

УНИВЕРЗИТЕТ У БЕОГРАДУ  
ЕЛЕКТРОТЕХНИЧКИ ФАКУЛТЕТ



## **АНАЛИЗА СОЦИЈАЛНИХ МРЕЖА**

Пројектни задатак

Ментори:

др Марко Мишић, доцент

др Јелица Протић, редовни професор

Предраг Обрадовић, асистент

Кандидати:

Никола Ристић 2021/3281

Београд, јул 2022.

# САДРЖАЈ

<b>САДРЖАЈ .....</b>	<b>2</b>
<b>1. УВОД.....</b>	<b>3</b>
<b>2. ЧИШЋЕЊЕ ПОДАТАКА.....</b>	<b>4</b>
<b>3. СТАТИЧКА ОБРАДА ПОДАТАКА .....</b>	<b>5</b>
<b>4. SUBREDDIT NETWORK (SNET) .....</b>	<b>7</b>
4.1. КРЕИРАЊЕ МРЕЖЕ .....	7
4.2. ИЗРАЧУНАВАЊЕ МЕТРИКА МРЕЖЕ .....	7
4.3. ИЗРАЧУНАВАЊЕ ЦЕНТРАЛНОСТИ .....	9
4.4. ДЕТЕКЦИЈА КОМУНА .....	10
<b>5. SUBREDDIT NETWORK (SNETT).....</b>	<b>13</b>
5.1. КРЕИРАЊЕ МРЕЖЕ .....	13
5.2. ИЗРАЧУНАВАЊЕ МЕТРИКА МРЕЖЕ .....	13
5.3. ИЗРАЧУНАВАЊЕ ЦЕНТРАЛНОСТИ .....	15
5.4. ДЕТЕКЦИЈА КОМУНА .....	16
<b>6. SUBREDDIT NETWORK (SNETF).....</b>	<b>20</b>
6.1. КРЕИРАЊЕ МРЕЖЕ .....	20
6.2. ИЗРАЧУНАВАЊЕ МЕТРИКА МРЕЖЕ .....	20
6.3. ИЗРАЧУНАВАЊЕ ЦЕНТРАЛНОСТИ .....	22
6.4. ДЕТЕКЦИЈА КОМУНА .....	23
<b>7. USERNET.....</b>	<b>27</b>
7.1. КРЕИРАЊЕ МРЕЖЕ .....	27
7.2. ИЗРАЧУНАВАЊЕ МЕТРИКА МРЕЖЕ .....	27
7.3. ИЗРАЧУНАВАЊЕ ЦЕНТРАЛНОСТИ .....	30
7.4. ДЕТЕКЦИЈА КОМУНА .....	30
<b>8. ПОРЕЂЕЊЕ SNET И SNETT .....</b>	<b>32</b>

# 1. Увод

Овај документ односи се на израду пројектног задатка на предмету Анализа социјалних мрежа. У оквиру пројектног задатка анализирани су подаци са друштвене мреже *Reddit* из периода 2008. године. Анализиране су објаве корисника (означене са *submission*) и коментари (означени са *comment*).

Анализа података је вршена уз помоћ програмског језика Python и његових библиотека. Језик је изабран због своје велике заступљености и лакоће коришћења приликом обраде великог сета података. Обрада примарног скупа података је обављена коришћењем библиотеке *pandas*. За визуелизацију мреже је коришћен алат *Gephi*.

У поглављу број два се говори о чишћењу података и манипулацијом над подацима како би се добио секундарни скуп података који се касније користи приликом анализе и креирања социјалних мрежа.

Поглавље број три говори о статичкој обради података и даје одговоре на нека од питања која су постављена приликом анализе.

Поглавља број четири, пет, шест и седам говоре о мрежама *SNet*, *SNetT*, *SNetF*, *UserNet* респективно. У поглављима се могу прочитати детаљније информације о одговарајућој мрежи, као и њеним карактеристикама и закључцима који су донети на основу те анализе.

Поглавље број осам се бави упоређивањем мреже *SNet* и *SNetT*.

## 2. ЧИШЋЕЊЕ ПОДАТАКА

У овом поглављу је описан процес обраде примарних података који су покупљени са сајта *Reddit*. Подаци су распоређени у два фолдера. У једном фолдеру се налазе подаци који се односе на објаве, док се у другом фолдеру налазе подаци о коментарима корисника. Сваки фолдер садржи 12 фајлова у CSV формату, а сваки од фајлова се односи на један месец у 2008. години.

Спајањем фајлова о објавама (уз помоћ библиотеке *pandas*) креира се једна табела. Исти процес је потребно одрадiti и са подацима о коментарима. У табеле улазе само подаци који су од значаја за анализу, а што је дефинисано пројектним задатком. Из табеле о објавама уклоњени су подаци који се налазе у колонама *distinguished*, *domain*, *stickied*, *locked*, *hide\_score*. Из табеле о коментарима уклоњене су следеће колоне: *score*, *distinguished*, *gilded*, *controversiality*.

Над подацима су извршене и одређене провере да би се утврдиле евентуалне грешке које би могле да утичу на анализу. Овако очишћени подаци су након извршених провера сачувани на диску (у оквиру пројектног задатка на путањи *data/cleaned/*). У наставку ће ови подаци бити коришћени за креирање и анализу више различитих социјалних мрежа, као и за статичку обраду података.

### 3. СТАТИЧКА ОБРАДА ПОДАТАКА

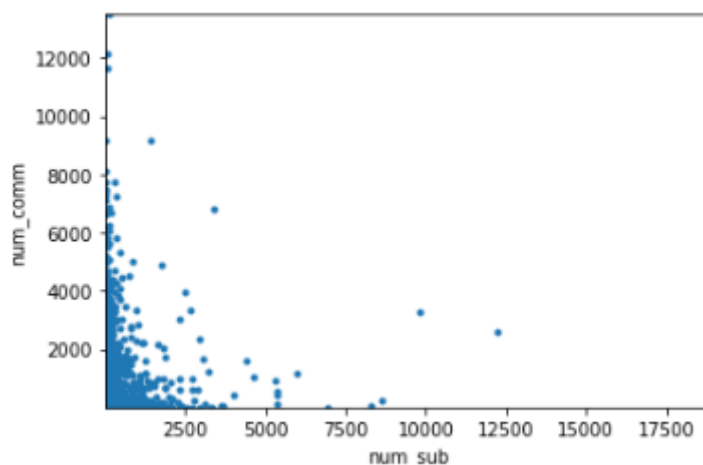
У овом поглављу је извршена статичка анализа података над секундарним (очишћеним) скупом података. Обрада је извршена да би се добио почетни увид у то са каквим подацима радимо, и да би се стекао осећај о томе какви су подаци.

Број различитих субредита је **4359**. Овај податак ће приликом креирања прве мреже означавати број чворова у мрежи. Сваки од субредита ће представљати посебан чвор у мрежи. Најважнији субредити по броју корисника су: reddit.com (163779 корисника), politics (38374), pics (29753), technology (28337), funny (28186). Са највећим бројем коментара издвајају се следећи субредити: reddit.com (1143184), politics (801396) programming (345997), pics (286192), science (238291).

Просечан број активних корисника износи 148.6314. Корисници са највећим бројем објава су (дати у формату корисничко име (број објава)): gst (18870), qgyh2 (12238), rmuser (9822), twolf1 (8597), IAmperfectlyCalm (8308). Корисници са највећим бројем коментара (дати у формату корисничко име (број коментара)) су: NoMoreNicksLeft (13480), malcontent (12159), matts2 (11672), mexicodoug (9169), 7oby (9161).

Корисници који су активни на највећем броју субредита коментара (дати у формату корисничко име (број активности)) су: MrKlaatu (181), Escafane (154), omfgninja (122), codepoet (111), scientologist2 (111).

Пирсонов коефицијент колерације броја објава и броја коментара износи 0.146.



Слика 1 График броја објава и броја коментара

На Х оси су приказани бројеви објава, док су на Y оси приказани бројеви коментара.

