



زن زندگی آزادی

دانشگاه تهران



دانشکده مهندسی برق و کامپیوتر

## درس شبکه‌های اجتماعی پروژه نهایی

|            |                    |
|------------|--------------------|
| سیاوش رزمی | نام و نام خانوادگی |
| ۸۱۰۱۰۰۳۵۲  | شماره دانشجویی     |
| ۱۴۰۱/۱۱/۲۴ | تاریخ ارسال گزارش  |

## فهرست

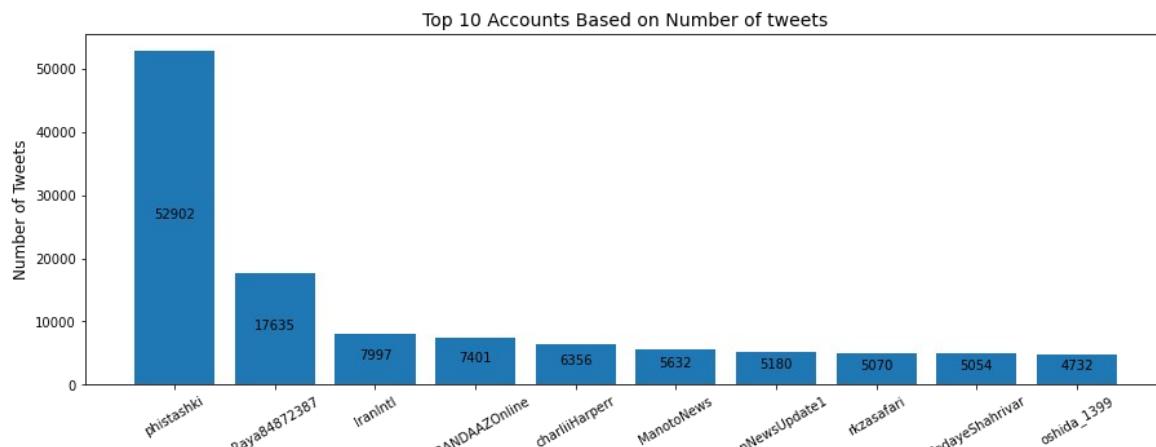
|         |  |
|---------|--|
| ۱.....  | پاسخ ۱. بررسی اکانت ها از معیار های مختلف..... |
| ۸.....  | پاسخ ۲ - بررسی شاخص های Centrality.....        |
| ۸.....  | ۲-۱. اکانت ها با بیشترین مقدار Hub             |
| ۹.....  | ۲-۲. اکانت ها با بیشترین Authority:            |
| ۹.....  | ۲-۳. اکانت ها با بیشترین PageRank:             |
| ۱۰..... | ۲-۴. اکانت ها با بیشترین In Degree:            |
| ۱۱..... | پاسخ ۳- بررسی شاخص K-shell                     |
| ۱۱..... | ۳-۱. توضیح کد:                                 |
| ۱۵..... | پاسخ ۴ - خوشه بندی با استفاده از Louvian       |
| ۱۹..... | پاسخ ۵ - خوشه بندی با استفاده از Infomap       |
| ۲۴..... | پاسخ ۶ - بررسی ویژگی Robbot و BTSexo           |
| ۲۹..... | پاسخ ۷ - بررسی سن اکانتها                      |
| ۳۱..... | پاسخ ۸ - بررسی میزان فعالیت کاربران            |

## پاسخ ۱. بررسی اکانت ها از معیار های مختلف

لیست ۱۰ اکانت برتر در ویژگی های نودها را بباید و محتوای توییتهای این اکانت هارا در توییتر بررسی کنید.

با مرتب سازی اکانت ها بر اساس هر کدام از ویژگی ها میتوان اکانت های برتر در هر ویژگی را بررسی کرد، در ابتدا ویژگی تعداد توییت را مورد بررسی قرار دادیم:

با مشاهده نمودار شکل ۱ میتوان به وضوح مشاهده کرد که تعداد توییت های اکانت phistashki از مابقی اکانت ها با اختلاف بیشتر است، این موضوع موجب شد که محتوی این اکانت را در توییتر بررسی کنیم.



شکل ۱: ده اکانت برتر از نظر تعداد توییت



شکل ۲: اکانت **phistashki** در توییتر

با بررسی محتوی این اکانت، احتمال ربات بودن آن بسیار زیاد به نظر می‌رسد، آخرین توییت این کاربر مربوط به ۲۲ زانویه و تمام پست های آن صرفاً ۲ بیت شعر با هشتگ های mahsaAmini و مهساامینی است، البته مقدار ویژگی ربات برای این کاربر ۰.۱۲ است، با بررسی محتوی توییت ها در بازه ۲۳ شهریور تا ۲۵ آبان در توییتر تعداد پست های کاربر بسیار محدود است که احتمال این می‌رود که بعد از این بازه توییت های خود را پاک کرده باشد.

اکانت barandaazonline ساسپند شده، اکانت charliiharperr نیز در تمامی پست های خود صرفاً یک پست با ماهیت (برای...) با هشتگ های ثابت دارد البته احتمال ربات بودن این اکانت حدود ۰.۰۸ است با توجه به اینکه این اکانت هر پست خود را شماره زده کاملاً واضح است که تعداد زیادی از پست های وی پاک شده که دلیل آن کاملاً مشخص نیست.

