јер је просечна дистанца у мрежи је 1.973, а у Erdos-Reny мрежи димензија као гигантска компонента је 1.897285.
2. Ограничен дијаметар у односу на величину мреже – важи за гигантску компоненту SNetF мреже јер је дијаметар 5, а величина мреже 766, док је у Erdos-Reny мрежи истих димензија као гигантска компонента дијаметар 3.
3. Висок степен кластеризације – према резултатима питања 10) закључујемо да важи за гигантску компоненту SNetF мреже.

Из наведеног важи да гигантска компонента SNetF мреже испољава особине малог света.

12. Коефицијент асортативности SNetF мреже на основу тежинског степена чвора је -0.0047.

Коефицијент асортативности SNetF мреже на основу степена чвора је - 0.5945074611306468.

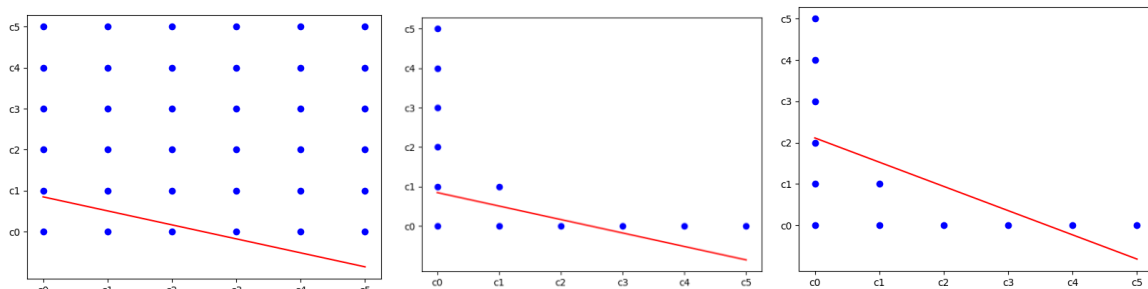
Чворови графа су категорички груписани у скупове c0-c5 једнаких величина. Изабрано је 6 категоричких гурпа, да би чворови тежинског степена 1-5 били сврстани у једну групу, јер њих има 129 и представљају шестину мреже. Прерачуната је матрица асортативног мешања која је

приказана у наредној табели. У табели су поред имена скупова приказани и опсези степене чворова који припадају тој групи.

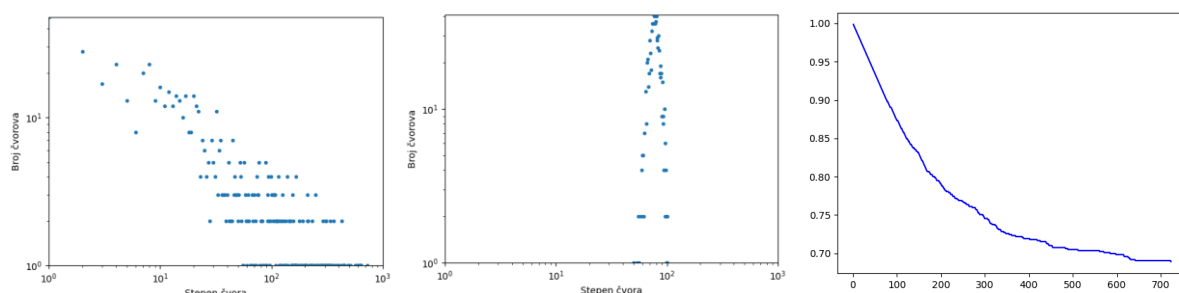
	c0 (151, 724)	c1 (67, 150)	c2 (29, 67)	c3 (14, 29)	c4 (5, 14)	c5 (1, 5)
c0	0.2705528	0.1978033	0.0917061	0.0407748	0.019519	0.0039534
c1	0.1978033	0.0169716	0.0009925	0.0003639	6.62e-05	3.31e-05
c2	0.0917061	0.0009925	9.92e-05	4.96e-05	0.0	8.27e-05
c3	0.0407748	0.0003639	4.96e-05	9.92e-05	0.0001323	0.0003143
c4	0.019519	6.62e-05	0.0	0.0001323	6.62e-05	0.000182
c5	0.0039534	3.31e-05	8.27e-05	0.0003143	0.000182	0.0002647

Коефицијент асортативности за категорички подељене чворова према њиховим степенима чвора је -0.2882. Сабирањем елемента главе дијагонале матрице добијамо да је вероватноћа 0.2881 да су чворови истих група међусобно повезани. За Хавел-Хакими мрежу са истом расподелом чворова по степену као код посматране мреже добијамо коефицијент асортативности -0.583.

На наредним сликама је прво представљена корелација између поменутих група чворова, затим су визуелно занемарене оне комбинације група код којих је вероватноћа мешања мања од 0.001. На трећој слици је приказана корелација када би се вероватноћа мешања мања од 0.001 заменила са 0, па би се добио и нешто другачији Пирсонов коефицијент односно коефицијент праве.



13. Проучавањем степена чворова закључујемо да сабрeдит reddit.com има виши степен чвора од осталих чворова (724). Нарeдних 10 највeћих сабрeдита по степену чвора имају степен у опсегу (562, 643) – {politics, pics, science, technology, worldnews, funny, WTF, entertainment, programming, business}, па је наредни највeћи степен 513. То што сабрeдит reddit.com има висок степен у односу на остале доводи до закључка да су корисници који су активни на reddit.com били активни и на другим сабрeдитима у неком тренутку. Значи нови корисници који долазе а немају налог највероватније ће прво коментарисати и објављивати на главном сабрeдиту reddit.com, или на неком од популарних сабрeдита, а потом и на неким од осталих сабрeдита. Стога се може закључити да ће степени чворова који већ имају висок степен расти брже од преосталих. Тиме закључујемо да мрежа испољава феномен клуба богатих.
14. На наредним сликама су дате дистрибуције чворова по степену за SNetF мрежу, и за Erdos-Reny мрежу истих димензија, и комплементарна кумулативна расподела степена чвора, која није налик правој. Стога не делује као да би могла да се апроксимирала power law расподелом.



Power law пакет је прерачунао да се најбоље фитује са power law међутим тек од вредности $x_{\min}=272$, $\alpha=3.869347965$, $\sigma=0.36738236$.

Поређењем power law модела редом са *truncated power law*, *exponential*, *lognormal*, *stretched_exponential* добијамо негативне вредности за сличност (око -2), а статистички значај се не добија мање од 0.05. Испитаће се и поређење *truncated power law* са осталим моделима, што је приказано следећом табелом (R - Loglikelihood ratio, p - Statistical significance).

truncated_power_law exponential: R: -0.20261360499475956 p: 0.5294703798029627	truncated_power_law lognormal: R: 0.11945372408456123 p: 0.8014912660959591	truncated_power_law stretched_exponential: R: -0.30049758097268064 p: 0.6517487420139914
---	--	---

R је код *exponential*, *stretched_exponential* модела благо негативан, а p није мањи од 0.05, тако да се не може са сигурношћу тврдити да је неки од та два бољи од *truncated power law*.

Поређење *exponential*, *stretched_exponential* модела:

Loglikelihood ratio: 0.09788397597792109

Statistical significance: 0.6581587400780612

Добија се благо позитивна предност за *stretched_exponential*, мада се не може тврдити због малог статистичког значаја. Прерачунавањем параметра за *stretched_exponential* опет се добија да $x_{\min}=272$, $\beta=1.23499$.

- Ауторитети и хабови су карактеристични за усмерене мреже. Стога их нема потребе испитивати над гигантском компонентом SNetF мреже јер је она неусмерена.

4.2. Мере централности

- У следећим табелама су приказани чворови са највећим вредностима централности по степену, блискости и релационе централности, респективно.

Чвор	Сабредит	Вредност
t5_6	reddit.com	0.946405
t5_2cneq	politics	0.840523
t5_2qh0u	pics	0.826144
t5_mouw	science	0.823529

Чвор	Сабредит	Вредност
t5_6	reddit.com	0.947955
t5_2cneq	politics	0.848115
t5_2qh0u	pics	0.846239
t5_2qh16	technology	0.842511

t5_2qh16	technology	0.819608
t5_2qh13	worldnews	0.807843
t5_2qh33	funny	0.806536
t5_2qh61	WTF	0.783007
t5_2qh0f	entertainment	0.766013
t5_2fwo	programming	0.747712

t5_mouw	science	0.83698
t5_2qh13	worldnews	0.834242
t5_2qh33	funny	0.832427
t5_2qh61	WTF	0.806962
t5_2qh0f	entertainment	0.806112
t5_2fwo	programming	0.794393

Чвор	Сабредит	Вредност
t5_6	reddit.com	0.191027
t5_2cneq	politics	0.060944
t5_2qh0u	pics	0.054255
t5_2cneq	politics	0.051584
t5_2qh16	technology	0.051336
t5_mouw	science	0.049068
t5_2fwo	programming	0.048668
t5_2qh33	funny	0.047433
t5_2qh61	WTF	0.041519
t5_2qh0f	entertainment	0.036219

17. Централност по својственом вектору

Чвор	Сабредит	Вредност	Расподела суседа чвора по централности по степену	Централност по степену
t5_6	reddit.com	0.374686	0.8x8, 0.7x3, 0.6x10, 0.5x9, 0.4x30, 0.3x33, 0.2x72, 0.1x168, 0.0x391	0.946405
t5_2cneq	politics	0.312671	0.9x1, 0.8x7, 0.7x3, 0.6x10, 0.5x9, 0.4x30, 0.3x33, 0.2x72, 0.1x168, 0.0x310	0.840523
t5_2qh0u	pics	0.291858	0.9x1, 0.8x7, 0.7x3, 0.6x10, 0.5x9, 0.4x30, 0.3x33, 0.2x72, 0.1x168, 0.0x299	0.826144
t5_2qh33	funny	0.279571	0.9x1, 0.8x7, 0.7x3, 0.6x10, 0.5x9, 0.4x30, 0.3x33, 0.2x72, 0.1x168, 0.0x284	0.806536

t5_mouw	science	0.273051	0.9x1, 0.8x7, 0.7x3, 0.6x10, 0.5x9, 0.4x30, 0.3x33, 0.2x72, 0.1x168, 0.0x297	0.823529
t5_2qh16	technology	0.263953	0.9x1, 0.8x7, 0.7x3, 0.6x10, 0.5x9, 0.4x30, 0.3x33, 0.2x72, 0.1x168, 0.0x294	0.819608
t5_2qh13	worldnews	0.258878	0.9x1, 0.8x7, 0.7x3, 0.6x10, 0.5x9, 0.4x30, 0.3x33, 0.2x72, 0.1x168, 0.0x285	0.807843
t5_2qh61	WTF	0.257433	0.9x1, 0.8x7, 0.7x3, 0.6x10, 0.5x9, 0.4x30, 0.3x33, 0.2x72, 0.1x168, 0.0x266	0.783007
t5_2qh0f	entertainment	0.247073	0.9x1, 0.8x7, 0.7x3, 0.6x10, 0.5x9, 0.4x30, 0.3x33, 0.2x72, 0.1x168, 0.0x253	0.766013
t5_2fwo	programming	0.220669	0.9x1, 0.8x8, 0.7x2, 0.6x10, 0.5x9, 0.4x30, 0.3x33, 0.2x72, 0.1x168, 0.0x239	0.747712

Посматрањем четврте колоне може се закључити да је сваки сабрeдит повезан са више сабрeдита ниске централности по степену него високе. Из тога се закључује да су сабрeдити са високом вредношћу централности по вектору моћнији, јер њихови суседи нису повезани са великим бројем суседа тј. претежно имају малу централност по степену.