Објаве са највећим бројем коментара су:

- **Got\_six\_weeks\_try\_the\_hundred\_push\_ups\_training** (објава која се бави тренирањем у теретани, линк до објаве: <http://hundredpushups.com/>)
- **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?** (говори о држави у којој су живели људи црне расе без текуће воде. Линк: <http://www.time.com/time/nation/article/0,8599,1822455,00.html>)
- **i\_am\_constantly\_wondering\_what\_everyone\_on\_here\_looks\_like.Post\_your\_pictures!** (објава се бави изгледом других корисника на друштвеној мрежи *Reddit*)

## 4. SUBREDDIT NETWORK (SNET)

У овом поглављу описана је мрежа *SNet*. Такође поред тога су изложени и резултати спроведених анализа мреже.

### 4.1. Креирање мреже

Мрежа је креирана коришћењем секундарног скупа података уз помоћ библиотеке *network* која је подржана у оквиру програмског језика *Python*.

Мрежу карактерише **4359** различитих чворова. Представљена је нетежинским графом. Сваки чвор унутар мреже је један сабредит, док се гране између чворова успостављају уколико је један корисник био активан на оба та сабредита. Тежине грана представљају број корисника који су били активни на оба сабредита које повезује дата грана. Након креирања мреже можемо приметити да је успостављено **156959** грана између чворова.

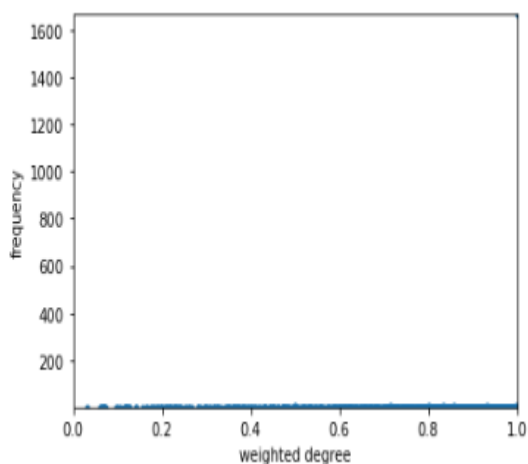
### 4.2. Израчунавање метрика мреже

Метрике мреже се могу израчунати на више начина. Неке израчунавања су подржана кроз алат *Gephi* након учитавања потребног фајла са одговарајућом екстензијом (*.gml* у овом случају). У оквиру библиотеке *networkX* подржана су израчунавања свих потребних параметара за анализу мреже.

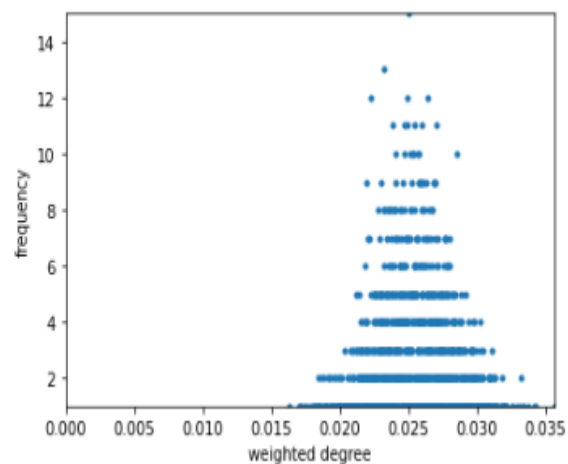
Густина мреже износи 0.025. Дијаметар мреже износи 5, док је просечна дужина путање у мрежи 2.099.

Број повезаних компоненти износи 814. Највећа повезана компонента је величине 3524, стога се може закључити да постоји гигантска компонента. Гигантска компонента се издваја из графа, и над њом се врше све потребне анализе мреже. На диску је сачувана под именом *sent\_giant.gml*.

Глобални коефицијент кластеризације износи 0.26730, док просечни ниво кластеризације износи 0.77054.



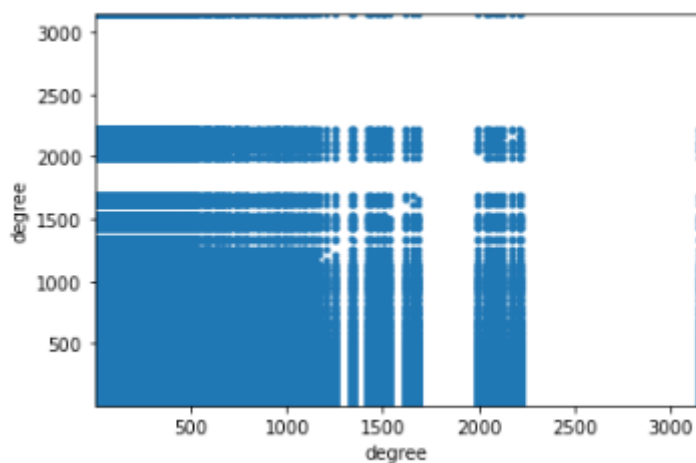
Слика 2 Расподела коефицијента кластеризације (*SNet*)



Слика 3 Расподела коефицијента кластеризације (*Erdos-Renyi*)

Са графика се може закључити да је мрежа са леве стране доста више кластерисана у поређењу са *Erdos-Renyi* мрежом приказаном десно. Због високог степена кластеризације и мале просечне дужине пута можемо закључити да мрежа изражава особине малог света.

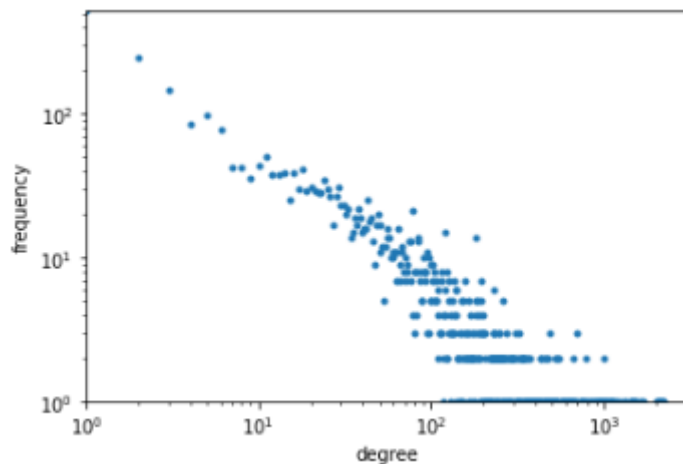
Коефицијент асортативности износи  $-0.43670$ . Асортативно мешање није изражено.



Слика 4 Асортативност мреже (*SNet*)



Поређење расподела је извршено коришћењем неколико различитих функција. Коришћене су power law расподеле и truncated power law расподеле. Због мање  $p$  вредности код truncated power law расподеле закључујемо да мрежа много боље прати ту расподелу.



Слика 5 График расподеле степена чвора (*SNet*)

Код неусмереног графа нема смисла одређивати вредности хабова и ауторитета у мрежи. Резултати за хабове и ауторитете у неусмереној мрежи ће бити идентични за сваки чвор. Има смисла одређивати те вредности за усмерене мреже.

### 4.3. Израчунавање централности

Анализа централности нам може дати више информација о потенцијалним улогама одређених чворова у мрежи. Постоји више врста централности. У овом раду су обрађене следеће: централност по степену, централност по блискости, релациона централност и централност по сопственом вектору.

Након израчунавања централности по степену (*DC*), чворови који су се издвојили као најважнији су: *reddit.com*, *technology*, *politics*, *pics*, *funny*.

Након израчунавања централности по блискости (*CC*), чворови који су се издвојили као најважнији су: *reddit.com*, *technology*, *politics*, *pics*, *funny*.

Након израчунавања релационе централности (*BC*), чворови који су се издвојили као најважнији су: *reddit.com*, *technology*, *programming*, *politics*, *business*.

Након израчунавања централности по сопственом вектору (*EVC*), чворови који су се издвојили као најважнији су: *reddit.com*, *technology*, *politics*, *pics*, *science*. Ова централност нам

говори меру утицаја чвора. Чворови са већом централношћу по вектору су утицајнији у мрежи од осталих.

Након израчунавања Кацове скале ( $KC$ ), најважнији чворови су: *reddit.com*, *politics*, *pics*, *funny*, *science*.