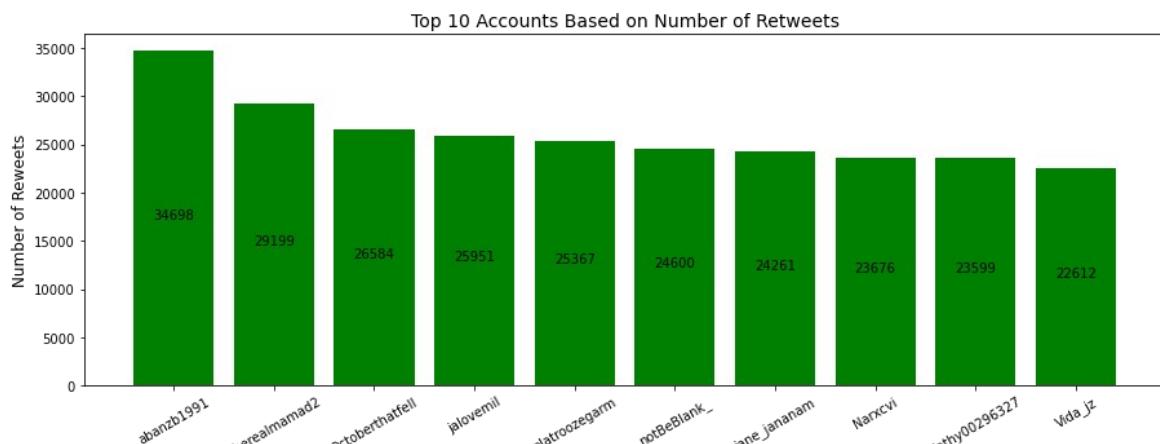
اکانت oshida\_1399 اما به نظر کاملاً واقعی میرسد و محتوی اکثر توییت ها در آن بازه ویدیوهای اعتراضات با هشتگ های قیام تا سرنگونی، اعتراضات سراسری و مهسا امینی است برای بررسی بهتر حتی پست های قبل از بازه ۲۳ شهریور را نیز بررسی کردیم که به نوعی روزمره نویسی بود ولی گهگاه با پست های اعتراضی نیز همراه میشد.

اکانت rkzasafari از ابتدا به پست های سیاسی و مباحث مربوط به خاورمیانه، جنگ اکراین و اخبار مربوط به اسرائیل میگذاشته و بعد از ۲۳ شهریور نیز کاملاً بروی اعتراضات ایران تمکن داشته است.

نکته عجیب در بررسی اکانت ها در توییت این بود که در صورتی که پست های آن هارا در بازه ۲۳ شهریور تا ۲۵ آبان بررسی میکردیم تعداد پست بسیار کمی بازگردانده میشد که مطابق با آمار های داده شده در داده های پروژه نبود، دلیل این مسأله ممکن است پاک شدن پست توسط صاحب اکانت باشد این سؤال پیش می آید که چرا این عمل انجام شده است؟

مابقی اکانت ها اما اکثراً کانال های خبری هستند و ماهیت سازمانی دارند که با تعداد بالای آمار توییت منطبق است.

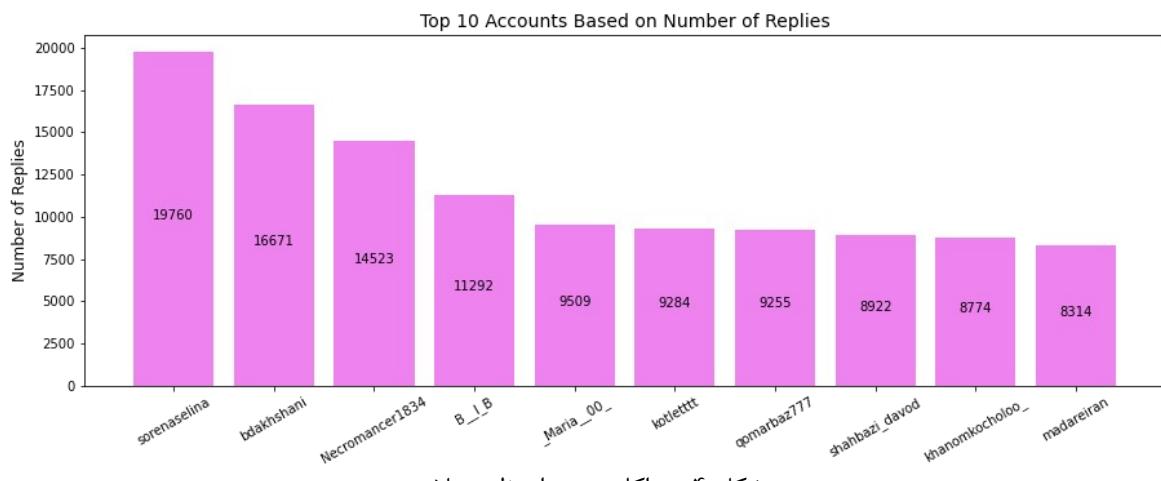
## ۲-۱. تعداد ریتیویت ها



شکل ۳: ده اکانت برتر از نظر ریتیویت

اکانت abanzb1991 که بیشترین میزان ریتیویت را داشته عموماً محتوی آموزشی مربوط به اعتراضات در آن بازه منتشر می کرده است و احتمال ربات بودن آن ۰ است، مابقی اکانت های این لیست اما احتمال ربات بودن به نسبت بالایی دارند اکانت Therealmamad2 اخبار اعتراضات را منتشر میکرده، اکانت octoberthatfell قبل از شروع اعتراضات محتوی غیر سیاسی داشته اما از ۲۳ شهریور شروع به انتشار اخبار و پوشش اعتراضات کرده، اکانت jalovemil سلطنت طلب بوده و بعد از شروع اعتراضات اخبار اعتراضات را منتشر می کرده است، محتوای alatoozegarm بیشتر محتوای اعتراضی و خبر رسانی اعتراضات را منتشر می کرده است، محتوای اکثر این اکانت های اعتراضی و این کاربران تعداد فالور زیادی ندارند که احتمال ربات بودن آن هم تا حدی بالا می برد.

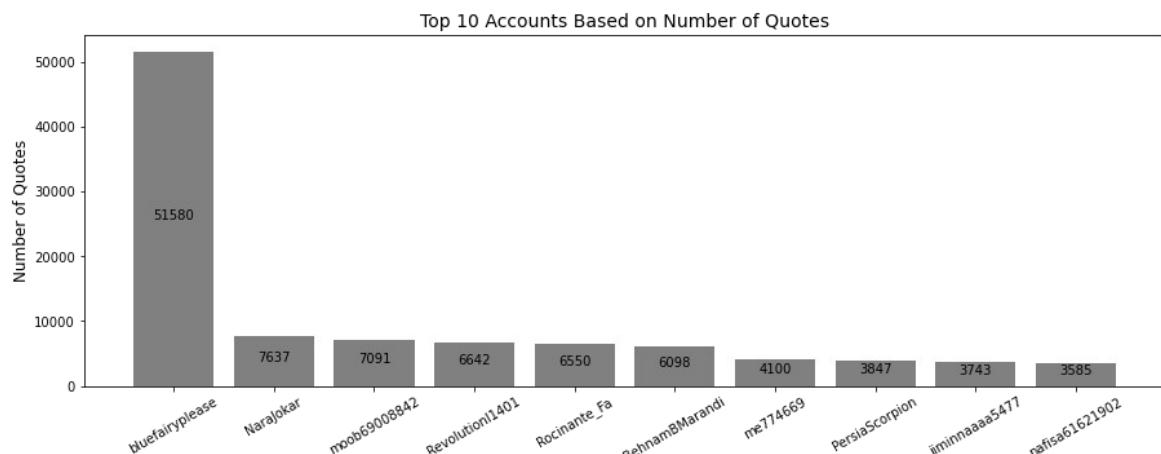
### ۱-۳. تعداد ریپلای:



شکل ۴: ده اکانت برتر از نظر ریپلای

اکانت های برتر از لحاظ ریپلای نیز احتمال ربات بودن نسبتاً بالایی دارند از لحاظ محتوی نیز محتوی آموزشی، خبررسانی و یا تشویق به اعتراض و اعتراض داشتند. اکانت sorenaselina بیشتر به نظر به ریپلای کردن یک سری عبارت مشخص در پست های دیگران مشغول بوده است، قبل از bdakhshani، necromancer1834 اعتراضات بیشتر مطالب متفرقه منتشر میکرده و فعالیتش شخصی است، در توابیت وجود ندارد، shahbazi\_davod ساسپند شده و مابقی اکانت ها محتوی خاصی برای تحلیل نداشتند.

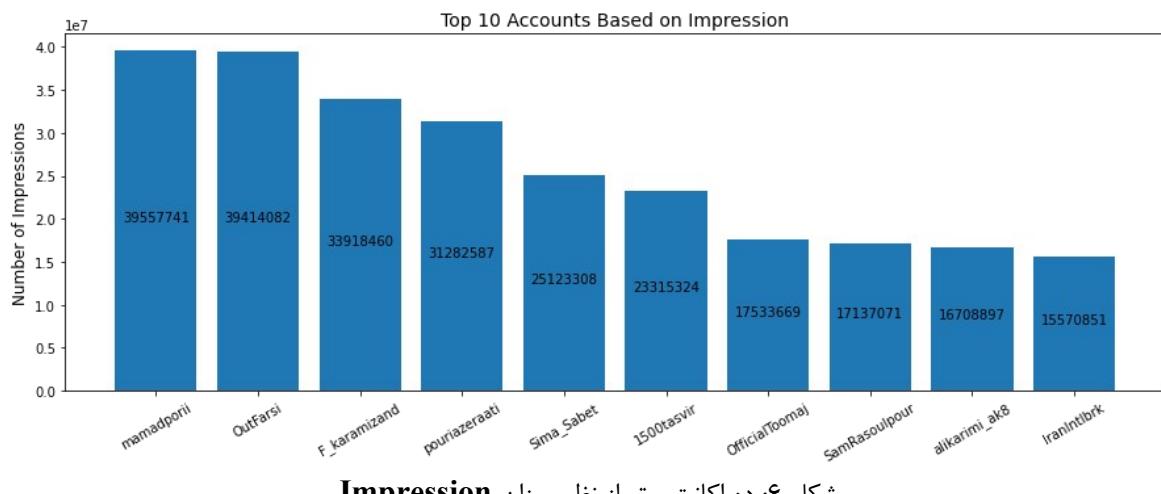
### ۱-۴. تعداد Quote



شکل ۵: ده اکانت برتر از نظر Quote

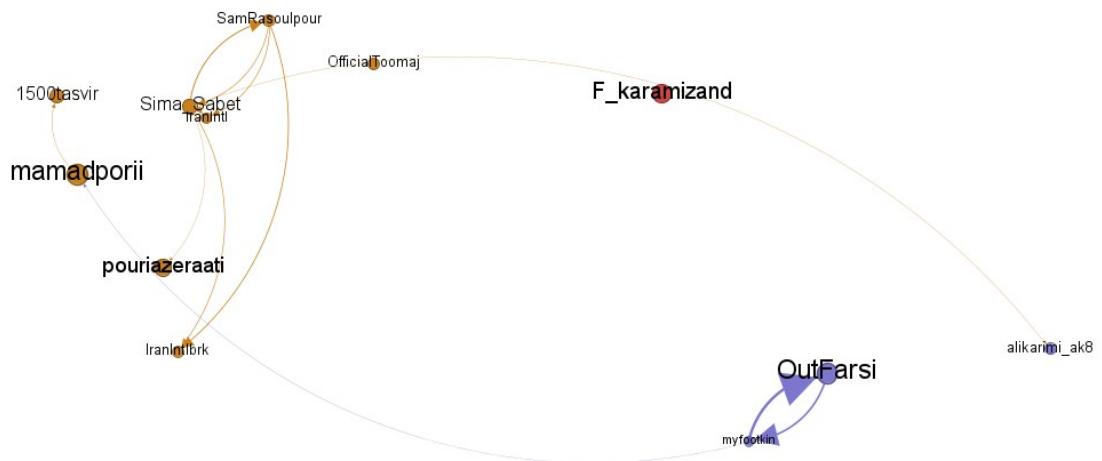
کاربر bluefairyplease با اختلاف زیادی از مابقی کاربران بالاتر بود، با بررسی این اکانت اما متوجه می شویم که صرفاً باز نشر کردن توابیت های دیگران به شکل quote میزان آنرا بالا برد و محتوی از خود تولید نکرده است، محتوی مانند قبل آموزش و ترغیب به اعتراض بود. مابقی کاربران البته محتوی قابل توجهی نداشتند