18. Израчунат је $\alpha_{max} = (5.452e-06+0j)$, и за α је узимана вредност $1/3 \alpha_{max}$ и $2/3 \alpha_{max}$, а за β вредности из скупа $\{1,2,10,100\}$. Комбинација ових параметара у Katz-овој централности приказана је у наредним табелама. Повећањем параметра β за сабрeдит *reddit.com* повећава се и централност тог чвора и постаје значајније већа од централности осталих чворова.

Табела 5.2.6 $\alpha=1.799\text{e-}06$, $\beta=10$ за reddit.com, за остале 1

Чвор	Сабредит	Вредност
t5_6	reddit.com	0.072241
t5_2cneq	politics	0.065751
t5_2qh0u	pics	0.064011
t5_2qh33	funny	0.062799
t5_mouw	science	0.06229

Табела 5.2.10 $\alpha=3.598\text{e-}06$, $\beta=10$ за reddit.com, за остале 1

Чвор	Сабредит	Вредност
t5_6	reddit.com	0.161934
t5_2cneq	politics	0.139707
t5_2qh0u	pics	0.133009
t5_2qh33	funny	0.128728
t5_mouw	science	0.126689

19.

Табела 5.2.7 $\alpha=1.799\text{e-}06$, $\beta=100$ за reddit.com, за остале 1

Чвор	Сабредит	Вредност
t5_6	reddit.com	0.107702
t5_2cneq	politics	0.067925
t5_2qh0u	pics	0.065813
t5_2qh33	funny	0.064489
t5_mouw	science	0.063918

Табела 5.2.11 $\alpha=3.598\text{e-}06$, $\beta=100$ за reddit.com, за остале 1

Чвор	Сабредит	Вредност
t5_6	reddit.com	0.195466
t5_2cneq	politics	0.14502
t5_2qh0u	pics	0.137578
t5_2qh33	funny	0.133036
t5_mouw	science	0.130851

Табела 5.2.6 $\alpha=1.799\text{e-}06$, $\beta=10$ за reddit.com, за остале 1

Чвор	Сабредит	Вредност
t5_6	reddit.com	0.363432
t5_2cneq	politics	0.08111
t5_2qh0u	pics	0.076357
t5_2qh33	funny	0.07427
t5_mouw	science	0.073264

Табела 5.2.10 $\alpha=3.598\text{e-}06$, $\beta=10$ за reddit.com, за остале 1

Чвор	Сабредит	Вредност
t5_6	reddit.com	0.410062
t5_2cneq	politics	0.174609
t5_2qh0u	pics	0.162525
t5_2qh33	funny	0.156431
t5_mouw	science	0.15336

Табела 5.2.7 $\alpha=1.837\text{e-}06$, $\beta=100$ за reddit.com, за остале 1

Чвор	Сабредит	Вредност
t5_6	reddit.com	0.938119
t5_2cneq	politics	0.081384
t5_2qh0u	pics	0.071166
t5_2qh33	funny	0.067837
t5_mouw	science	0.066035

Табела 5.2.11 $\alpha=3.598\text{e-}06$, $\beta=100$ за reddit.com, за остале 1

Чвор	Сабредит	Вредност
t5_6	reddit.com	0.805401
t5_2cneq	politics	0.194056
t5_2qh0u	pics	0.174374
t5_2qh33	funny	0.16635
t5_mouw	science	0.162057

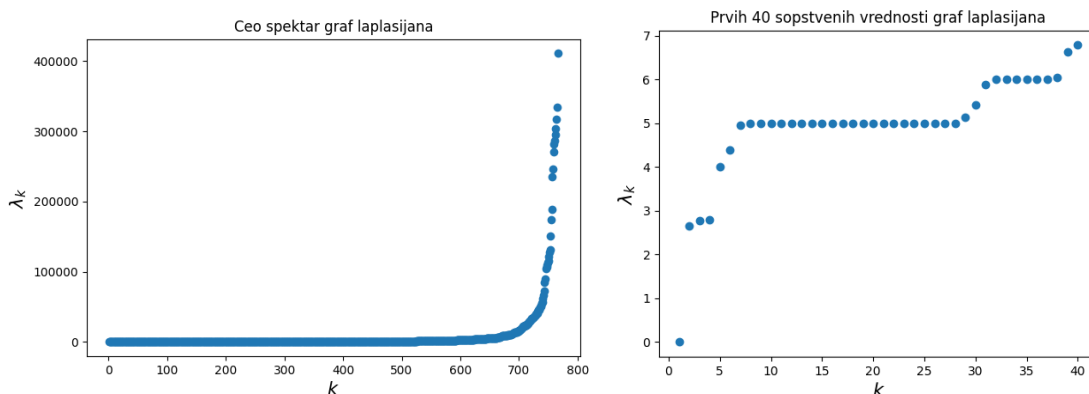
19. 10 чворова гигантске компоненте SNetF мреже са највећом вредношћу композитног ранга

Чвор	Сабредит	DC ранг	CC ранг	BC ранг	EVC ранг	Композитни ранг
t5_6	reddit.com	1.0	1.0	1.0	1.0	1.0
t5_2cneq	politics	2.0	2.0	2.0	2.0	16.0
t5_2qh0u	pics	3.0	3.0	3.0	3.0	81.0
t5_2qh16	technology	5.0	4.0	4.0	6.0	480.0
t5_mouw	science	4.0	5.0	5.0	5.0	500.0
t5_2qh33	funny	7.0	7.0	8.0	4.0	1568.0

t5_2qh13	worldnews	6.0	6.0	7.0	7.0	1764.0
t5_2qh61	WTF	8.0	8.0	9.0	8.0	4608.0
t5_2qh0f	entertainment	10.0	10.0	6.0	10.0	6000.0
t5_2fwo	programming	9.0	9.0	10.0	9.0	7290.0

4.3. Детекција комуна мреже *SnetF*

20. Цео спектар граф лапласијана и првих 40 сопствених вредности дати су на наредним сликама. Са графика првих 40 сопствених вредности граф лапласијана видмо да скокови постоје за вредност k из скупа $\{2,5,6,7,31,39\}$.



Подела на 2: величине компонената су [763, 3]

Подела на 5: величине компонената су [747, 2, 2, 3, 12]

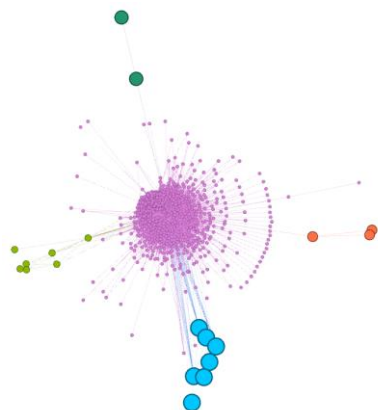
Подела на 6: величине компонената су [740, 2, 2, 7, 3, 12]

Подела на 7: величине компонената су [737, 2, 2, 3, 7, 3, 12]

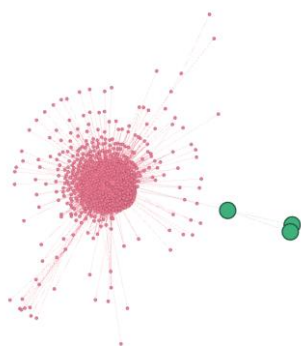
Подела на 31: величине компонената су [63, 196, 2, 2, 1, 2, 2, 2, 3, 29, 3, 3, 4, 3, 2, 3, 2, 2, 2, 2, 7, 128, 5, 2, 237, 2, 26, 18, 9, 3, 1]

Подела на 39: величине компонената су [165, 66, 2, 3, 2, 2, 1, 3, 2, 3, 2, 3, 3, 3, 14, 2, 2, 35, 6, 2, 45, 2, 3, 1, 7, 10, 21, 228, 2, 33, 3, 11, 1, 3, 20, 17, 13, 15, 10]

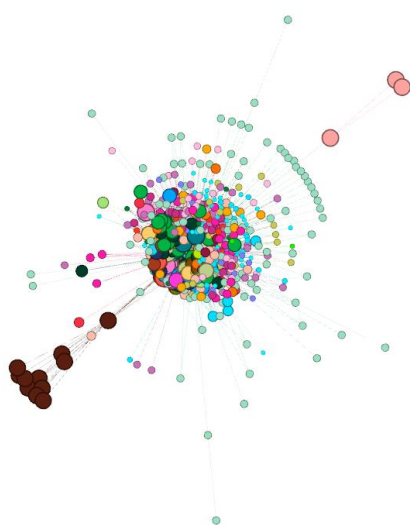
21. Ако се за параметар резолуције узме вредност 6 или већа добија се једна комуна што је тривијалан случај. За параметар резолуције 1.25 добија се 5 комуна величина [742, 12, 3, 2, 7].



За параметар резолуције добијају се 2 комуне величина [763, 3], што се подудара са резултатом спектралне анализе. Ту су се издвојили сабредити као зелени чворови {7moode, arabic, arabicblogs} чије су тежине грана [7, 20].



Ако се узме параметар резолуције 0.2525 добија се 39 комуна величина [41, 21, 23, 33, 25, 12, 38, 41, 42, 110, 25, 25, 28, 26, 14, 17, 20, 21, 7, 8, 14, 22, 12, 7, 7, 14, 22, 12, 12, 14, 9, 8, 7, 4, 12, 3, 4, 3, 3].



22. У наредној табели су приказана поређења коришћених метода за поделу мреже на комуне. Добија се да је модуларност за сваку мала, али да је ипак највећа за спектралну анализу која је поделила мрежу на 31 комуна, па је тај случај разматран.