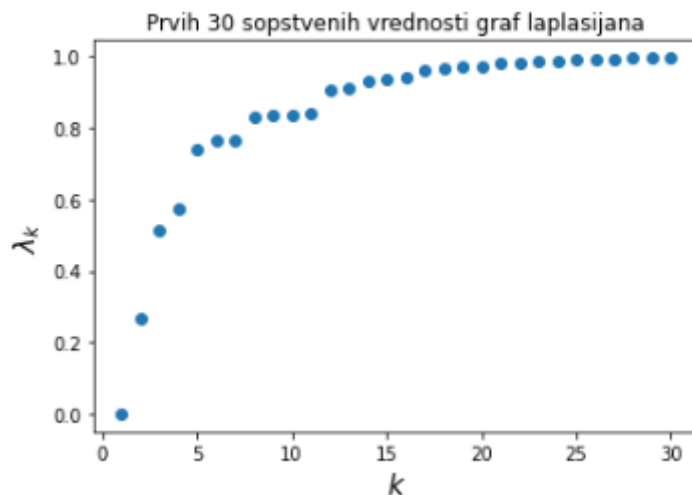
Композитни ранк је креиран следећом формулом:

$$rank = DC * CC * BC * EVC * KC$$

Најважнији актери по композитном ранку су: *reddit.com*, *technology*, *politics*, *pics*, *funny*.

#### 4.4. Детекција комуна

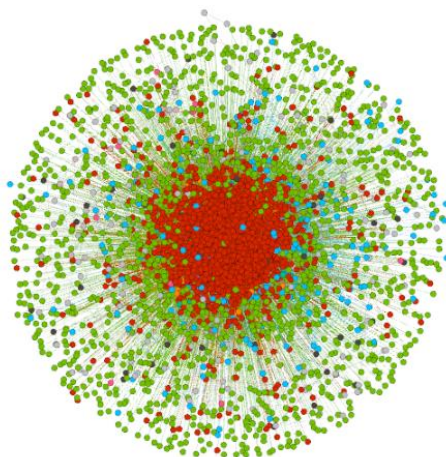
Спектарном анализом можемо одредити који су потенцијални кандидати за комуне у мрежи. То се најбоље види са сопствених вредности где је приказана подела на кластере у распону од 2 до 30 кластера.



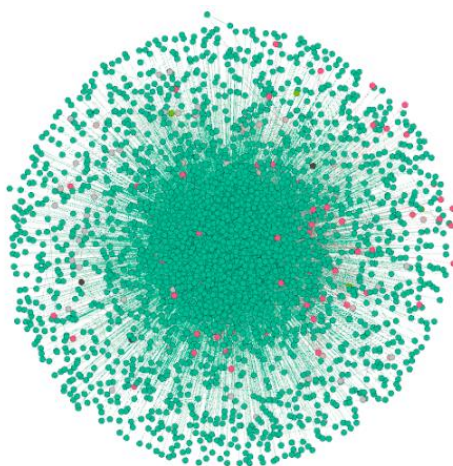
Слика 6 Првих 30 сопствених вредности лапласијана за (*SNet*)

Сваки скок на графу значи да би на том делу могли да направимо пресек кластера. Са графа се види да су добри потенцијални кандидати за комуне у мрежи на 2, 3, 5, 8, 12.

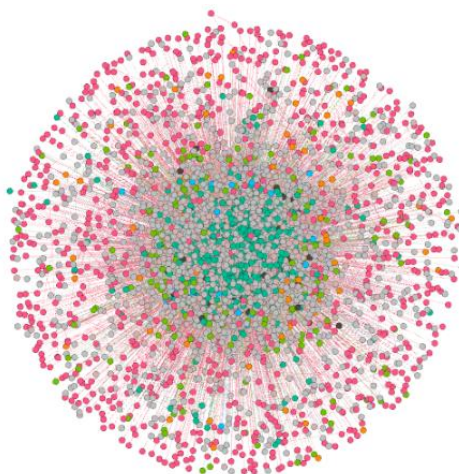
Детекција комуна је извршена коришћењем алата Gerhi. Алат користи Лувенску методу за детекцију комуна приближне сложености која је реда величине  $O(n \log n)$ .



Слика 7 Лувенска метода (*SNet*). Резолуција 1.0. Број комуна 55  
8



Слика 9 Лувенска метода (*SNet*). Резолуција 2.0. Број комуна 15

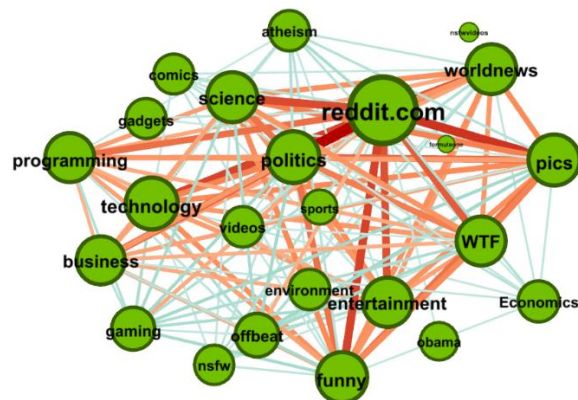


Слика 10 Лувенска метода (*SNet*). Резолуција 0.5. Број комуна 59

Неки од најважнијих чворова комуна су дати на следећим сликама:



Слика 11 Лувенска метода (*SNet*). Резолуција 1.0. Најважнији чворови највеће комуне



Слика 12 Лувенска метода (*SNet*). Резолуција 1.0. Најважнији чворови друге највеће комуне

## 5. SUBREDDIT NETWORK (SNETT)

У овом поглављу описана је мрежа *SNetT*. Такође поред тога су изложени и резултати спроведених анализа мреже.

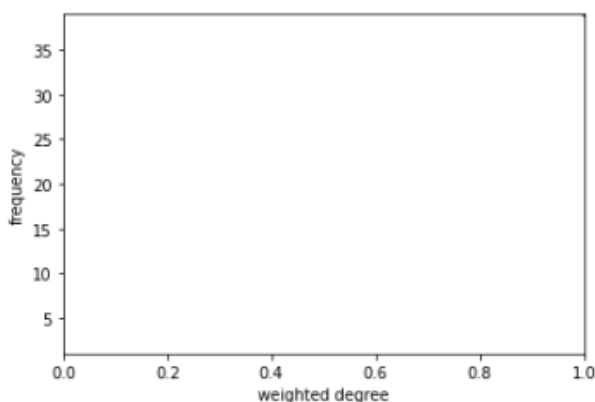
### 5.1. Креирање мреже

Мрежу карактерише **39** различитих чворова. Представљена је нетежинским графом. Сваки чвор унутар мреже је један сабредит, док се гране између чворова успостављају уколико је један корисник био активан на оба та сабредита. Мрежа је креирана као подфраф мреже *Sent*, тако што су изабрани чворови који су од значаја за анализу. Чворови мреже су: (*reddit.com*, *pics*, *worldnews*, *programming*, *business*, *politics*, *obama*, *science*, *technology*, *WTF*, *AskReddit*, *netsec*, *philosophy*, *videos*, *offbeat*, *funny*, *entertainment*, *linux*, *geek*, *gaming*, *comics*, *gadgets*, *nsfw*, *news*, *environment*, *atheism*, *canada*, *math*, *Economics*, *scifi*, *bestof*, *cogsci*, *joel*, *Health*, *guns*, *photography*, *software*, *history*, *ideas*) Након креирања мреже можемо приметити да је успостављено **741** грана између чворова.

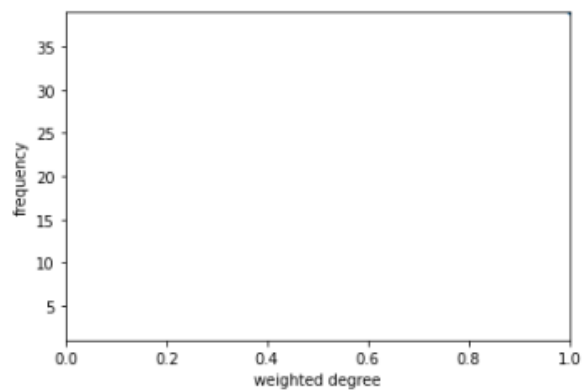
### 5.2. Израчунавање метрика мреже

Густина мреже износи 1. Дијаметар мреже износи 1. док је просечна дужина путање у мрежи 1. Број повезаних компоненти износи 1. Највећа повезана компонента је величине 39, стога се може закључити да постоји гигантска компонента.

Глобални коефицијент кластеризације износи 1, док просечни ниво кластеризације износи 1. Мрежа је потпуно повезана, па је због тога коефицијент кластеризације 1.