## ۱-۵ میزان Impression



شکل ۶: ده اکانت برتر از نظر میزان Impression

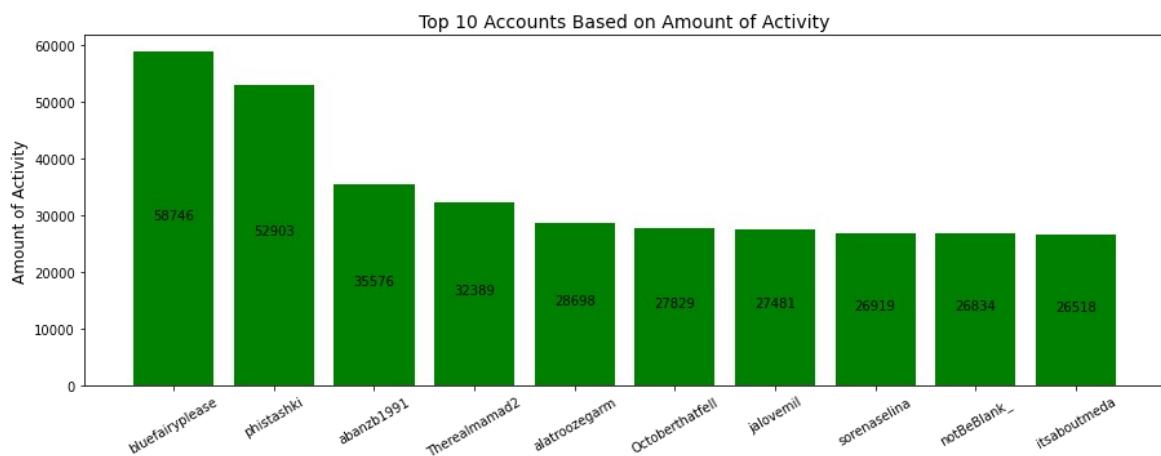
کاربران با بیشترین **impression** همگی از اکانت های مطرح توییتر محسوب می شوند که سهم اصلی را در توییتر در زمان اعتراضات داشتند، در ابتدا به آماری کاربران می پردازیم، همگی کاربران حدود ۶۴ داشتند بنابرین از ابتدای اعتراضات در توییتر حضور فعال داشته اند، کم سن ترین این اکانت ها یعنی **official\_toomaj** و **myfootkin**، **f\_karamizand** همگی از یک سال بیشتر بود و هیچ کدام آنها اکانت جدید نبودند و علاوه بر این اکثر آنها از قبیل علی کریمی، سام رسول پور، توماج صالحی، سیما ثابت پور یا زراعتی چهره های شناخته شده یا به اصطلاح سلبریتی هستند که شهرت خود در توییتر را مدیون فعالیت های بیرون از این فضا هستند، **Outfarsi** اکانت طنز نویس بود که بعد از اعتراضات به طور اختصاصی فعالیت سیاسی انجام میدهد، پور یا زراعتی، سیما ثابت، سام رسول پور همگی خبرنگاران شبکه ایران اینترنشنال هستند که به طور حرفة ایی فعالیت سیاسی انجام می دهند، ۱۵۰۰ تصویر قبل از ۲۳ شهریور نیز فعالیت حرفه ایی سیاسی انجام می دهد، توماج صالحی رپر اعتراضی است که چندیnen سال است فعالیت سیاسی انجام می دهد، **mamadporii** قبل از اعتراضات هم فعالیت سیاسی اجتماعی داشته و خود را **social activist** معرفی می کند اما بعد از اعتراضات تمرکز خود را بر روی اخبار اعتراضات گذاشته است.



شکل ۷: ارتباط گره های با بیشترین میزان **Impression**

در شکل فوق ارتباط اکانت ها بر اساس میزان ریتویت آورده شده است، رنگ گره ها براساس انجمانی است که بر اساس الگوریتم **Louvian** گره در آن قرار دارد، ضخامت یال ها بر اساس وزن به شکل **rescale** شده و رنگ آن ها نیز بر اساس رنگ گره مبدأ است، نکته جالب توجه این است که اکثر اکانت های برتر مربوط به شبکه ایران اینترنشنال بوده و یا با آن ها ارتباط نزدیک دارد، اکانت فریبرز کرمی زند تقریباً با هیچ کدام از این ۹ اکانت ارتباطی نداشته و ایمپرشن بالای آن بخاطر محتوی آموزشی که خود تولید میکند است، اکانت توماج صالحی نیز به همین شکل ارتباط خاصی با باقی اکانت ها نداشته است، اکانت های مربوط به ایران اینترنشنال اما همگی با همدیگر ارتباط نسبتاً نزدیک دارند، اکانت های **Outfarsi** و **myfootkin** قوی ترین لینک ها را بین یکدیگر دارند و نیز رابطه نسبتاً ضعیفی با **mamadporii** دارد، سیما ثابت بیشترین میزان ریتویت را از اکانت های ایران اینترنشنال و ایران اینترنشنال اسپورت و سام رسول پور را داشته و علی کریمی نیز ارتباط ضعیفی با سیما ثابت دارد.

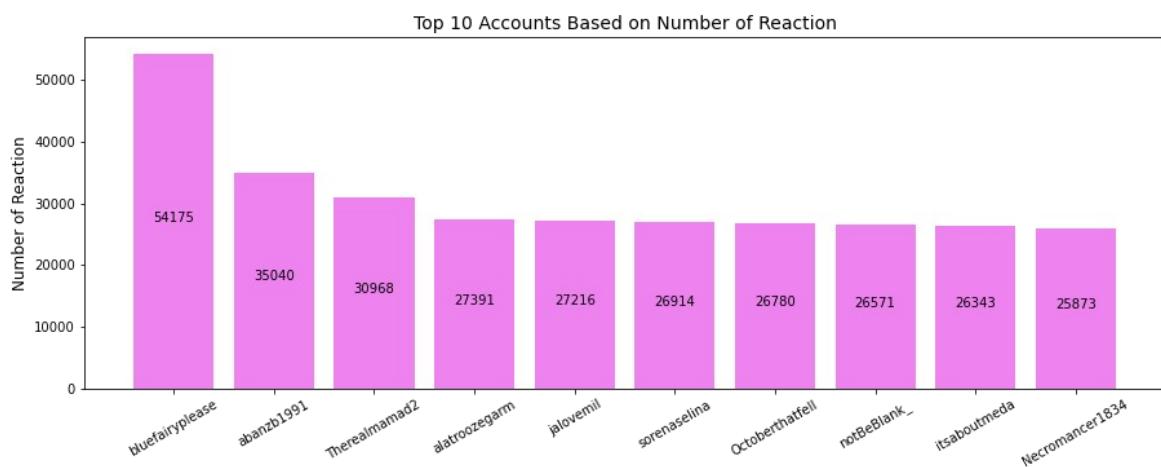
## ۱-۵. میزان Activity



شکل ۸: ده اکانت برتر از نظر میزان Activity

با توجه به اینکه معیار Activity از جمع ستون‌های دیگر بدست می‌آید لیست ۱۰ اکانت برتر همپوشانی زیادی با دیگر اکانت‌های برتر ویژگی‌های تشکیل دهنده‌ی آن دارد که البته اکانت‌های هارا در قسمت‌های قبلی بررسی کرده ایم.