Метода	Број комуна	Величине комуна	Модуларност
Асинхрона пропација лабела	3	[751, 12, 3]	0.000114223
Синхрона пропација лабела	5	[754, 5, 3, 2, 2]	4.85466e-05
Girvan-Newman	672	[88x1, 4x1, 2x4, 1x666]	0.000562923
Louvian	2	[7, 20]	1.542138e-05
Louvian	5	[742, 12, 3, 2, 7]	0.000175879
Louvian	39	[110 .. 3]	-0.018637961
Спектрално кластерисање	2	[763, 3]	1.542138e-05
Спектрално кластерисање	5	[747, 2, 2, 3, 12]	0.000121076
Спектрално кластерисање	6	[740, 2, 2, 7, 3, 12]	0.000179305

Спектрално кластерисање	7	[737, 2, 2, 3, 7, 3, 12]	0.000189579
Спектрално кластерисање	31	[237 .. 1]	0.022104973
Спектрално кластерисање	39	[228 .. 1]	0.013545307

У наставку су приказани сабрeдити који припадају комунама, и претпоставке зашто су тако груписани.

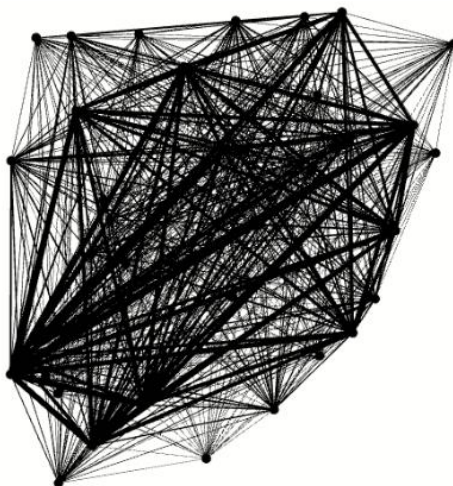
- [nederlands, nl] – Holandija i skraćenica za Holandiju. Okuplja korisnike koji su zainteresovani za Holandiju.
- [france, fr] – Francuska i njena skraćenica. Okuplja korisnike koji imaju veze ili interesovanja sa Francuskom.
- [ru, russia] – Rusija i skraćenica za Rusiju. Korisnici koji su zainteresovani za Rusiju imaju tendenciju da budu aktivni na oba sabredita.
- [mexico, latinoamerica] – Meksiko se nalazi u latinskoj Americi. Ova komuna okuplja korisnike koji su zainteresovani za Meksiko i latinsku Ameriku.
- [argentina, es, spanishlanguagenerds] – Ovi sabrediti privlače korisnike koji žele da upoznaju Argentinu, a samim tim i španski jezik.
- [braga, brasil] – Brada je opština u državi Brazil, pa su ova dva sabredita usko povezana i imaju zajedničke korisnike.
- [7moode, arabic, arabicblogs] – Komuna privlači Arape, jer prvi sabredit ima dosta obajva i komentara na arapskom, a druga dva su jasna.
- [ANTISHARIA, Islamofascism, antiterrorism, islam] – Sabrediti koji su bave temama terorizma, šerijata, Islama.
- [it, businessnews] – Komuna okuplja korisnike koji se bave biznisom i informacionim tehnologijama.
- [bloggerdal, tr, geyik] - bloggerdal je turski blog za muziku, tr je skraćenica za Tursku, geyik je turska reč za jelena, a postoji i turski košarkaš sa tim prezimenom. Ova komuna je interesantna korisnicima koji imaju veze sa Turskom.
- [math, compsci, haskell, lisp, erlang, cpp, codeprojects, ocaml, C_Programming, asm, haskell_proposals, types, scheme, functional, scala, git, agi, smalltalk] – Sabrediti vezani za matematiku, programiranje, programske jezike.
- [Libertarian, Anarchism, Military, conspiracy, economy, guns, ronpaul, Israel, government, AmericanGovernment, conspiracies, socialism, AmericanPolitics, 911truth, BlackOps, Liberty, Cryptogon, WorldWideNews, CommonLaw, ResistTyranny, antiwar, AmericanHistory, GlobalistNews, WorldWidePolitics, Conservative, mises, finance, joos, MMJ] – sabrediti koji se bave tajnim operacijama vlade i vojnih organizacija (BlackOps), kao i terorističkim napadima (911truth), i svime što ide uz to, kao što je politika, ekonomija, sloboda koje tom prilikom mogu biti narušene.
- [2008Olympics, travel, celebrities, Design, socialmedia, PalinProblem, WebNews, nyc, cute, fashion, Fitness, blogs, gossip, marketing, television, social_bookmarking, lifestyle, Autos, SEO, search, celebrity, Blogging, InternetMarketing, FinanceNews, ExDiggers, just4callie] – komuna koja okuplja korisnike koje zanimaju svakodnevne teme, kao što su vesti, moda, životni stil, dešavanja poznatih.
- [Infos] – sabredit predstavlja zasebnu komunu.
- [Descary] – sabredit predstavlja zasebnu komunu.
- [lgbt, MensRights, egalitarian, giveaways, SuicideWatch, Socialize, lovereddit, hotmanflesh, likeredit] – Sabrediti koji se bave ljudskim pravima, socijalnim problemima i podržavaju mišljenje da su svi jednaki.
- [deutschland, de, AnnikRubens] – Nazivi ovih sabredita predstavljaju Nemačku, skraćenicu za tu državu, i novinarku te zemlje, Larisu Vasilian koja ima podcast gde nastupa pod imenom Annik Rubens.

- [Gamer, Webbook] – Komuna obuhvata korisnike koji vole da igraju igrice, i zainteresovani su za webbook tj. prenosne računare.
- [ruby, rails, Clojure, Cricket, merb] – Ruby je open source programski jezik, Rails je razvojni okvir za veb napisan na programskom jeziku Ruby. Clojure je programski jezik opšte namene sa naglaskom na funkcionalnom programiranju. Merb je preteča Rails-a.
- [Ambientalmente, italy] – Komuna koja privlači korisnike koji se bave ekologijom i zaštitom životne sredine u Italiji.
- [Chinese, SinoBeta, cnbeta, Josie, tosoft, lamengao, XYON] – cnBeta je vodeći sajt u Kini za vesti iz lokalne i globalne IT industrije. Kineska kompanija SinoBeta je razvila pretraživač Chinasearch.com.
- [hardware, apple, google, linux, web_design, opensource, browsers, bugs, hackers, netsec, startups, torrents, worstof, Python, javascript, l33t, astro, robotics, tech, ComputerSecurity, firefox, productivity, software, joel, Ubuntu, ASPNET, microsoft, askXKCD, redditdev, perl, osx, BSD, django, iphone, PHP, mac, TestMySite, meta, kde, hacking, softwaredevelopment, dotnet, ReverseEngineering, security, Database, electronics, java, emacs, ITnews, typography, agile, mobile, cheats, floss, windows, Webmaster, encryption, dotfiles, gnu, vim, virtualization, management, onlycode] – sabrediti koji predstavljaju hardver, softver, IT industriju.

23. Брокери су они чворови у мрежи који имају релативно ниска мрежа ограничења, а високу релациону централност. За проналажење брокера коришћења је функција `bridges` из библиотеке `networkx.algorithms.bridges`, која проналази ламбда скуп тј. оне гране чијим би се уклањањем мрежа поделила на више неповезаних компоненти. Сабредит `reddit.com` се појављује у највећем броју грана, и он има највећу релациону централност, тако да он дефинитивно представља мост у мрежи. Од њега зависи 25 сабредита. Сабредит `programming` је 6. по важности што се тиче релационе централности. Његовим уклањањем добили бисмо још 3 неповезане компоненте. Међутим, он припада језгру мреже, тако да комуникација у том делу би могла да се обави и без њега. Од сабредита *Chinese* директно зависи 6 чворова, тако да и он представља брокера.

5. МОДЕЛОВАЊЕ SNETT МРЕЖЕ

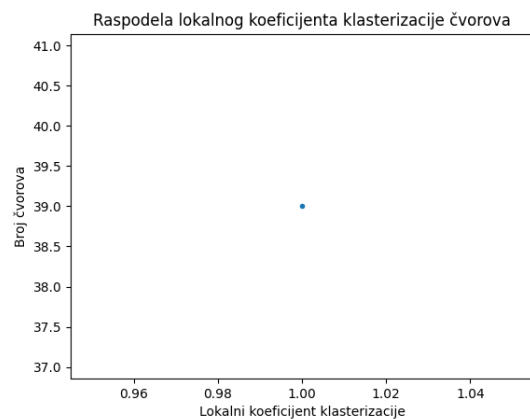
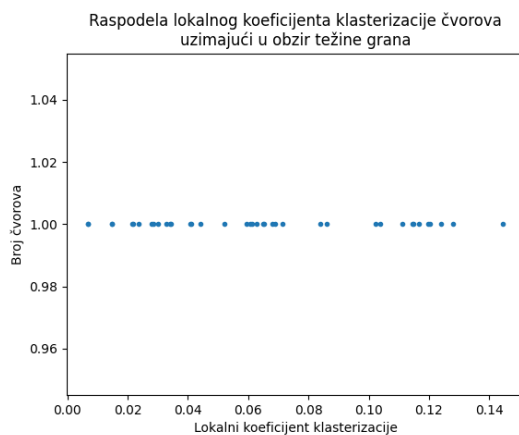
Мрежа SnetT је генерисана као подграф мреже Snet, где су чворови из скупа сабредита чија је тематика везана за економску кризу. Подграф се састоји од 39 чворова и 741 грана. У наредним потпоглављима је извршена анализа мреже SnetT, а на слици 5.1 је приказан модел мреже генерисан у *Gephy* алату.



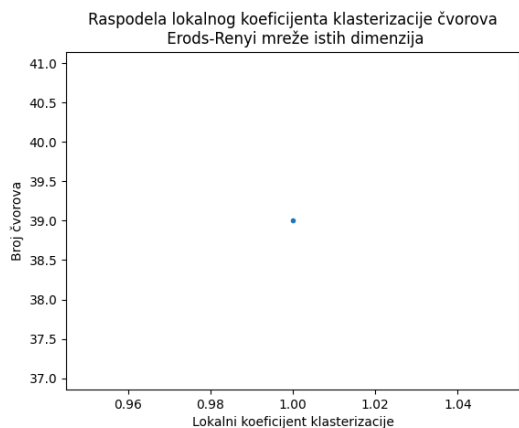
Слика 5.1 SnetT мрежа

5.1. Основне особине мреже

7. Густина мреже је 1.0.
8. Дијаметар мреже је 1, а просечне дистанце у оквиру мреже су 1.0. Просечна дистанца у Erdos-Reny мрежи истих димензија је 1.0.
9. Број повезаних компоненти је 1. Цела мрежа је повезана и комплетна, и представља гигантску компоненту мреже.
10. Просечни коефицијент кластеризације SNetT мреже је 1.0, док је просечни коефицијент кластеризације узимајући у обзир тежине грана 0.0648. Глобални коефицијент кластеризације је 1.0. На наредним сликама је приказана расподела локалног коефицијента кластеризације њених чворова узимајући и не узимајући у обзир тежине грана.