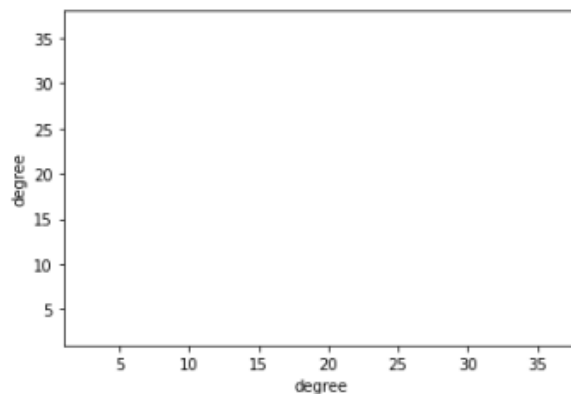
Слика 13 Расподела коефицијента  
Кластеризације (*SNetT*)



**Слика 14** Расподела коефицијента кластеризације (Erdos-Renyi)

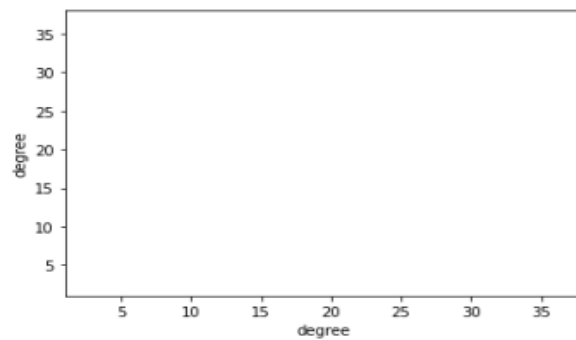
Због високог степена кластеризације и мале просечне дужине пута можемо закључити да мрежа изражава особине малог света.

Коефицијент асортативности није могуће одредити због потпуне повезаности мреже.



**Слика 15** Асортативност мреже (*SNetT*)

Мрежа не изражава феномен клуба богатих јер су сви чворови међусобно повезани. Са графика се може видети да мрежа не прати ниједну расподелу.



Слика 16 График расподеле степена чвора (*SNetT*)

Код неусмереног графа нема смисла одређивати вредности хабова и ауторитета у мрежи. Резултати за хабове и ауторитете у неусмереној мрежи ће бити идентични за сваки чвор. Има смисла одређивати те вредности за усмерене мреже.

### 5.3. Израчунавање централности

Након израчунавања централности по степену (*DC*), чворови који су се издвојили као најважнији су: *comics, funny, guns, joel, Economics*.

Након израчунавања централности по блискости (*CC*), чворови који су се издвојили као најважнији су: *comics, funny, guns, joel, Economics*.

Након израчунавања релационе централности (*BC*), чворови који су се издвојили као најважнији су: *comics, funny, guns, joel, Economics*..

Након израчунавања централности по сопственом вектору (*EVC*), чворови који су се издвојили као најважнији су: *reddit.com, politics, pics, funny, science*. Ова централност нам говори меру утицаја чвора. Чворови са већом централношћу по вектору су утицајнији у мрежи од осталих.

Након израчунавања Кацове скале (*KC*), најважнији чворови су: *reddit.com, politics, pics, funny, science*.

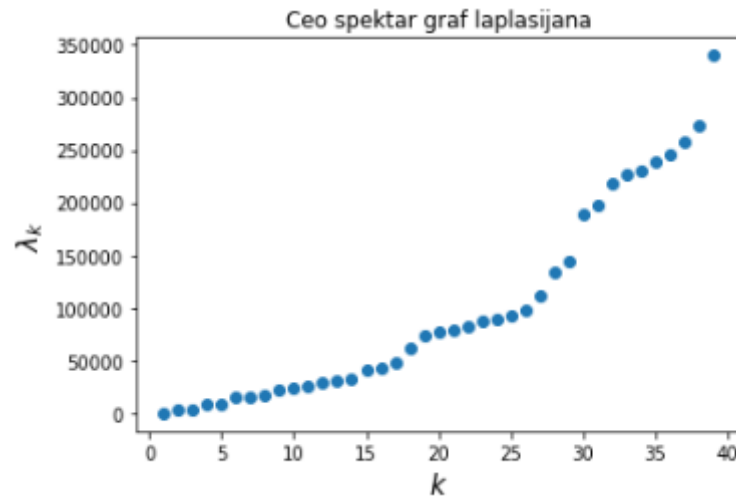
Композитни ранк је креиран следећом формулом:

$$rank = DC * CC * BC * EVC * KC$$

Најважнији актери по композитном ранку су: *reddit.com, politics, pics, funny, science*.

## 5.4. Детекција комуна

Спектарном анализом можемо одредити који су потенцијални кандидати за комуне у мрежи. То се најбоље види са сопствених вредности где је приказана подела на кластере у распону од 2 до 30 кластера.

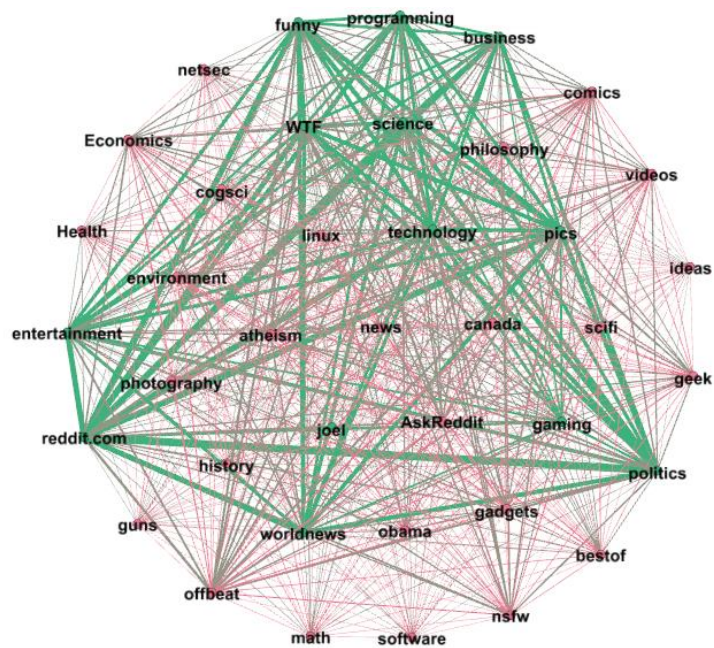


Слика 17 Првих 30 сопствених вредности лапласијана за (SNetT)

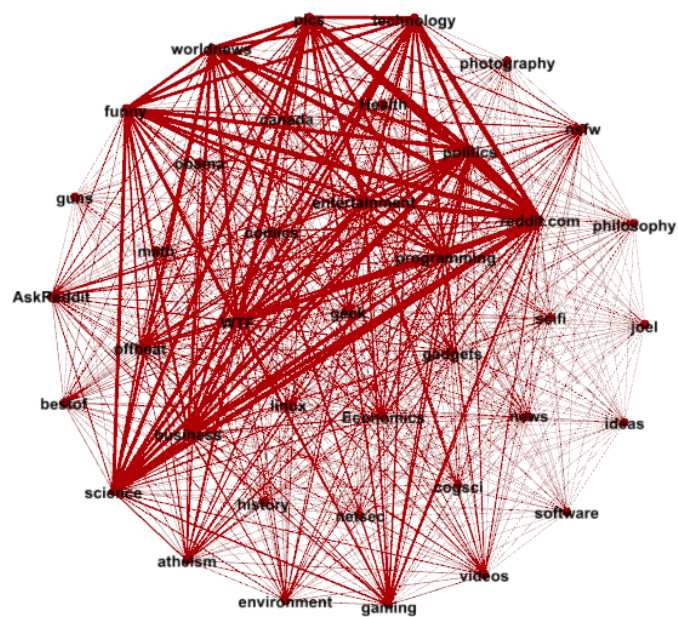
Сваки скок на графу значи да би на том делу могли да направимо пресек кластера. Са графа се види да су добри потенцијални кандидати за комуне у мрежи на 2, 4, 9, 18, 28, 30, 39.

Детекција комуна је извршена коришћењем алата Gephi. Алат користи Лувенску методу за детекцију комуна приближне сложености која је реда величине  $O(n \log n)$ .