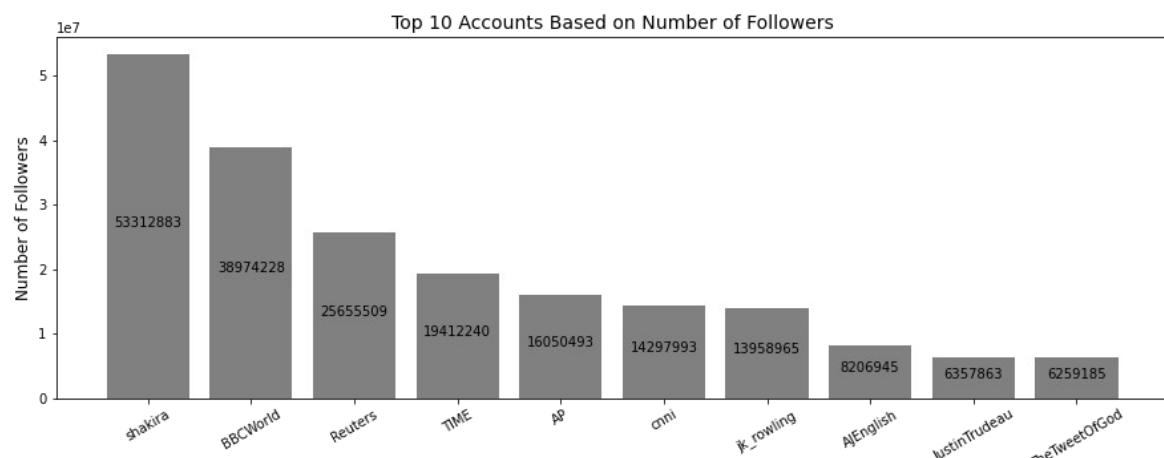
## ۱-۶. تعداد Reaction



شکل ۹: ده اکانت برتر از نظر تعداد Reaction

ویژگی reaction نیز تقریباً شبیه به quote و activity است که قبلاً بررسی کردیم.

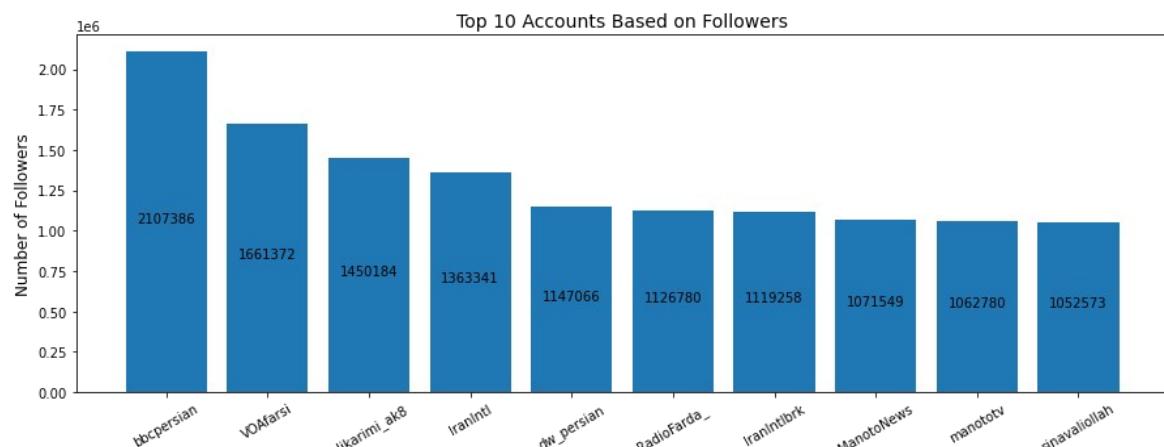
## ۱-۴. تعداد Follower



شکل ۱۰: ده اکانت برتر از نظر تعداد Follower

اکانت های با بیشترین دنبال کننده همگی اکانت های خارجی بودند که فعالیت های بسیار کمی هم در زمینه اعتراضات داشتند از اشخاص حقیقی به طور مثال اکانت شکیرا به طور کل ۱ توییت در زمینه مهسا امینی زده است و مابقی اکانت ها نیز اکانت های خبرگزاری های جهانی است که آن ها نیز به نسبت خبرگزاری های داخلی فعالیت بسیار کمتری داشته اند بنابرین به نظر مناسب می رسد که ۱۰ اکانت فارسی را از لحاظ دنبال کننده بررسی کنیم.