На наредној слици је приказана расподела локалног коефицијента кластеризације Erdos-Reny мреже истих димензија. Просечни коефицијент кластеризације Erdos-Reny мреже је 1.0, а глобални коефицијент кластеризације је такође 1.0. Пошто су коефицијенти исти закључујемо да је кластерисање изражено.



11. Феномен малог света је карактеристика реалних мрежа за које важи:

- 1) Мала просечна удаљеност било која два чвора – важи за SNetT јер је просечна дистанца у мрежи је 1.0.
- 2) Ограничен дијаметар у односу на величину мреже – важи за SNetT јер је дијаметар 1, а величина мреже 39.
- 3) Висок степен кластеризације – према резултатима питања 10) закључујемо да важи за SNetT.

Из наведеног следи да SNetT мрежа испољава особине малог света.

12. Коефицијент асортативности на основу тежинског степена чвора је 0.0001.

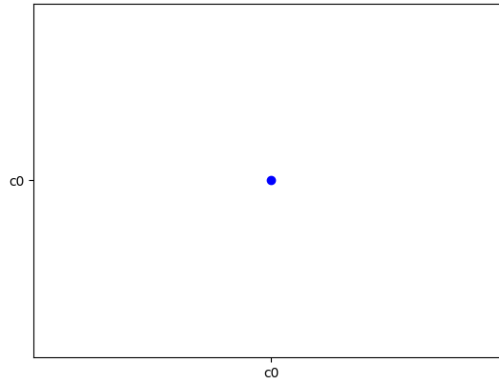
Коефицијент асортативности на основу степена чвора је NaN (*not a number*).

Коефицијент асортативности представља Пирсонов коефицијент корелације и рачуна

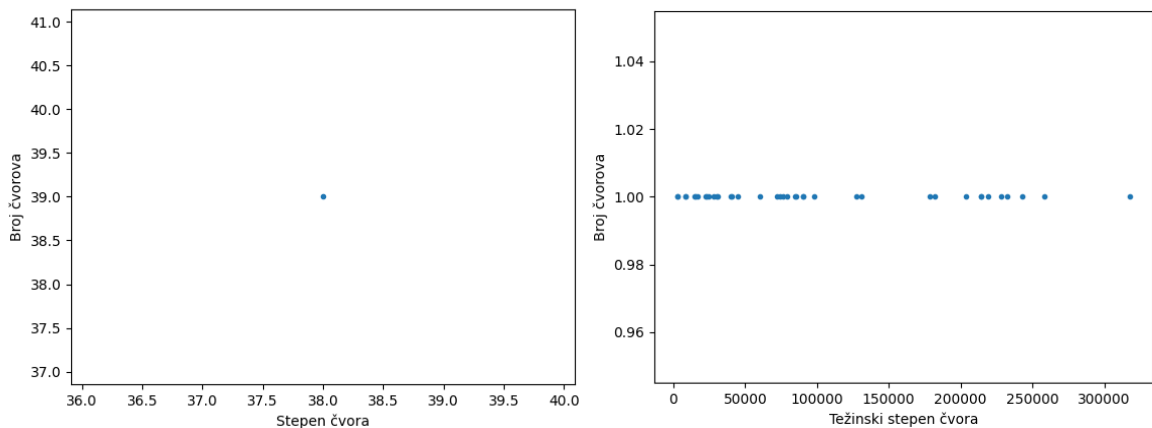
се по формули $\rho_{X,Y} = \frac{cov(X,Y)}{\sigma_X \sigma_Y}$, а стандардна девијација је $\sigma = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - \mu)^2}$. μ

аритметичка средина чвора и једнака је 38, степен сваког чвора x_i је такође 38. Стога добијамо да је стандардна девијација једнака 0, па је Пирсонов коефицијент недефинисан због дељења нулом. Пошто је граф комплетан закључује се да су чворови

истог степена међусобно повезани. Сви чворови имају исти степен чвора 38, па самим тим и припадају истој класи c0. Визуелизација је дата на следећој слици.



13. Из визуелизације SNetT мреже закључено је да структура мреже не изгледа као да испољава феномен клуба богатих. SNetT мрежа нема мали број јако добро повезаних чворова и велики број чворова са малим степеном, јер уопште нема чворова са малим степеном, сви имају максимални степен. Додавањем нових чворова не би се могло закључити да ће степени неких чворова расти знатно брже од осталих. Тиме закључујемо да мрежа не испољава феномен клуба богатих.
14. На наредним сликама је приказана дистрибуција чворова по степеном и дистрибуција чворова по тежинском степеном.



Дистрибуција чворова по степеном код случајно генерисане мреже са истим бројем чворова и грана (гнм генератор) и код Erdos-Reny мреже је иста као дистрибуција чворова по степеном SNetT мреже, и не прати power-law расподелу. Дистрибуција чворова по степеном би могла да се представи Дираковом функцијом $P(k) = \delta(k - 38) = \begin{cases} +\infty, & k = 38 \\ 0, & k \neq 38 \end{cases}$, а дистрибуција чворова по тежинском степеном функцијом $P(k) = \frac{1}{38} \sum_{i=0}^{38} \delta(k - d_{ti})$, где је d_{ti} тежински степен чвора i .

15. Ауторитети и хабови су карактеристични за усмерене мреже. Стога их нема потребе испитивати над SNetT мрежом јер је она неусмерена.

5.2. Анализа мера централности

16. Централност по степену за сваки чвор износи 1, јер је у питању комплетан граф (сваки чвор има степен $N-1$, и због нормализације се тај број дели са $N-1$, и добијамо вредност 1). Мера централности по блискости такође има вредност 1 (за сваки чвор се вредност $N-1$ дели са збиром најкраћих растојања од тог до сваког другог чвора што је $1*(N-1)$, па количник има вредност 1). Релациона централност је за сваки чвор једнака 0. Ако посматрамо један чвор и најкраће путеве између свака два преостала чвора, ниједан од најкраћих путева не пролази кроз посматрани чвор. Централности по степену, блискости и релациона централност за 10 чворова SNetT мреже су приказане у наредним табелама, и имају исте вредности за све чворове.

Табела 5.2.1 Централност по степену(лево) и централност по блискости (десно)

Чвор	Сабредит	Вредност	Чвор	Сабредит	Вредност
t5_2qh16l.0	technology	1.0	t5_2qh16	technology	1.0
t5_1rqwi	netsec	1.0	t5_1rqwi	netsec	1.0
t5_2qh03	gaming	1.0	t5_2qh03	gaming	1.0
t5_3b8o	joel	1.0	t5_3b8o	joel	1.0
t5_2qh0f	entertainment	1.0	t5_2qh0f	entertainment	1.0
t5_2qh61	WTF	1.0	t5_2qh61	WTF	1.0
t5_1a8ah	obama	1.0	t5_1a8ah	obama	1.0
t5_2qh1n	environment	1.0	t5_2qh1n	environment	1.0
t5_2qh33	funny	1.0	t5_2qh33	funny	1.0
t5_2qh9z	Health	1.0	t5_2qh9z	Health	1.0

Табела 5.2.2 Релациона централност

Чвор	Сабредит	Вредност
t5_2qh16	technology	0.0
t5_1rqwi	netsec	0.0
t5_2qh03	gaming	0.0
t5_3b8o	joel	0.0
t5_2qh0f	entertainment	0.0

t5_2qh61	WTF	0.0
t5_1a8ah	obama	0.0
t5_2qh1n	environment	0.0
t5_2qh33	funny	0.0
t5_2qh9z	Health	0.0

17. Централност по сопственом вектору

Суседи сваког посматраног чвора имају исти број суседа. Они чворови који имају веће тежине грана које их повезују са суседима имаће и већу вредност централности по сопственом вектору. Ниједан чвор није моћнији од осталих, јер суседи сваког чвора имају велики број суседа, и ниједан чвор не зависи од неког другог. Чворови са већом вредности централности по сопственом вектору су утицајни ако се разматра да ће информације брже да се шире кроз гране већих тежина.

Табела 5.2.3 Централност по сопственом вектору

Чвор	Сабредит	Вредност
t5_6	reddit.com	0.379795
t5_2cneq	politics	0.316792
t5_2qh0u	pics	0.295384
t5_2qh33	funny	0.282877
t5_mouw	science	0.276318
t5_2qh16	technology	0.267164
t5_2qh13	worldnews	0.262058
t5_2qh61	WTF	0.260284
t5_2qh0f	entertainment	0.249796
t5_2fwo	programming	0.223551

18. Израчунат је $\alpha_{\max} = (5.566e-06+0j)$, и за α је узимана вредност $\frac{1}{3}\alpha_{\max}$ и $\frac{2}{3}\alpha_{\max}$, а за β вредности из скупа $\{1,2,10,100\}$. Комбинација ових параметара у Katz-овој централности приказана је у наредним табелама. Повећањем параметра β за сабрeдит reddit.com повећава се и централност тог чвора и постаје значајније већа од централности осталих чворова.