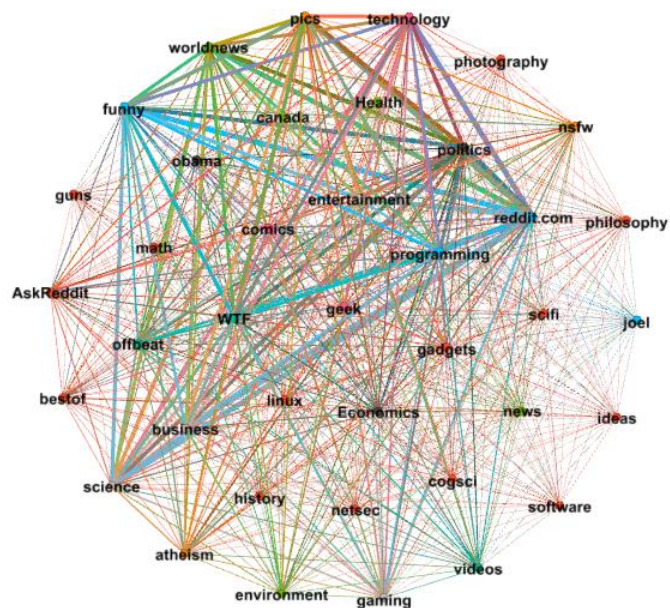


Слика 18 Лувенска метода (*SNetT*). Резолуција 1.0. Број комуна 2



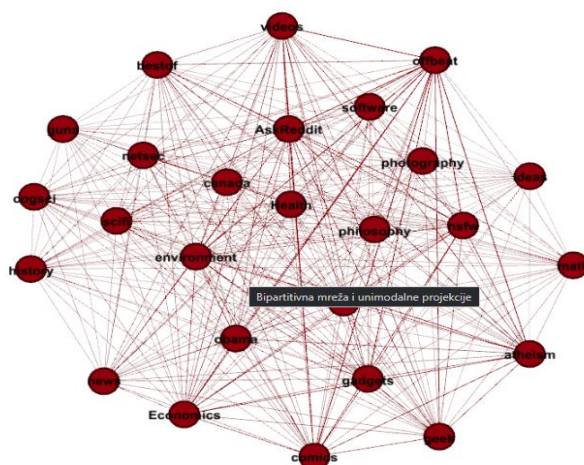
Слика 19 Лувенска метода (*SNetT*). Резолуција 2.0. Број комуна 1

За резолуцију 0.5 детектовано је 9 различитих комуна.

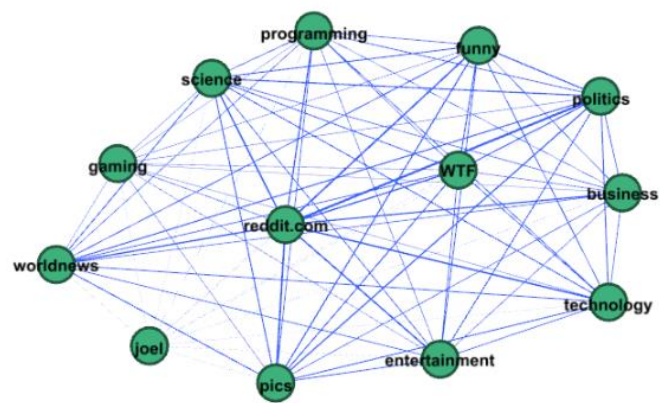


Слика 20 Лувенска метода (*SNetT*). Резолуција 0.5 Број комуна 9

Неки од најважнијих чворова комуна су дати на следећим сликама:



Слика 21 Лувенска метода (*SNetT*). Резолуција 1.0. Најважнији чворови највеће комуне



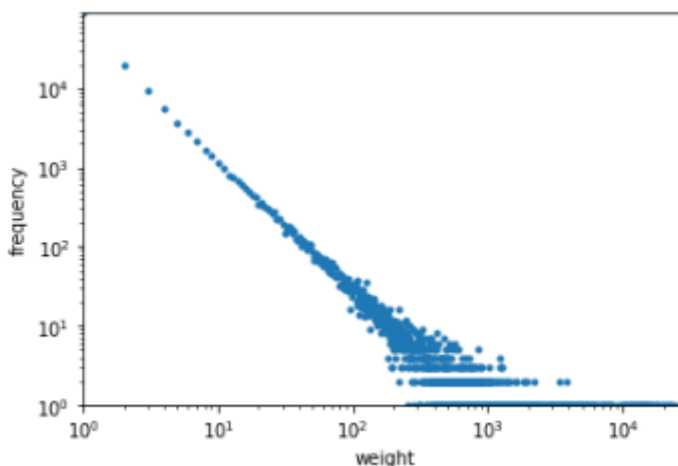
Слика 22 Лувенска метода (*SNetT*). Резолуција 1.0. Најважнији чворови друге највеће комуне

## 6. SUBREDDIT NETWORK (SNetF)

У овом поглављу описана је мрежа *SNetF*. Такође поред тога су изложени и резултати спроведених анализа мреже.

### 6.1. Креирање мреже

Мрежу карактерише **4359** различитих чворова. Представљена је нетежинским графом. Сваки чвор унутар мреже је један сабредит. Мрежа је креирана као одбацивањем грана са малом тежином мреже *SNet*.



Слика 23 Расподеле тежине грана (*SNetF*)

Можемо приметити да велика већина чворова садржи гране које су мале тежине. Да бисмо добили што боље податке, одбацићемо гране са малом тежином и задржати само гране које имају тежину већу од неке задате вредности. То се може представити следећом формулом.

$$\text{Одбацивање грана: } \begin{cases} w \leq T, & \text{грана се брише из графа} \\ w > T, & \text{грана остаје у графу} \end{cases}$$

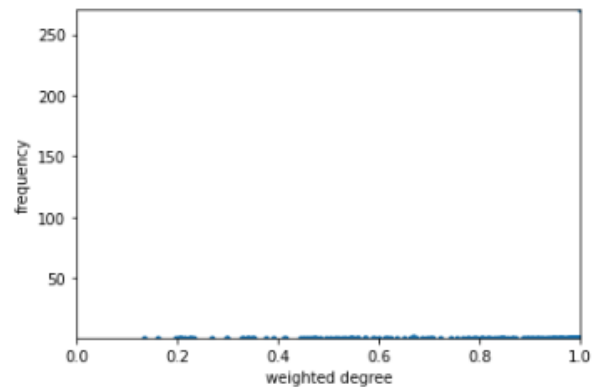
Уз помоћ графика је одлучено да се за *w\_threshold* је изабере вредност 10. Све гране које имају тежину мању од изабране вредности се уклањају из мреже. Након уклањања грана добијена је финална верзија графа која има **18640** грана.

### 6.2. Израчунавање метрика мреже

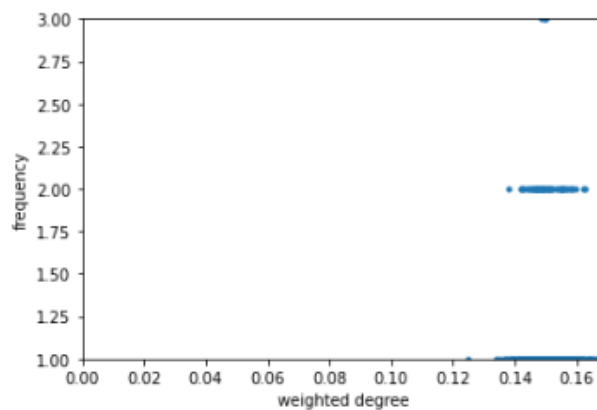
Густина мреже износи 0,151. Дијаметар мреже износи 4. док је просечна дужина путање у мрежи 1.894.

Број повезаних компоненти износи 3863. Највећа повезана компонента је величине 497, стога се може закључити да постоји гигантска компонента.

Глобални коефицијент кластеризације износи 0,49724, док просечни ниво кластеризације износи 0.84913.