### اکانت های فارسی از لحاظ تعداد دنبال کننده:

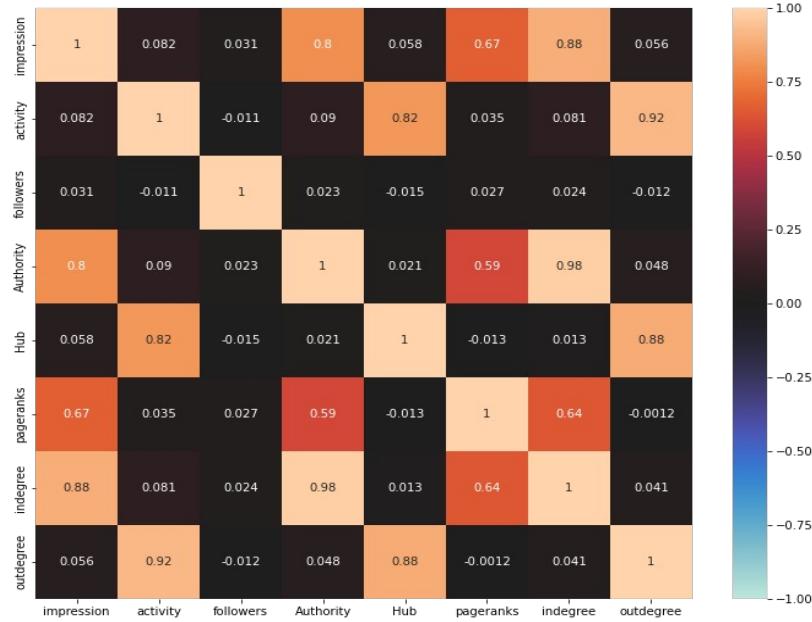


شکل ۱۱: ده اکانت برتر فارسی از لحاظ تعداد دنبال کننده

در ده اکانت فارسی با بیشترین دنبال کننده به جز اکانت علی کریمی و سینا ولی الله همگی متعلق به خبرگزاری ها بود، در این میان اکانت ایران اینترنشنال بیشترین فعالیت را در میان این اکانت ها داشت.

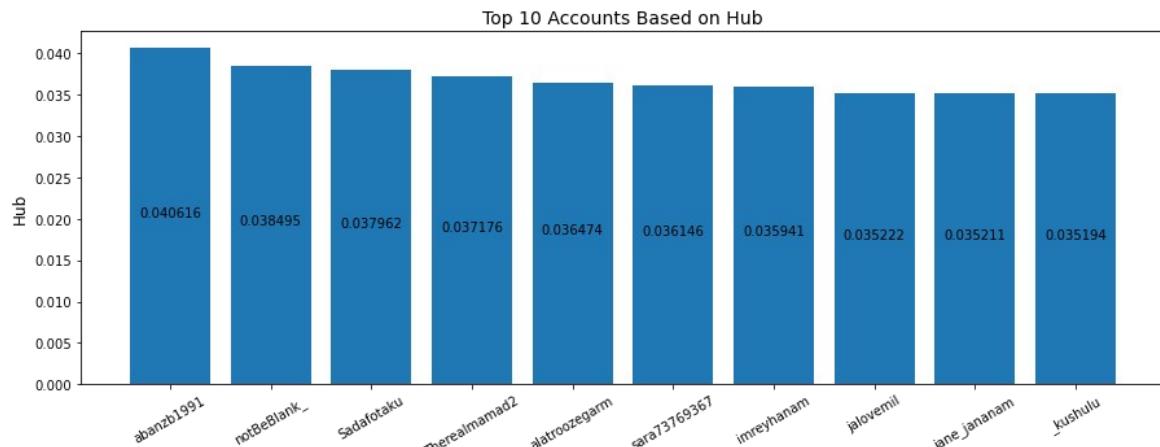
## پاسخ ۲ - بررسی شاخص های Centrality

در ابتدا با رسم ماتریس همبستگی به بررسی آماری شاخص های بدست آمده می پردازیم



شکل ۱۱: ماتریس همبستگی شاخص های محاسبه شده

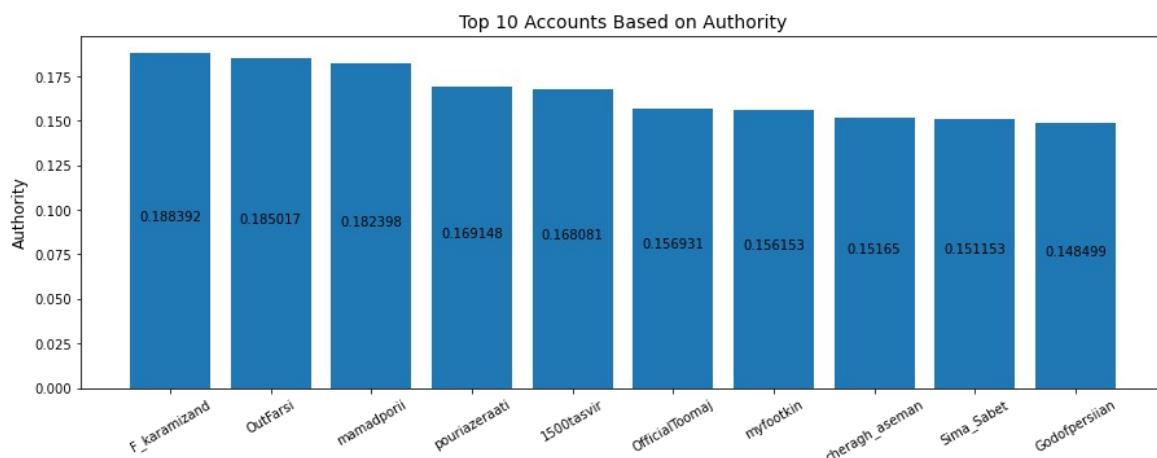
### ۱-۲. اکانت ها با بیشترین مقدار Hub



شکل ۱۲: ده اکانت برتر از لحاظ میزان Hub

با توجه به جدول همبستگی شاخص Hub بیشترین همبستگی را با outdegree دارد به این معنا که اکانت هایی که شاخص hub بالایی دارند میزان ریتویت بالایی هم از اکانت های دیگر داشته اند، که این مسئله با بررسی محتوی اکانت ها در قسمت قبل هم مشهود بود، این کاربران خودشان محتوی ارزشمندی تولید نمیکنند اما از اکانت های دیگر محتوی را ریتویت میکنند.