$\alpha=1.837e-06$, $\beta=1$ за reddit.com, за остале 1

Чвор	Сабредит	Вредност
t5_6	reddit.com	0.230875
t5_2cneq	politics	0.211636
t5_2qh0u	pics	0.20613
t5_2qh33	funny	0.202516
t5_mouw	science	0.200952

$\alpha=3.674e-06$, $\beta=1$ за reddit.com, за остале 1

Чвор	Сабредит	Вредност
t5_6	reddit.com	0.309799
t5_2cneq	politics	0.268519
t5_2qh0u	pics	0.25561
t5_2qh33	funny	0.247629
t5_mouw	science	0.24381

$\alpha=1.837e-06$, $\beta=2$ за reddit.com, за остале 1

Чвор	Сабредит	Вредност
t5_6	reddit.com	0.338972
t5_2cneq	politics	0.209189
t5_2qh0u	pics	0.202667
t5_2qh33	funny	0.198838
t5_mouw	science	0.197134

$\alpha=3.674e-06$, $\beta=2$ за reddit.com, за остале 1

Чвор	Сабредит	Вредност
t5_6	reddit.com	0.366797
t5_2cneq	politics	0.267207
t5_2qh0u	pics	0.253287
t5_2qh33	funny	0.245113
t5_mouw	science	0.241157

$\alpha=1.837e-06$, $\beta=10$ за reddit.com, за остале 1

Чвор	Сабредит	Вредност
t5_6	reddit.com	0.764918
t5_2cneq	politics	0.163547
t5_2qh0u	pics	0.153573
t5_2qh33	funny	0.149425
t5_mouw	science	0.147374

$\alpha=3.674e-06$, $\beta=10$ за reddit.com, за остале 1

Чвор	Сабредит	Вредност
t5_6	reddit.com	0.609041
t5_2cneq	politics	0.247615
t5_2qh0u	pics	0.229774
t5_2qh33	funny	0.221135
t5_mouw	science	0.216754

$\alpha=1.837e-06$, $\beta=100$ за reddit.com, за остале 1

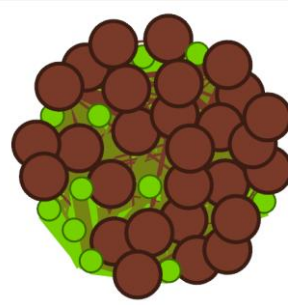
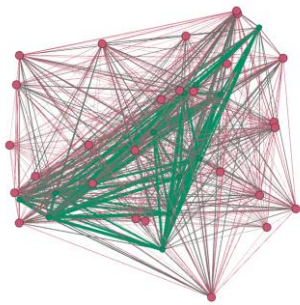
Чвор	Сабредит	Вредност
t5_6	reddit.com	0.970099
t5_2cneq	politics	0.084244
t5_2qh0u	pics	0.073489
t5_2qh33	funny	0.070021
t5_mouw	science	0.068137

$\alpha=3.674e-06$, $\beta=100$ за reddit.com, за остале 1

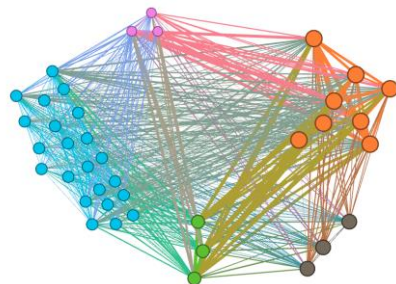
Чвор	Сабредит	Вредност
t5_6	reddit.com	0.825465
t5_2cneq	politics	0.198943
t5_2qh0u	pics	0.178438
t5_2qh33	funny	0.170167
t5_mouw	science	0.165754

19. 10 чворова са највећом вредношћу композитног ранга:

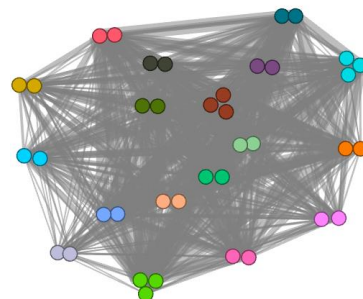
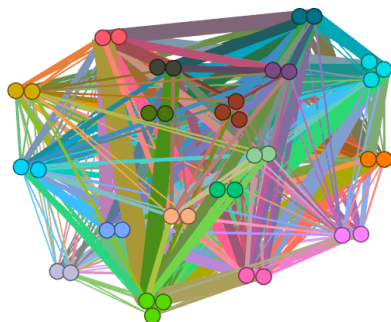
Чвор	Сабредит	DC ранг	CC ранг	BC ранг	EVC ранг	Композитни ранг
t5_6	reddit.com	20.0	20.0	20.0	1.0	8000.0
t5_2cneq	politics	20.0	20.0	20.0	2.0	16000.0
t5_2qh0u	pics	20.0	20.0	20.0	3.0	24000.0
t5_2qh33	funny	20.0	20.0	20.0	4.0	32000.0
t5_mouw	science	20.0	20.0	20.0	5.0	40000.0
t5_2qh16	technology	20.0	20.0	20.0	6.0	48000.0
t5_2qh13	worldnews	20.0	20.0	20.0	7.0	56000.0



За параметар резолуције 0.8 добијамо пет комуна величина [3, 3, 3, 8, 22] чворова. Међутим, ту се добија негативна модуларност од -0.065. Визуелизација мреже подељене на 5 комуна дата је на наредној слици.



За параметар резолуције 0.08 добијамо 18 комуна од којих су три комуне величине три чвора, а петнаест комуна величине два чвора, модуларности -0.065. Добијамо један од броја кластера који смо добили спектралном анализом. Визелизаија са и без акцента на гране дата је на наредним сликама.



22. Ако мрежу делимо на 2 комуне (јер је Лувенском методом једино у том случају добијена позитивна вредност), једна има субредите {*technology, science, politics, worldnews, funny, WTF, business, entertainment, pics, programming, reddit.com* }, а друга преостале. 11 чворова који су се издвојили у првој комуни имају највеће централности по сопственом вектору у интервалу (0.221315, 0.379795), док су централности по сопственом вектору за чворове из друге комуне у опсегу (0.002385, 0.150658). Подела на већи број комуна од 2 би имала модуларност мању од 0.

Програмским путем је испитано кластерисање применом Girman-Newman метода као и алгоритмима пропагације лабела и асинхроном пропагацијом лабела. У сва три случаја добијено је да је број комуна 1, тј. сви чворови припадају једној комуни, а самим тим се добија и модуларност од 0.

23. Брокери су они чворови у мрежи који имају релативна ниска мрежа ограничења, а високу релациону централност. Што су суседи актера слабије повезани, то он има више могућности за деловање. У SnetT мрежи су сви суседи свих чворова добро повезани, и ниједан чвор нема релациону централност већу од нуле, што доводи до

закључка да се ниједан чвор не може издвојити као брокер.

24. Поређење SNet и SNetT мрежа

	Гигантска компонента SNet мреже	SNet
Густина	0.025268 (0.017 за целу мрежу)	1
Просечне дистанце	2.099	1
Дијаметар	5	1
Број повизаних компоненти у целој мрежи	814	1
Просечни коефицијент кластеризације узимајући у обзир тежине грана	0.0003	0.0648
Просечни коефицијент кластеризације нерачунајући тежине грана	0.7706	1
Глобални коефицијент кластеризације	0.0011369691	1
Мрежа исказује феномен малог света	Да	Да
Коефицијент асортативности на основу тежинског степена чвора	-0.0323	0.0001
Коефицијент асортативности на основу степена чвора	-0.43665334856397703	Сви чворови истог степена, и сви међусобно повезани
Мрежа испољава феномен клуба богатих	Да	Не

25. Мрежа SnetT је подграф k -језгра мреже Snet, где је k 192, што значи да чворови SnetT мреже припадају k -језгру Snet мреже тј. постоји најмање 192 везе између чланова језгра.

6. МОДЕЛОВАЊЕ USERNET МРЕЖЕ

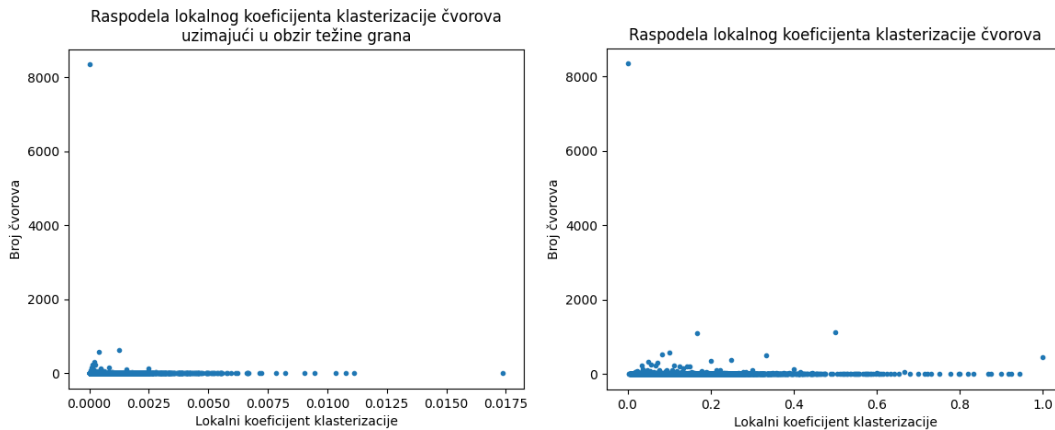
Мрежа UserNet представља мрежу корисника платформе *Reddit*. Корисник се налази у мрежи као чвор ако је коментарисао нечију објаву или коментар, или ако је постављао објаве, па су његове објаве коментарисане. Интеракције између корисника су моделоване вишеструким (*multiplex*) гранама, којима се памти да ли је неки корисник коментарисао објаву или коментар другог корисника. Гране су усмерене од корисника који коментарише ка оном кориснику на чију објаву или коментар је постављен коментар. Тежина грана одређена је агрегацијом бројањем.

Дакле, за конструисање мреже је коришћена функција *nx.MultiDiGraph()*. Тип гране (*parameter key*) може бити *submission* или *comment*, а тежина гране представља број коментара који је један корисник направио на објавама односно коментарима другог корисника.

У првом потпоглављу је извршена основна карактеризација мреже, док је у другом делу дата анализа мера централности. Треће потпоглавље обрађује тему детекција комуна.

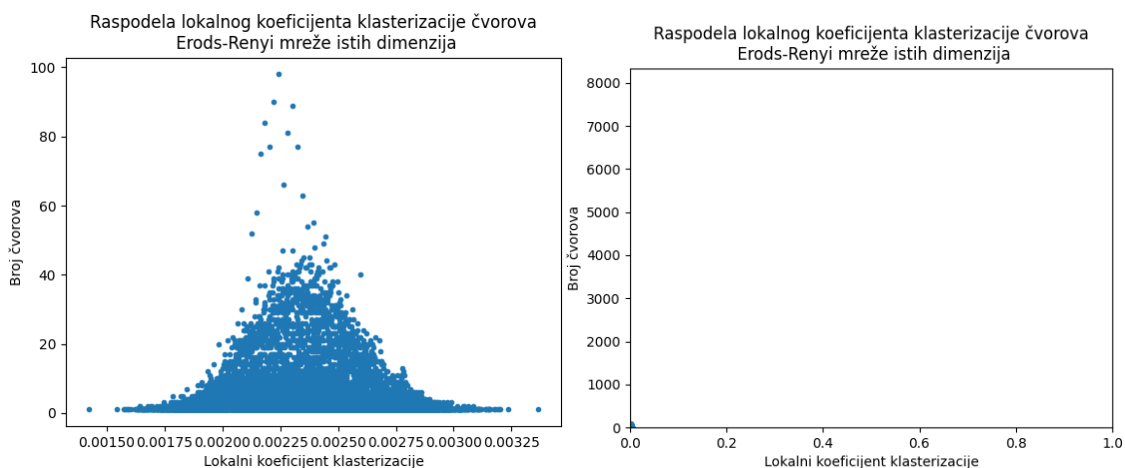
6.1. Основна карактеризација UserNet мреже