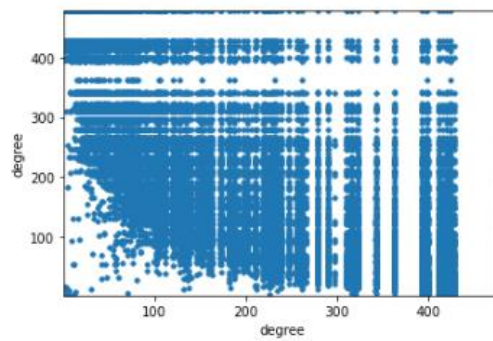
Слика 24 Расподеле коефицијента кластеризације (*SNetF*)



Слика 25 Расподела коефицијента кластеризације (Erdos-Renyi)

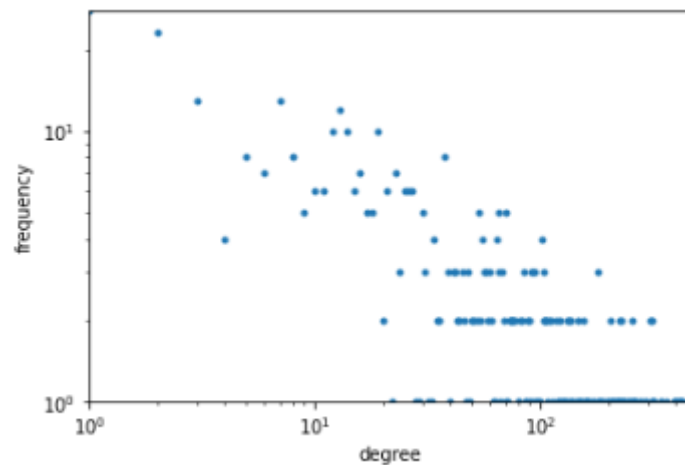
Због високог степена кластеризације и мале просечне дужине пута можемо закључити да мрежа изражава особине малог света.

Коефицијент асортативности износи -0.61150. Асортативно мешање није изражено.



Слика 26 Асортативност мреже (*SNetF*)

Мрежа не изражава феномен клуба богатих јер су сви чворови међусобно повезани.



Слика 27 График расподеле степена чвора (*SNetF*)

Код неусмереног графа нема смисла одређивати вредности хабова и ауторитета у мрежи. Резултати за хабове и ауторитете у неусмереној мрежи ће бити идентични за сваки чвор. Има смисла одређивати те вредности за усмерене мреже.

### 6.3. Израчунавање централности

Након израчунавања централности по степену (*DC*), чворови који су се издвојили као најважнији су: *reddit.com*, *politics*, *pics*, *science*, *technology*.

Након израчунавања централности по блискости (*CC*), чворови који су се издвојили као најважнији су: *reddit.com*, *politics*, *pics*, *science*, *technology*.

Након израчунавања релационе централности (*BC*), чворови који су се издвојили као најважнији су: *reddit.com*, *politics*, *pics*, *science*, *technology*.



Након израчунавања централности по сопственом вектору (*EVC*), чворови који су се издвојили као најважнији су: *reddit.com*, *politics*, *pics*, *science*, *technology*. Ова централност нам говори меру утицаја чвора. Чворови са већом централношћу по вектору су утицајнији у мрежи од осталих.

Након израчунавања Кацове скале (*KC*), најважнији чворови су: *reddit.com*, *politics*, *pics*, *funny*, *science*.

Композитни ранк је креиран следећом формулом:

$$rank = DC * CC * BC * EVC * KC$$

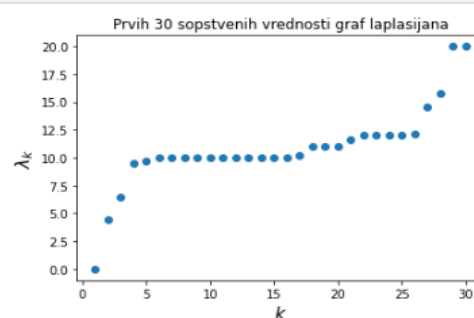
Најважнији актери по композитном ранку су: *reddit.com*, *politics*, *pics*, *science*, *technology*.

## 6.4. Детекција комуна

Спектарном анализом можемо одредити који су потенцијални кандидати за комуне у мрежи. То се најбоље види са сопствених вредности где је приказана подела на кластере у распону од 2 до 30 кластера.



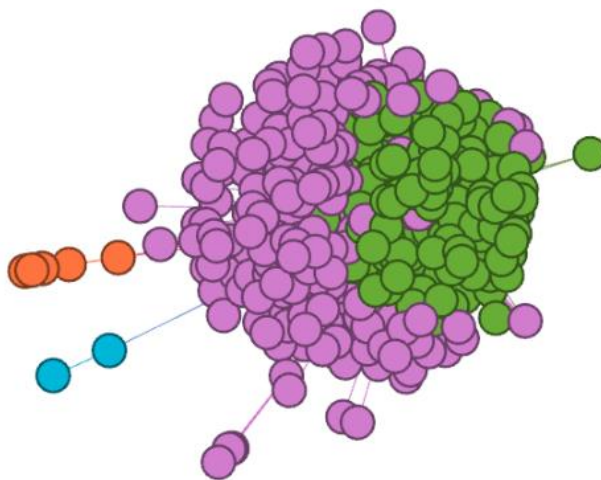
Слика 28 График целог спектра графа (*SNetF*)



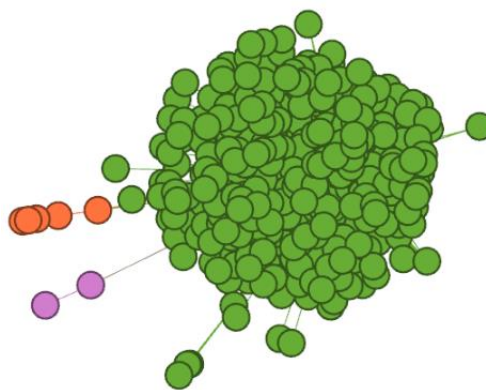
Слика 29 Првих 30 сопствених вредности лапласијана за (*SNetF*)

Сваки скок на графу значи да би на том делу могли да направимо пресек кластера. Са графа се види да су добри потенцијални кандидати за комуне у мрежи на 2, 3, 4, 6, 18, 27, 29.

Детекција комуна је извршена коришћењем алата Gephi. Алат користи Лувенску методу за детекцију комуна приближне сложености која је реда величине  $O(n \log n)$ .

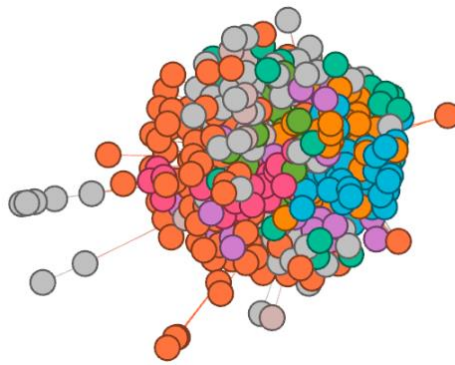


Слика 30 Лувенска метода (*SNetF*). Резолуција 1.0. Број комуна 4



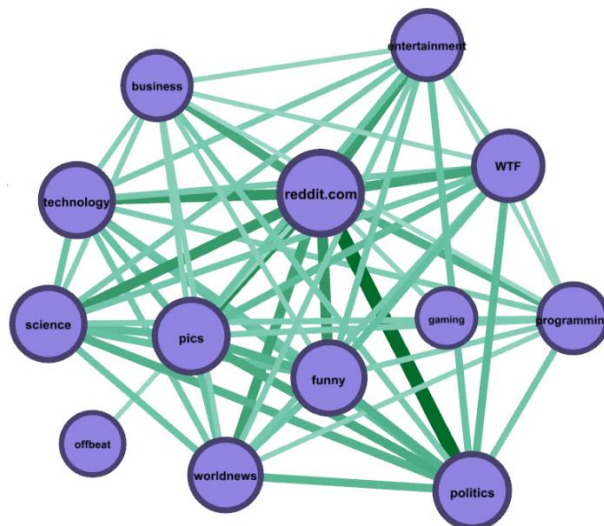
Слика 31 Лувенска метода (*SNetF*). Резолуција 2.0. Број комуна 3



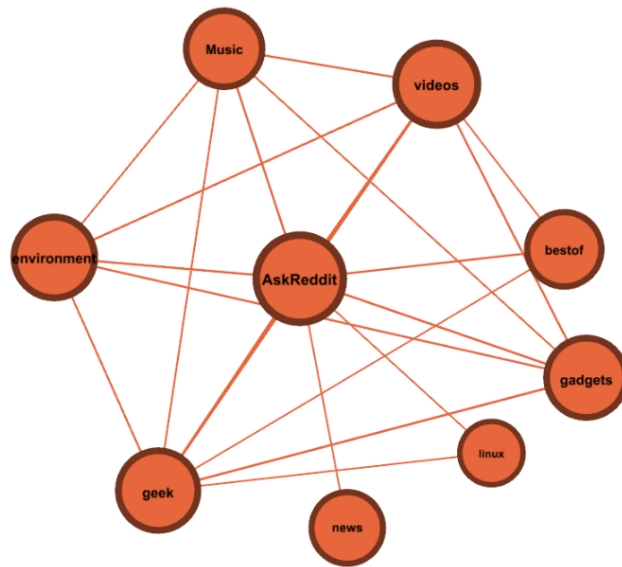


Слика 32 Лувенска метода (*SNetF*). Резолуција 0.5. Број комуна 99

Неки од најважнијих чворова комуна су дати на следећим сликама:



Слика 33 Лувенска метода (*SNetF*). Резолуција 1.0. Најважнији чворови највеће комуне



Слика 34 Лувенска метода (*SNetF*) . Резолуција 1.0. Најважнији чворови друге највеће комуне

## 7. USERNET

У овом поглављу описана је мрежа *UserNet*. Такође поред тога су изложени и резултати спроведених анализа мреже.