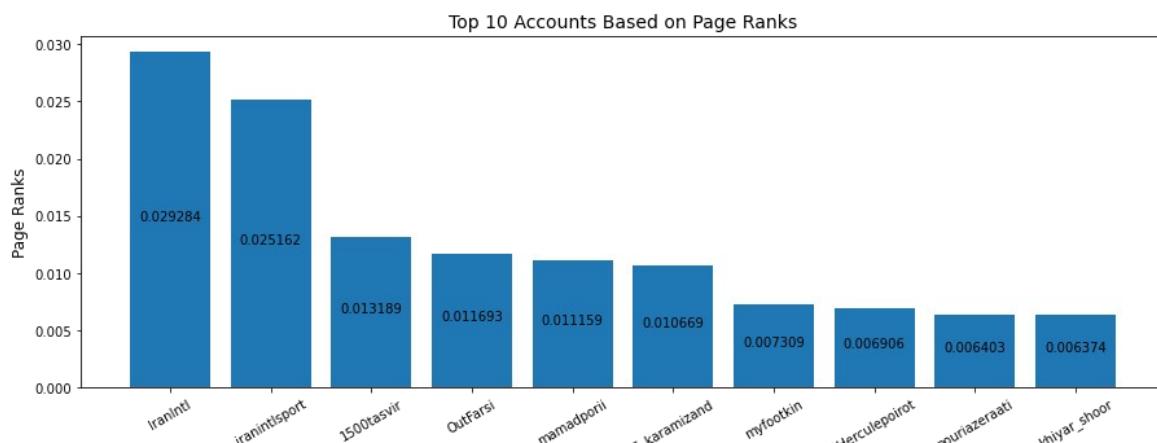
## ۲-۲. اکانت ها با بیشترین Authority



شکل ۱۳: ده اکانت برتر از نظر میزان Authority

ویژگی Authority بیشترین میزان همبستگی با Impression و Indegree دارد به این معنا که این شاخص بیشتری ارتباط را با میزان ریتوبیت پست هایشان دارد. با بررسی اکانت ها به نظر می رسد اکانت هایی بیشترین Authority را دارند که محتوى Genuine دارد و خودشان محتوى اصيل تولید میکنند به این معنا که محتوى توییت ها را از جایی کپی یا محتوى تکراری نمی گذارند و خبر های دسته اول پست می کنند.

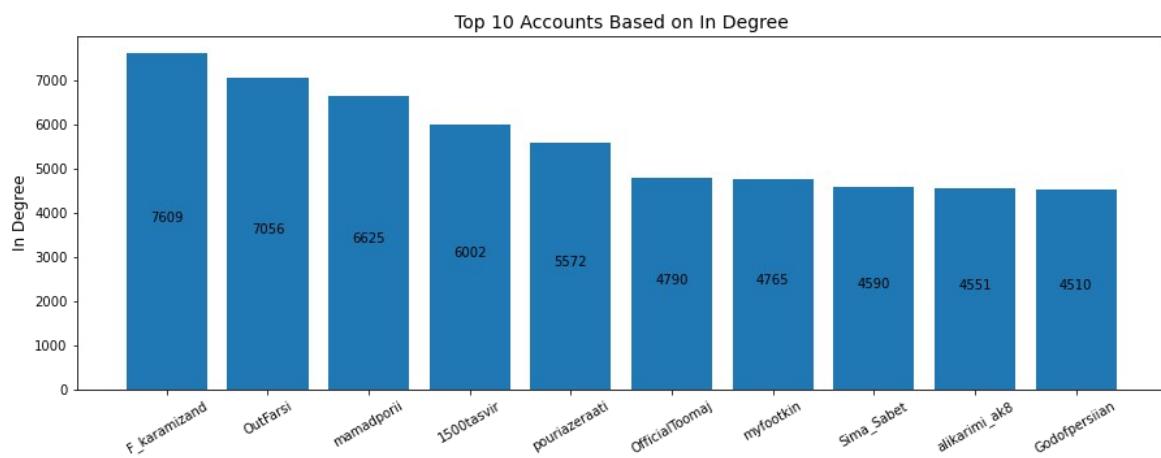
## ۲-۳. اکانت ها با بیشترین PageRank



شکل ۱۴: ده اکانت برتر از نظر PageRank

طبق تعریف اکانت هایی بیشترین Pagerank را دارند که میزان پست هایی که بقیه از آنها ریتوبیت میکنند از تعداد پست هایی که آنها از بقیه ریتوبیت میکنند بیشتر باشد یا در مثال اقتصادی میزان دخلشان از خرج آنها بیشتر باشد، اکانت هایی که در شکل بالا Pagerank بالایی دارند اکانت هایی بودند که به عنوان خبرگزاری فعالیت می کردند و یا اخبار و ویدیوهای اعتراضات را اشتراک می گذاشتند یا آموزش اعتراض می دادند. با توجه به جدول همبستگی نیز PageRank با Indegree و Impression مثبت دارد.

## ۴-۲. اکانت ها با بیشترین In Degree



شکل ۱۵: اکانت های برتر از نظر In degree

همانطور که در بالا گفته شد شاخص In degree همبستگی بسیار زیادی با Authority و Impression دارد و دلیل آن نیز به این خاطر است که اکانت هایی که Authority بالای دارند محتوی تولیدی آنها ریتویت زیادی دارد و این باعث می شود که In degree آنها بالا رود همچنین شاخص Impression نیز از روی میزان ریتویت محاسبه می شود. اکانت های با Pagerank بالا نیز با این دو لیست نسبتاً اشتراک کمتری دارد که دلیل آن میتوان این باشد که این اکانت ها ریتویت زیادی دارند اما چون علاوه بر In degree، شاخص Outdegree نیز برای آنها در نظر گرفته می شود بنابرین ممکن است برخی کاربران در لیست قرار نگیرند.