7. Густина UserNet мреже је 0.000301.
8. Мрежа није повезана, стога су просечна дужина и дијаметар мреже бесконачни. Просечна дистанца у Erdos-Reny мрежи истих димензија као UserNet мреже је 3.094588, а дијаметар је 4. Густина гигантске компоненте мреже је 0.001172. Просечна дистанца у гигантској компоненти UserNet мреже је 3.185, док је дијаметар 13. Просечна дистанца у Erdos-Reny мрежи истих димензија као гигантска компонента UserNet мреже је 2.761968, а дијаметар је 3. Закључујемо да је дијаметар UserNet мреже велики у односу на дијаметар гигантске компоненте UserNet мреже.
9. Број повезаних компоненти UserNet мреже је 48935. Величина компоненти је дата као низ где је први члан представља величину компоненте, а други број компоненти које су те величине: 48501x1, 11x1, 6x1, 5x2, 4x2, 3x12, 2x44, 1x48872. Постоји гигантска компонента мреже јер постоји језгро величине 48501 чворова и представља највећи део мреже тј. 49.73% мреже.
10. Просечни коефицијент кластеризације UserNet мреже је 0.117, док је просечни коефицијент кластеризације узимајући у обзир тежине грана 0.0004. Глобални коефицијент кластеризације је 3.1581423283315703e-06. На наредним сликама су приказане расподеле локалног коефицијента кластеризације, узимајући у обзир тежине грана и не узимајући их у разматрање. Са друге слике уочавамо да има преко 8000 чворова који не припадају ниједном троуглу.



Што се тиче испитивања да ли су чворови са високим локалним коефицијентом кластеризације међусобно повезани, програмским путем је испитиван подграф гигантске компоненте који се састоји само од чворова чији је локални степен кластеризације једнак 1, и њихових грана. Закључује се да тај подграф има 446 чворова и једну грану, и да је његова густина 0.00001, што значи да су ти чворови слабо међусобно повезани. Исти поступак је примењен и за релаксирани критеријум да се испита подграф којег чине чворови чији је локални коефицијент кластеризације већи или једнак 0.5. У том случају добијамом подграф од 1836 чворова, 28 грана, и густине 0.0000166. Дакле чворови високг локалног коефицијента кластеризације нису изразито повезани међусобно.

Просечни коефицијент кластеризације Erdos-Reny мреже истих димензија је 0.002, док је глобални коефицијент кластеризације 1.029431395910057e-07. На наредним сликама је приказана расподела локалног коефицијента кластеризације Erdos-Reny мреже. Опсег броја чворова који имају неки локални коефицијент кластеризације иде од 0 до 100, што је и приказано на левој слици. На десној слици је приказана иста дистрибуција чворова по локалном коефицијенту кластеризације, с тим што се посматра у односу на у осу максималне вредности 8000, због лакшег поређења тог графика и графика расподеле локалног коефицијента кластеризације чворова гигантске компоненте.



Поређењем графика и коефицијената кластеризације гигантске компоненте UserNet и Erdos-Reny мреже, добијамо да је кластерисање више изражено у гигантској компоненти UserNet мреже.

11. Феномен малог света је карактеристика реалних мрежа за које важи:

1. Мала просечна удаљеност било која два чвора – важи за гигантску компоненту UserNet мреже јер је просечна дистанца у гигантској компоненти 3.094588, а у Erdos-Reny мрежи димензија као гигантска компонента је је 2.761968.
2. Ограничен дијаметар у односу на величину мреже – важи за гигантску компоненту UserNet мреже јер је дијаметар 13, а величина мреже 48501, иако је у Erdos-Reny мрежи истих димензија као гигантска компонента дијаметар нешто мањи и износи 3.
3. Висок степен кластеризације – према резултатима питања 10) закључујемо да важи за гигантску компоненту UserNet мреже.

Из наведеног важи да гигантска компонента UserNet мреже испољава особине малог света.

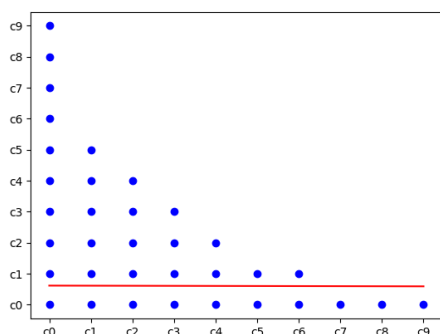
12. Коефицијент асортативности UserNet мреже на основу улазног тежинског степена чвора је -0.0196026761, а на основу излазног тежинског степена чвора је 0.0030034557. Коефицијент асортативности UserNet мреже на основу улазног степена чвора је -0.0254513238, а на основу излазног степена чвора је 0.0074803688.

Асортативно мешање по улазном степену чвора

Чворови графа су категориички груписани према улазном степену чвора у скупове c0-c9 једнаких величина. Прерачуната је матрица асортативног мешања која је приказана у наредној табели. У табели су поред имена скупова приказани и опсези степене чворова који припадају тој групи.

	c0 (132, 8434)	c1 (53, 132)	c2 (26, 53)	c3 (14, 26)	c4 (8, 14)	c5 (5, 8)	c6 (3, 5)	c7 (2, 3)	c8 (1, 2)	c9 (1, 1)
c0	0.4669478	0.1015357	0.0450527	0.0230817	0.0126757	0.0073173	0.0042939	0.0025862	0.0014317	0.0012096
c1	0.1148666	0.0231724	0.0102938	0.0053476	0.0029559	0.0016457	0.0009673	0.0005896	0.0003153	0.0002601
c2	0.0531015	0.0108203	0.0047768	0.0024327	0.0013606	0.0007946	0.0004557	0.000279	0.0001538	0.0001259
c3	0.0287068	0.0058091	0.002609	0.0013482	0.0007655	0.0004234	0.0002427	0.0001495	8.02e-05	6.53e-05
c4	0.0163529	0.0033114	0.0015264	0.0008055	0.0004405	0.0002514	0.0001488	8.45e-05	4.43e-05	4.1e-05
c5	0.0097619	0.0019175	0.0009295	0.0004542	0.0002656	0.0001524	8.64e-05	6.86e-05	3.19e-05	1.96e-05
c6	0.0063068	0.0012651	0.0005656	0.000283	0.0001673	9.65e-05	6.35e-05	3.37e-05	1.96e-05	1.6e-05
c7	0.0041539	0.0008312	0.0003886	0.000201	0.0001154	7.29e-05	4.17e-05	3.27e-05	1.71e-05	1.45e-05
c8	0.0028173	0.0005801	0.0002569	0.0001411	8.34e-05	4.39e-05	3.16e-05	1.89e-05	1.34e-05	9.1e-06
c9	0.0024817	0.0005388	0.0002344	0.0001263	8.27e-05	4.75e-05	2.9e-05	1.6e-05	1.05e-05	1.02e-05

Сабирањем елемента главе дијагонале матрице добијамо да је вероватноћа 0.497 да су чворови истих група међусобно повезани. Визуелизација је дата на наредној слици. Коефицијент асортативности за категориички подељене чворова према њиховим уазним степенима чвора је -0.0081. Како је добијена негативна вредност, то асортативност по улазном степену није изражена. Поделом чворова по улазном степену и на 15 и на 30 класа добијане су негативне вредности блиске нули.

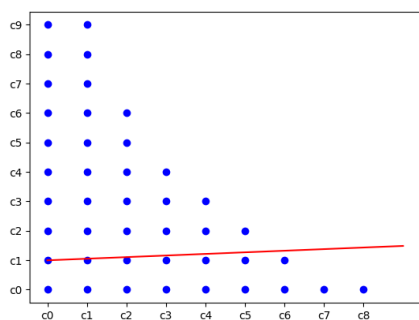


Асортативно мешање по излазном степену чвора

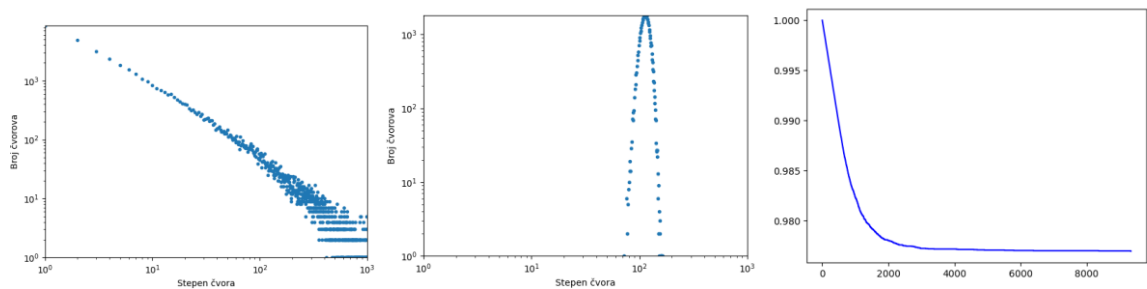
Чворови графа су категорички груписани у скупове c0-c9 једнаких величина према излазном степену чвора. Прерачуната је матрица асортативног мешања која је приказана у наредној табели. У табели су поред имена скупова приказани и опсези излазног степена чворова који припадају тој групи.

	c0 (131, 3747)	c1 (51, 131)	c2 (24, 51)	c3 (13, 24)	c4 (8, 13)	c5 (5, 8)	c6 (3, 5)	c7 (2, 3)	c8 (1, 2)	c9 (1, 1)
c0	0.4434017	0.1147534	0.0581493	0.0344378	0.0216859	0.0143657	0.0105043	0.008389	0.0060249	0.0051977
c1	0.08681	0.0233788	0.0125854	0.007708	0.0048041	0.0031028	0.0024418	0.0018841	0.0013497	0.0011019
c2	0.0365367	0.0101012	0.005583	0.0034627	0.0021294	0.0014433	0.0011944	0.0008668	0.0006505	0.0005275
c3	0.0182918	0.0052906	0.0029163	0.001899	0.0011574	0.0007409	0.0006698	0.0004608	0.0003661	0.0002765
c4	0.010091	0.0029755	0.0017397	0.0010819	0.0006581	0.0004503	0.0004031	0.0002609	0.0002072	0.0001556
c5	0.0060224	0.0017597	0.0010286	0.0006317	0.0004205	0.000279	0.0002663	0.0001477	0.0001321	0.0001041
c6	0.0036227	0.0010543	0.0006074	0.0003882	0.0002409	0.0001618	0.0001586	0.0001041	8.34e-05	7.51e-05
c7	0.0022281	0.0006407	0.0003762	0.0002485	0.0001549	0.0001023	0.0001092	6.46e-05	5.3e-05	4.93e-05
c8	0.0011987	0.0003625	0.0002064	0.0001477	9.61e-05	5.91e-05	6.86e-05	3.77e-05	3.45e-05	3.08e-05
c9	0.0009665	0.0002717	0.0001796	0.0001125	6.49e-05	4.64e-05	5.84e-05	2.76e-05	2.9e-05	2.14e-05

Сабирањем елемента главне дијагонале матрице добијамо да је вероватноћа 0.4755 да су чворови истих група међусобно повезани. Визуелизација је дата на наредној слици. Коефицијент асортативности за категорички подељене чворова према њиховим степенима чвора је 0.0144. Како је добијена ниска позитивна вредност, то асортативност по излазном степену није изражена. Поделом чворова по излазном степену и на 15 и на 30 класа добијане су ниске позитивне вредности.