### 7.1. Креирање мреже

Мрежа је креирана на основу интеракције корисника у оквиру друштвене мреже Reddit. Чворови мреже представљају кориснике у мрежи, док су гране мреже засноване на врсти комуникације између корисника. Уколико је један корисника коментарисао објаву другог корисника, смер гране је од корисника који коментарише ка креатору коментара. Уколико има више коментара од стране истог корисника, то се представља одговарајућом тежином гране. Може се закључити да је изабрана усмерена тежинска мрежа за моделовање мреже.

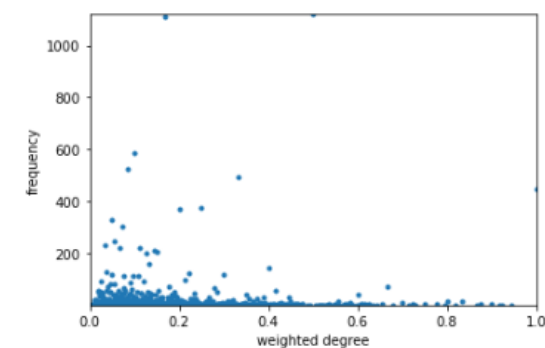
Број чворова у овако креираној мрежи за посматрани период је 80714. Број грана мреже износи 2803295.

### 7.2. Израчунавање метрика мреже

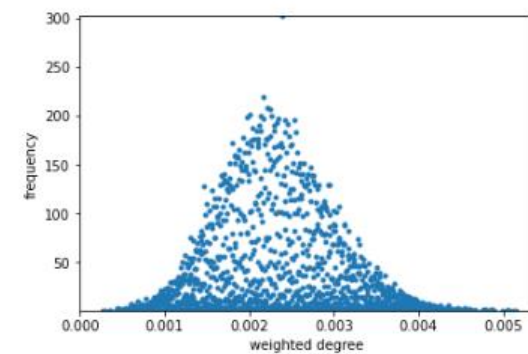
Густина мреже износи 0.001147. Дијаметар мреже износи 13. док је просечна дужина путање у мрежи 3.145.

Број повезаних компоненти износи 32116. Највећа повезана компонента је величине 48502, стога се може закључити да постоји гигантска компонента.

Глобални коефицијент кластеризације износи 0.071746 , док просечни ниво кластеризације износи 0.11697.



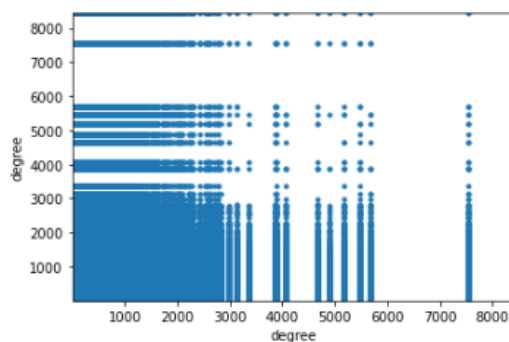
Слика 35 Расподеле коефицијента кластеризације (*UserNet*)



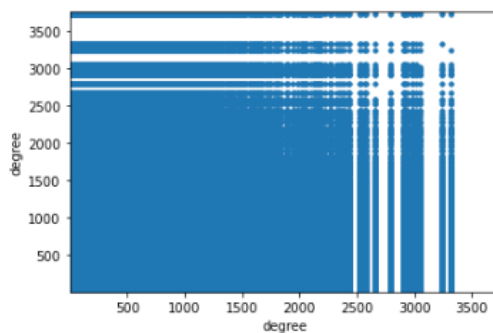
Слика 36 Расподеле коефицијента кластеризације (*UserNet*) Erdos-Renyi

Због високог степена кластеризације и мале просечне дужине пута можемо закључити да мрежа изражава особине малог света.

Коефицијент асортативности се код усмерене мреже одређује и за улазни, и за излазни степен. Коефицијент асортативности улазног степена износи -0.072866. Коефицијент асортативности излазног степена износи -0.028439.



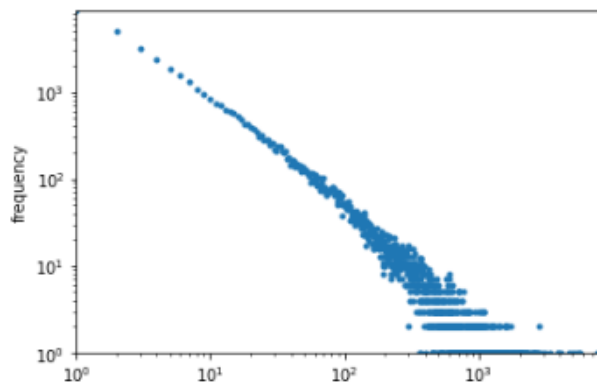
Слика 37 Асортативност мреже за улазни степен (*UserNet*)



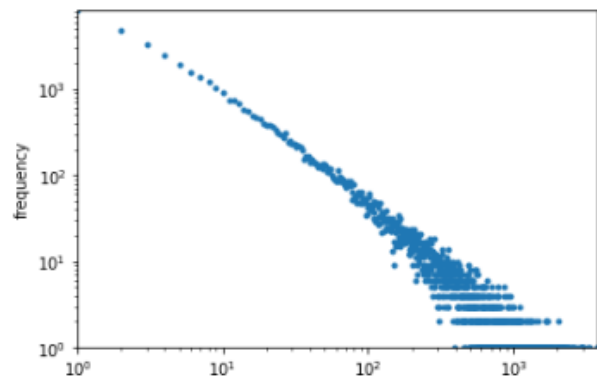
Слика 38 Асортативност мреже за излазни степен (*UserNet*)

Мрежа изражава феномен клуба богатих.

Поређење расподела је извршено коришћењем неколико различитих функција. Расподела је извршена и по улазном, и по излазном чвору. Коришћене су *power law* расподеле и *truncated power law* расподеле. Због мање  $p$  вредности код *truncated power law* расподеле закључујемо да мрежа много боље прати ту расподелу.



Слика 39 График расподеле улазног степена чвора



Слика 40 График расподеле излазног степена чвора

Код усмереног графа има смисла одређивати вредности хабова и ауторитета у мрежи. Резултати за хабове и ауторитете у усмереној мрежи неће бити идентичне за сваки чвор. Најважнији ауторитети у мрежи су: *qgyh2*, *rmuser*, *matts2*, *malcontent*, *allie*.

Најважнији хабови у мрежи су: *malcontent*, *matts2*, *NoMoreNicksLeft*, *glengyron*, *aletoledo*.

### 7.3. Израчунавање централности

Након израчунавања централности по улазном степену (*IDC*), чворови који су се издвојили као најважнији су: *rmuser*, *qgyh2*, *noname99*, *alllie*, *swampsparrow*.

Након израчунавања централности по излазном степену (*ODC*), чворови који су се издвојили као најважнији су: *NoMoreNicksLeft*, *mexicodoug*, *Toby*, *h0dg3s*, *randomb0y*.

Након израчунавања релационе централности (*BC*), чворови који су се издвојили као најважнији су: *alllie*, *Toby*, *NoMoreNicksLeft*, *malcontent*, *rmuser*.

Кацова скала се за улазне мреже рачуна и по улазном, и по излазном фактору. Након израчунавања улазне кацове централности (*IKC*), чворови који су се издвојили као најважнији су: *rmuser*, *qgyh2*, *alllie*, *gst*, *noname99*, *malcontent*. Након израчунавања излазне кацове централности (*OKC*), чворови који су се издвојили као најважнији су: *malcontent*, *matts2*, *NoMoreNicksLeft*, *mexicodoug*, *aletoledo*.

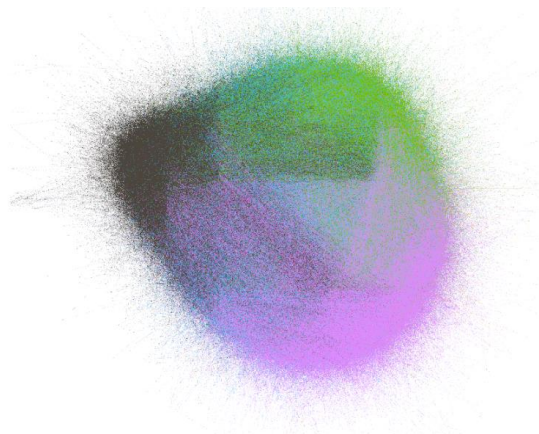
Композитни ранк је креиран следећом формулом:

$$\mathbf{rank} = IDC * ODC * BC * IKC * OKC$$

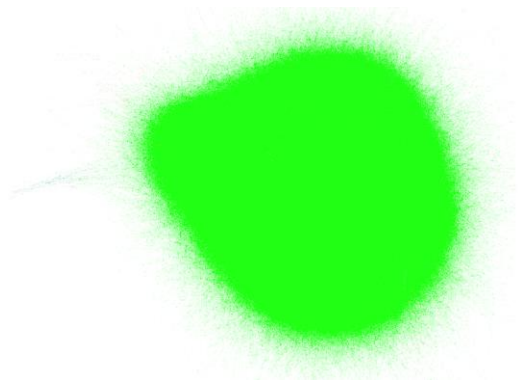
Најважнији актери по композитном ранку су: *malcontent*, *matts2*, *otakucode*, *deuteros*, *glengyron*.

### 7.4. Детекција комуна

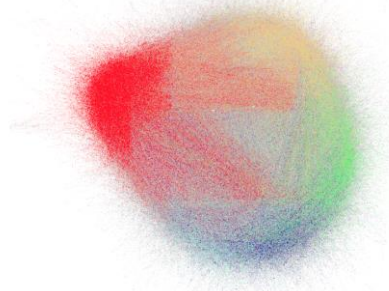
Усмерене мреже не подржавају спектарну анализу па је тако није могуће применити на задати скуп података.



**Слика 41** Лувенска метода (*UserNet*). Резолуција 1.0. Број комуна 41



**Слика 42** Лувенска метода (*UserNet*). Резолуција 2.0. Број комуна 27



**Слика 43** Лувенска метода (*UserNet*). Резолуција 2.0. Број комуна 143

Због величине мреже није било могуће визуелно приказати кључне чворове у мрежи. Те податке можемо претпоставити на основу високе релационе централности јер се чворови са већом вредношћу релационе централности налазе на већем броју најкраћих путева у мрежи.

## 8. ПОРЕЂЕЊЕ *SNet* И *SNetT*

*SNetT* је комплетно повезана мрежа. Због тога се ове две мреже разликују по својим метрикама.

	<i>SNet</i>	<i>SNetT</i>
Број чворова	4359	39
Број грана	156959	741
Густина	0.025	1
Дијаметар	5	1
Просечна дистанца	2.099	1

Табела 1 Метрике мрежа *SNet* и *SNetT*

Због мање ексцентричности чворова мреже *SNetT* можемо закључити да су чворови у језгру мреже *SNet*.

Након израчунавања активности можемо закључити да су корисници мреже *SNetT* доста активнији у мрежи.



## 9. СЛИКЕ

Слика 1 График броја објава и броја коментара .....	5
Слика 2 Расподела коефицијента кластеризације ( <i>SNet</i> ) .....	8
Слика 3 Расподела коефицијента кластеризације (Erdos-Renyi).....	8
Слика 4 Асортативност мреже ( <i>SNet</i> ) .....	8
Слика 5 График расподеле степена чвора ( <i>SNet</i> ) .....	9
Слика 6 Првих 30 сопствених вредности лапласијана за ( <i>SNet</i> ) .....	10
Слика 7 Лувенска метода ( <i>SNet</i> ). Резолуција 1.0. Број комуна 55 .....	11
8.....	11
Слика 9 Лувенска метода ( <i>SNet</i> ). Резолуција 2.0. Број комуна 15 .....	11
Слика 10 Лувенска метода ( <i>SNet</i> ). Резолуција 0.5. Број комуна 59 .....	11
Слика 11 Лувенска метода ( <i>SNet</i> ). Резолуција 1.0. Најважнији чворови највеће комуне.....	12
Слика 12 Лувенска метода ( <i>SNet</i> ). Резолуција 1.0. Најважнији чворови друге највеће комуне .....	12
Слика 13 Расподела коефицијента .....	13
Слика 14 Расподела коефицијента кластеризације (Erdos-Renyi).....	14
Слика 15 Асортативност мреже ( <i>SNetT</i> ) .....	14
Слика 16 График расподеле степена чвора ( <i>SNetT</i> ).....	15
Слика 17 Првих 30 сопствених вредности лапласијана за ( <i>SNetT</i> ) .....	16
Слика 18 Лувенска метода ( <i>SNetT</i> ). Резолуција 1.0. Број комуна 2 .....	17
Слика 19 Лувенска метода ( <i>SNetT</i> ). Резолуција 2.0. Број комуна 1 .....	17
Слика 20 Лувенска метода ( <i>SNetT</i> ). Резолуција 0.5 Број комуна 9 .....	18
Слика 21 Лувенска метода ( <i>SNetT</i> ). Резолуција 1.0. Најважнији чворови највеће комуне.....	18
Слика 22 Лувенска метода ( <i>SNetT</i> ). Резолуција 1.0. Најважнији чворови друге највеће комуне .....	19
Слика 23 Расподеле тежине грана ( <i>SNetF</i> ) .....	20
Слика 24 Расподеле коефицијента кластеризације ( <i>SNetF</i> ) .....	21
Слика 25 Расподела коефицијента кластеризације (Erdos-Renyi).....	21
Слика 26 Асортативност мреже ( <i>SNetF</i> ) .....	22
Слика 27 График расподеле степена чвора ( <i>SNetF</i> ).....	22
Слика 28 График целог спектра графа ( <i>SNetF</i> ) .....	23
Слика 29 Првих 30 сопствених.....	23
Слика 30 Лувенска метода ( <i>SNetF</i> ). Резолуција 1.0. Број комуна 4 .....	24
Слика 31 Лувенска метода ( <i>SNetF</i> ). Резолуција 2.0. Број комуна 3 .....	24
Слика 32 Лувенска метода ( <i>SNetF</i> ). Резолуција 0.5. Број комуна 99 .....	25
Слика 33 Лувенска метода ( <i>SNetF</i> ). Резолуција 1.0. Најважнији чворови највеће комуне .....	25
Слика 34 Лувенска метода ( <i>SNetF</i> ) . Резолуција 1.0. Најважнији чворови друге највеће комуне .....	26
Слика 35 Расподеле коефицијента кластеризације ( <i>UserNet</i> ).....	27
Слика 36 Расподеле коефицијента кластеризације ( <i>UserNet</i> ) Erdos-Renyi .....	28
Слика 37 Асортативност мреже за улазни степен ( <i>UserNet</i> ) .....	28
Слика 38 Асортативност мреже за излазни степен ( <i>UserNet</i> ).....	28
Слика 39 График расподеле улазног степена чвора .....	29
Слика 40 График расподеле излазног степена чвора .....	29
Слика 41 Лувенска метода ( <i>UserNet</i> ). Резолуција 1.0. Број комуна 41 .....	31

Слика 42 Лувенска метода ( <i>UserNet</i> ). Резолуција 2.0. Број комуна 27.....	31
Слика 43 Лувенска метода ( <i>UserNet</i> ). Резолуција 2.0. Број комуна 143.....	31

## 10. ТАБЕЛЕ

Табела 1 Метрике мрежа SNet и SNetT .....	32
---	----