13. Посматрањем улазног ненормализованог степена чвора, уочава се да постоје чворови вишег улазног степена, реда неколико хиљада, где је највећи степена 9335. То су чворови чије су објаве и коментари коментарисани од стране највећег броја људи. Постоји могућност да ће нови корисници који долазе на платформу прво уочити баш њихове обајаве и коментаре, и да ће и сами оставити коментар, и тиме створити везу прво ка тим ауторима. На исти начин анализирањем излазног ненормализованог степена чвора, уочава се да постоје чворови вишег излазног степена, реда неколико хиљада, где је највећи степена 4155. То су чворови који су оставили коментар највећем броју различитих људи. Вероватно је да ће они међу првима коментарисати и новим корисницима платформе, те да ће излазна грана од њих ка новим корисницима прва бити створена. Дакле, претпоставља се да ће чворови вишег улазног/излазног степена имати предност при доласку нових чланова и стварњу нови улазних/излазних грана. Тиме закључујемо мрежа UserNet испољава феномен клуба богатих.
14. На наредним сликама су дате дистрибуције чворова по улазном степену за UserNet мрежу, за Erdos-Reny мрежу истих димензија и комплементарна кумулативна расподела улазног степена чвора, која није налик правој. Стога не делује као да би могла да се апроксимирала power law расподелом.

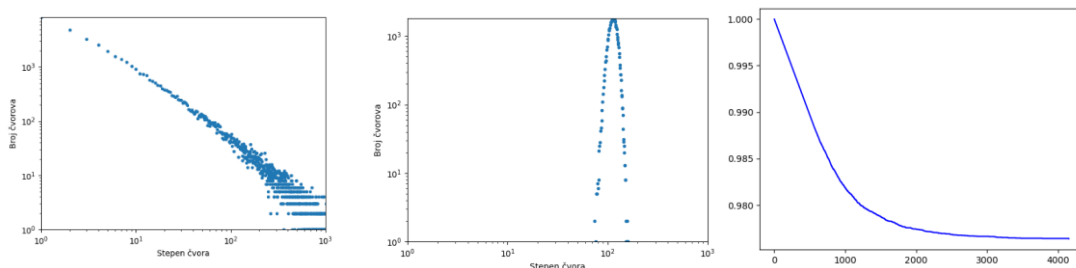


Power law paket је прерачунао да се најбоље фитује са power law, међутим тек од вредности $x_{min} = 1545$, $\alpha = 3.8395529$, $\sigma = 0.2719798$.

power_law truncated_power_law: R: 1.408328431651e-07 p: 0.9995765452464787	power_law exponential: R: 9.158367946487301 p: 0.0351827993061	power_law stretched_exponential: R: 0.466863111092011 p: 0.4277491543005
---	---	---

У свим поређењима добило се да је power law бољи, али се то не може са сигурношћу тврдити јер је статистички значај мањи од 0.05 само код поређења са експоненцијалним моделом.

Следе дистрибуције чворова по излазном степену за UserNet мрежу, за Erdos-Reny мрежу истих димензија и комплементарна кумулативна расподела излазног степена чвора, која није налик правој. Стога не делује као да би могла да се апроксимирала power law расподелом.



Power law пакет је прерачунао да се најбоље фитује са power law, међутим тек од вредности $x_{min}=886, \alpha=3.57429896, \sigma=0.1269808$.

Поређењем *power law* расподеле са осталим, увек добијамо негативно R, а p мање од 0.05. Даље се пореди *truncated power law* са осталим.

truncated_power_law lognormal: R: 1.1601647961829196 p: 0.0017392469247850	power_law exponential: R: 3.3876673742935486 p: 0.1435933856559439	power_law stretched_exponential: R: 0.4358648671032608 p: 0.0393704963995158
---	---	---

Пошто су све вредности R позитивне, а p је претежно мање од 0.05, узеће се да се расподела излазног степена чвора апроксимира *truncated power law* расподелом, где је $x_{min} = 886$, $\beta = 1.679575219$.

15. У следећим табелама су приказани 10 најзначајнијих хабова и ауторитета.

Чвор	Вредност	In-d	Out-d
malcontent	0.004142	2135	3102
matts2	0.003789	1869	3133
NoMoreNicksLeft	0.003225	2818	4155
glengyron	0.001957	1588	2255
aletoledo	0.001898	1762	2561
eleitl	0.001829	1150	2349
randomb0y	0.001824	2319	3296
alllie	0.001792	6388	2589
cartooncorpse	0.001689	647	1588
sakebomb69	0.001688	1116	1846

Чвор	Вредност	In-d	Out-d
qgyh2	0.006663	8360	1186
rmuser	0.006567	9335	1668
matts2	0.004418	1869	3133
malcontent	0.004336	2135	3102
alllie	0.004291	6388	2589
noname99	0.003076	5791	188
NoMoreNicksLeft	0.003014	2818	4155
maxwellhill	0.002995	4683	44
georgewashingtonblog	0.002905	4055	384
twolf1	0.00284	4101	134

Хабови су међусобно повезани јер чине подграф од 10 чворова и 90 грана. Низ [8, 9, 9, 8, 8, 8, 7, 5, 3, 7] представља број грана од једног хаба ка преосталим, тј. да први чвор има излазне гране ка преосталих 8 чворова, други са преосталих 9. Ауторитети су такође међусобно повезани јер чине подграф од 10 чворова и 104 грана. Низ [9, 9, 9, 9, 5, 9, 2, 8, 5] представља број грана од једног ауторитета ка преосталим, тј. да први чвор има излазне гране ка преосталих 8, други са преосталих 9. По броју улазних и излазних грана у и из чворова који представљају хабова и ауторитете види се да они нису на периферији већ централније у мрежи.

6.2. Анализа мера централности

16. Централности по степену, блискости и релациона централност

Табела 6.2.1 Централност по улазном степену

Чвор	Вредност
rmuser	0.192474
qgyh2	0.172371
alllie	0.131711
noname99	0.119402
swampsparrow	0.113567
gst	0.101155
maxwellhill	0.096557
7oby	0.09433
twolf1	0.084557
georgewashingtonblog	0.083608

Табела 6.2.2 Централност по излазном степену

Чвор	Вредност
NoMoreNicksLeft	0.08567
mexicodoug	0.077979
7oby	0.07167
h0dg3s	0.070206
mutatron	0.068062
randomb0y	0.067959
Poromenos	0.066206
nixonrichard	0.064784
matts2	0.064598
fingers	0.064289

Табела 6.2.3 Центр. по блискости (долазна удаљеност до чвора)

Чвор	Вредност
rmuser	0.516166
qgyh2	0.509486
noname99	0.492641
allli	0.488478
swampsparrow	0.484245
maxwellhill	0.480174
7oby	0.475719
gst	0.474439
twolf1	0.473744
georgewashingtonblog	0.470668

Табела 6.2.4. Центр. по блискости (излазна удаљеност до чвора)

Чвор	Вредност
mexicodoug	0.484008
NoMoreNicksLeft	0.482069
7oby	0.478422
h0dg3s	0.475873
Poromenos	0.473596
otakucod	0.473466
randomb0y	0.472318
BraveSirRobin	0.472061
nixonrichard	0.470732
malcontent	0.469888

Табела 6.2.5 Релациона централност

Чвор	Вредност
rmuser	0.023185
alllie	0.018251
7oby	0.017745
qgyh2	0.015374
NoMoreNicksLeft	0.011833
nixonrichard	0.010185
malcontent	0.010111
Poromenos	0.009149
swampsparrow	0.009115
mutatron	0.008909

17. 10 чворова са највећом вредношћу централности по сопственом вектору који одговарају улазним гранама:

Autor	EVC Vrednost	Raspodela suseda čvora po centralnosti po izl. stepenu	DC
malcontent	0.211252	0.1x20, 0.0x2919	0.107979
qgyh2	0.205532	0.1x14, 0.0x1067	0.196825
rmuser	0.183917	0.1x22, 0.0x1560	0.226866
NoMoreNicksLeft	0.169802	0.1x24, 0.0x3723	0.143773
matts2	0.156033	0.1x23, 0.0x2885	0.103134
gst	0.145475	0.0x3	0.101216
maxwellhill	0.143696	0.1x2, 0.0x39	0.097464
nixonrichard	0.140849	0.1x23, 0.0x2953	0.122289
noname99	0.135867	0.1x5, 0.0x174	0.123278
alllie	0.13354	0.1x20, 0.0x2367	0.185093

Као четврта колона табеле дато је са колико чворова и којих централности по излазном степену (заокружени на 1 децималу) је сваки од чворова са највећом вредношћу централности по сопственом вектору (који одговарају улазним гранама)

повезан. Гране које улазе у представљене аутора излазе из суседа тог аутора, а суседи имају ниску централност по излазном степену. Значи да суседи када преносе информације због мале централности по излазном степену преносе их малом броју чворова, па да ли ће њихове информације стићи на одредиште зависи од ранжираних чворова аутора. Закључује се да суседи зависе од ранжираних аутора, што ауторе чини моћнији.

10 чворова са највећом вредношћу централности по сопственом вектору који одговарају излазним гранама:

Autor	EVC Vrednost	Raspodela suseda čvora po centralnosti po ul. stepenu	DC
matts2	0.286577	0.2x2, 0.1x25, 0.0x2881	0.103134
NoMoreNicksLeft	0.19215	0.2x2, 0.1x26, 0.0x3719	0.143773
malcontent	0.191082	0.2x2, 0.1x24, 0.0x2913	0.107979
aletoledo	0.154657	0.2x2, 0.1x23, 0.0x2396	0.089134
redditcensoredme	0.142252	0.2x2, 0.1x16, 0.0x2044	0.07266
mexicodoug	0.129159	0.2x2, 0.1x23, 0.0x3701	0.113505
Toby	0.115168	0.2x2, 0.1x23, 0.0x3292	0.166
Bensch	0.107808	0.2x2, 0.1x20, 0.0x1655	0.056598
nixonrichard	0.1071	0.2x2, 0.1x25, 0.0x2949	0.122289
otakucode	0.104666	0.2x2, 0.1x25, 0.0x2882	0.109485

Као четврта колона табеле дато је са колико чворова и којих централности по улазном степену (заокружени на 1 децималу) је сваки од чворова са највећом вредношћу централности по сопственом вектору који одговарају (излазним гранама) повезан. Гране које излазе из представљених аутора улазе у суседе тог аутора, а суседи имају ниску централност по улазном степену. Закључује се да суседи зависе од ранжираних аутора, што ауторе чини моћнији.

18. У следећој табели су дати корисници који су били најактивнији на сабретиту reddit.com:

Корисник	Број активности
RugerRedhawk	3890
NoMoreNicksLeft	3551
mutatron	3341
mdinfo	3329
rmuser	3300

Katz-ова централност је испитана за најактивнијег корисника *RugerRedhawk*.

Коефицијент α_{\max} није могао програмским путем да се израчуна, јер није подржана функција $\max(nx.adjacency_spectrum(graph))$ која рачуна λ_{\max} , за граф величине као што је UserNet. Зато је узета вредност α 5.452e-06, за коју се показало да није већа од α_{\max} , тј. програмски код успева да се изврши.

За α је потом узимана вредност $1/3$ и $2/3$ претходно изабраног α , а за β вредности из скупа $\{1,2,10,100\}$. Када сви аутори имају исту вредност параметра $\beta=1$, аутор *RugerRedhawk* се не рангира у првих 5. Повећањем параметра β повећава се и централност тог чвора, који постаје рангиран у првих 5, а потом има значајно веће централности од осталих чворова.

$\alpha=1.799\text{e-}06$, $\beta=1$ за RugerRedhawk, за остале 1

Чвор	Вредност
rmuser	0.004688
qgyh2	0.004681
gst	0.004649
noname99	0.004631
alllie	0.004622

$\alpha=3.598\text{e-}06$, $\beta=1$ за RugerRedhawk за остале 1

Чвор	Вредност
rmuser	0.004836
qgyh2	0.004821
gst	0.004758
noname99	0.004721
alllie	0.004704

$\alpha=1.799\text{e-}06$, $\beta=2$ за RugerRedhawk, за остале 1

Чвор	Вредност
RugerRedhawk	0.009102
rmuser	0.004688
qgyh2	0.00468
gst	0.004649
noname99	0.004631

$\alpha=3.598\text{e-}06$, $\beta=2$ за RugerRedhawk, за остале 1

Чвор	Вредност
RugerRedhawk	0.009122
rmuser	0.004836
qgyh2	0.00482
gst	0.004758
noname99	0.004721

$\alpha=1.799\text{e-}06$, $\beta=10$ за RugerRedhawk, за остале 1

Чвор	Вредност
RugerRedhawk	0.045377
rmuser	0.004683
qgyh2	0.004676
gst	0.004645
noname99	0.004628

$\alpha=3.598\text{e-}06$, $\beta=10$ за RugerRedhawk, за остале 1

Чвор	Вредност
RugerRedhawk	0.045393
rmuser	0.004831
qgyh2	0.004816
gst	0.004753
noname99	0.004719

$\alpha=1.799\text{e-}06$, $\beta=100$ за RugerRedhawk, за остале 1

Чвор	Вредност
RugerRedhawk	0.413418
rmuser	0.004269
qgyh2	0.004263
gst	0.004235
noname99	0.004231

$\alpha=3.598\text{e-}06$, $\beta=100$ за RugerRedhawk за остале 1

Чвор	Вредност
RugerRedhawk	0.413388
rmuser	0.004405
qgyh2	0.004392
gst	0.004335
noname99	0.004327

19. Чворови гигантске компоненте мреже са највећом вредношћу композитног ранга

Autor	IN DC rang	OUT DC rang	IN CC rang	OUT CC rang	BC rang	Kompozitni rang
7oby	8.0	3.0	7.0	3.0	3.0	1512.0
NoMoreNicksLeft	20.0	1.0	12.0	2.0	5.0	2400.0
rmuser	1.0	92.0	1.0	100.0	1.0	9200.0
alllie	3.0	19.0	4.0	34.0	2.0	15504.0
nixonrichard	22.0	8.0	13.0	9.0	6.0	123552.0
mexicodoug	79.0	2.0	79.0	1.0	11.0	137302.0
Poromenos	28.0	7.0	26.0	5.0	8.0	203840.0
randomb0y	31.0	6.0	27.0	7.0	14.0	492156.0
mutatron	41.0	5.0	24.0	16.0	10.0	787200.0
otakucode	35.0	12.0	28.0	6.0	12.0	846720.0

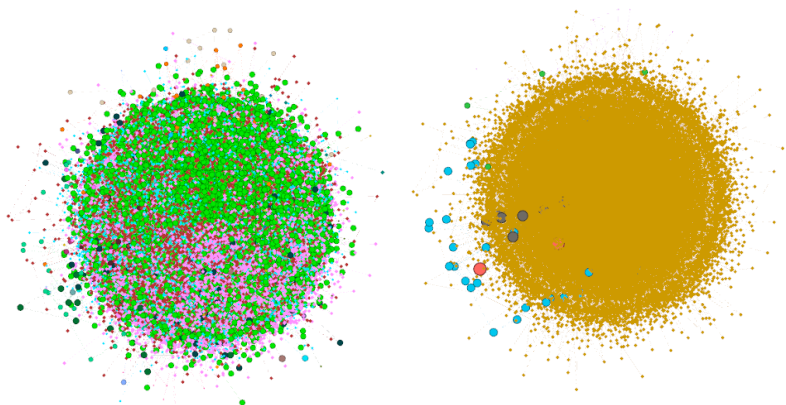
6.3. Детекција комуна мреже *UserNet*

20. Да би се спровела спектрална анализа неопходно је да се користи функција `nx.laplacian_matrix(graph)`. Она није имплементирана за усмерене графове. Стога је коришћен метод асинхроне пропагације лабела. Процељен је број комуна 44, а величине комуна су дате следећим низом - [48278, 23, 13, 4, 2, 2, 2, 2, 12, 6, 2, 6, 7, 2, 2, 3, 4, 2, 19, 30, 2, 2, 6, 2, 3, 2, 9, 3, 3, 14, 2, 2, 3, 2, 4, 2, 4, 2, 2, 3, 2, 2, 2, 2].

Модуларност која се добија коришћењем овог метода је 0.000788.

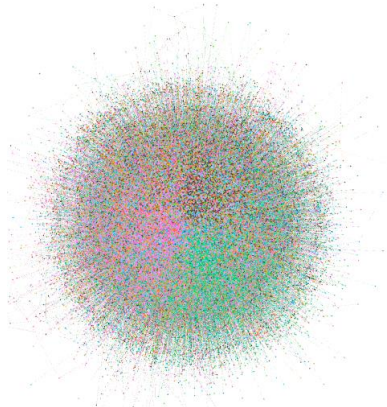
За параметра резолуције 10 добијају се две комуне и модуларност 0, док се за вредност 85 добија једна комуна, што је тривијалан сличај. За параметар резолуције 1.0 добија се подела графа на 39 комуна, модуларности 0.173, и величина [2, 3, 23, 9, 28, 2, 9, 9, 10, 3, 11517, 99, 2, 2, 2, 17, 10196, 15188, 10, 2, 3, 37, 2, 18, 3, 30, 4, 4, 3, 10, 3, 2, 11092, 2, 133, 3, 2, 15, 2]. Визуелизација је дата на наредној слици лево.

За параметар резолуције 5.0 добија се подела на 6 комуна, модуларности 0.0004220855, величина комуна [12, 48444, 6, 30, 7, 2]. Визуелизација је дата на наредној слици десно.



За параметар резолуције 0.7 добија се подела на 44 комуна, модуларност 0.13955320580942157. Величине добијених комуна су [1765, 3170, 132, 382, 723, 1253, 776, 135, 583, 212, 380, 269, 791, 4963, 502, 330, 256, 333, 5804, 4746, 6699, 186, 1815, 186, 3, 30, 161, 173, 108, 789, 793, 2167, 2033, 156, 1521, 156, 328, 15, 80, 255, 103, 229, 997, 2013].

Поређењем са методом асинхроне пропагације лабела (питање 20) којом је предложен исти број комуна, закључује се да је Лувенска метода дала бољу поделу.



- 21.** Из претходног питања се закључује да је најбоља подела по питању модуларности добијена Лувенском методом са параметром резолуције 10, и 39 комуна. Пошто неке од уочених комуна имају преко десет хиљада чворова, у наставку су приказане оне које су мање.

Комуне су детековане зато што постоје корисници сличних интересовања који посећују исте теме, и временом почну да коментаришу једни другима. Што је више коментара између њих, и то обостраних, и коментара на коментар и на објаву то су чворови склонији да се групишу у комуну.

- [fieldmouse, hesika]
- [allenmunroe, darrylhickson, willhasbro]
- [akadit, shiro, iwadon, kazuya, kawauchi, mayuge_born, mattn, yuseit, hengsu, jmur, tmiya, torus, flame, crane, ikegami--, iwata, bongole, syobosyobo, __hage, ogijun, bopper, antarest, unno]
- [dneuman911, Smitty1315, smitty13, Snorelock, rasbill, Slinky31, MasterAaron01, Mugger15, pasterhatt]
- [mhunt346, moosetime1230330, foxxyb069, ringlen837, donq321, gry_own, boooomsquad, bdoover234, ATsDIGwagon, jkustalicious, youboozeyoulose, grrtt_green, dhertz666, luvthymullet, ppot667, gbypass654, callturk245, tearsofmysoul99, obama4pres08, kyle9456, dizangfoo, noneMediocre29, jstalin456, starrj0548, dane_tmIn, time4drift, patrickvannypan, thelmajanemax]

- 22.** Брокери су они чворови у мрежи који имају релативно ниска мрежа ограничења, а високу релациону централност. Чворови највеће релационце централности су rmuser, allie, 7oby, qgyh2, NoMoreNicksLeft, nixonrichard, malcontent, Poromenos, swampsparrow, mutatron, и они су брокери мреже.

ЛИТЕРАТУРА

- [1] <https://networkx.org/documentation/stable/reference/algorithms/generated/networkx.algorithms.cluster.clustering.html#networkx.algorithms.cluster.clustering>
- [2] <https://rti.etf.bg.ac.rs/rti/mslasm/materijali/>