## پاسخ ۳- بررسی شاخص K-shell

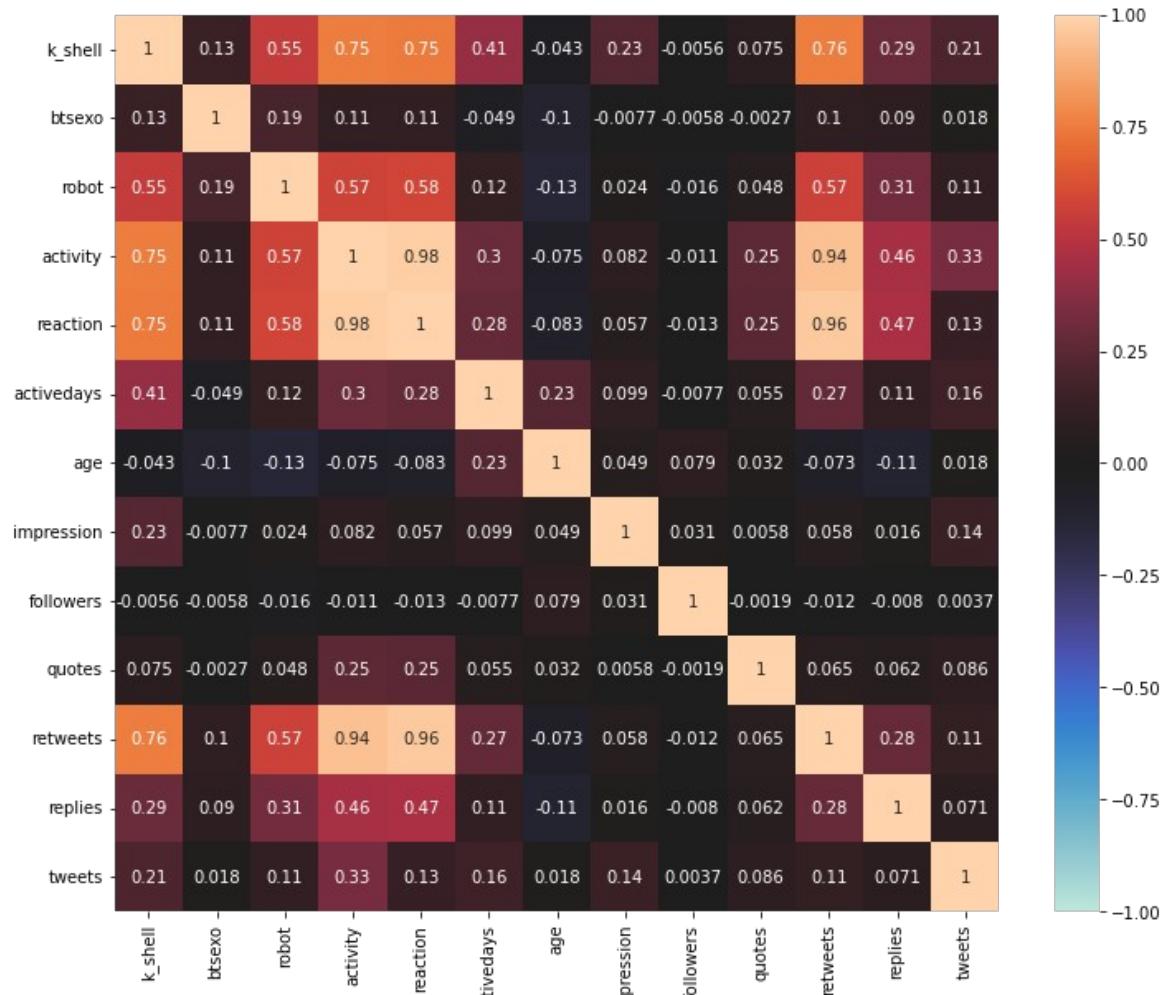
۱- توضیح کد:

در ابتدا با تبدیل گراف به حالت بی جهت آنرا از نرم افزار گفی خروجی گرفته و با استفاده از کتابخانه **networkx** مقدار **k\_shell** را برای همه گرهها به دست می آوریم:

```
1 G = nx.read_graphml('./final_demo/undirected.graphml')
2 k_shell = nx.core_number(G)
3 nx.set_node_attributes(G,k_shell,'k_shell')
4 nx.write_graphml(G,'./final_demo/graph_with_kshell.graphml')
```

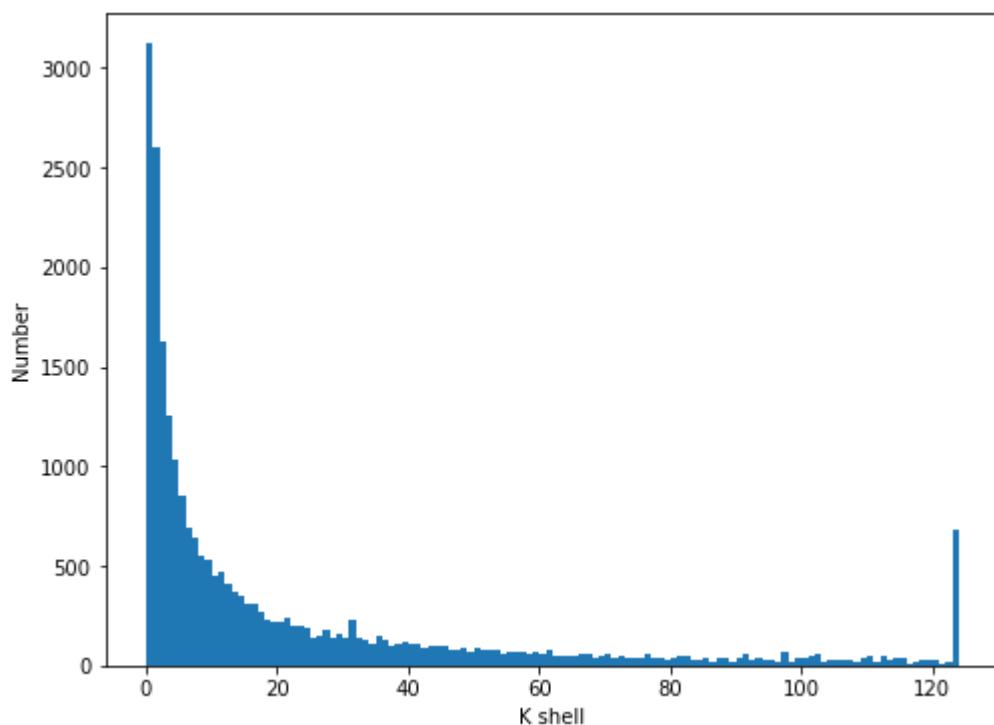
شکل ۱۶: کد استفاده شده جهت محاسبه **K shell**

برای بررسی آماری دقیق‌تر از ماتریس همبستگی برای مقایسه ویژگی‌های مختلف با **k shell** استفاده کردیم:



شکل ۱۷: ماتریس همبستگی ویژگی های مختلف

همانطور مشاهده می شود k shell بیشترین همبستگی را با ویژگی های reaction، retweet و activity دارد که البته با توجه به اینکه لینک های گره ها از طریق میزان ریتوات ریتوات محاسبه شده و در محاسبه k shell هم ما لینک بین گره هارا در نظر می گیریم بنابرین این همبستگی طبیعی است.



شکل ۱۸: هیستوگرام توزیع مقادیر **K shell**

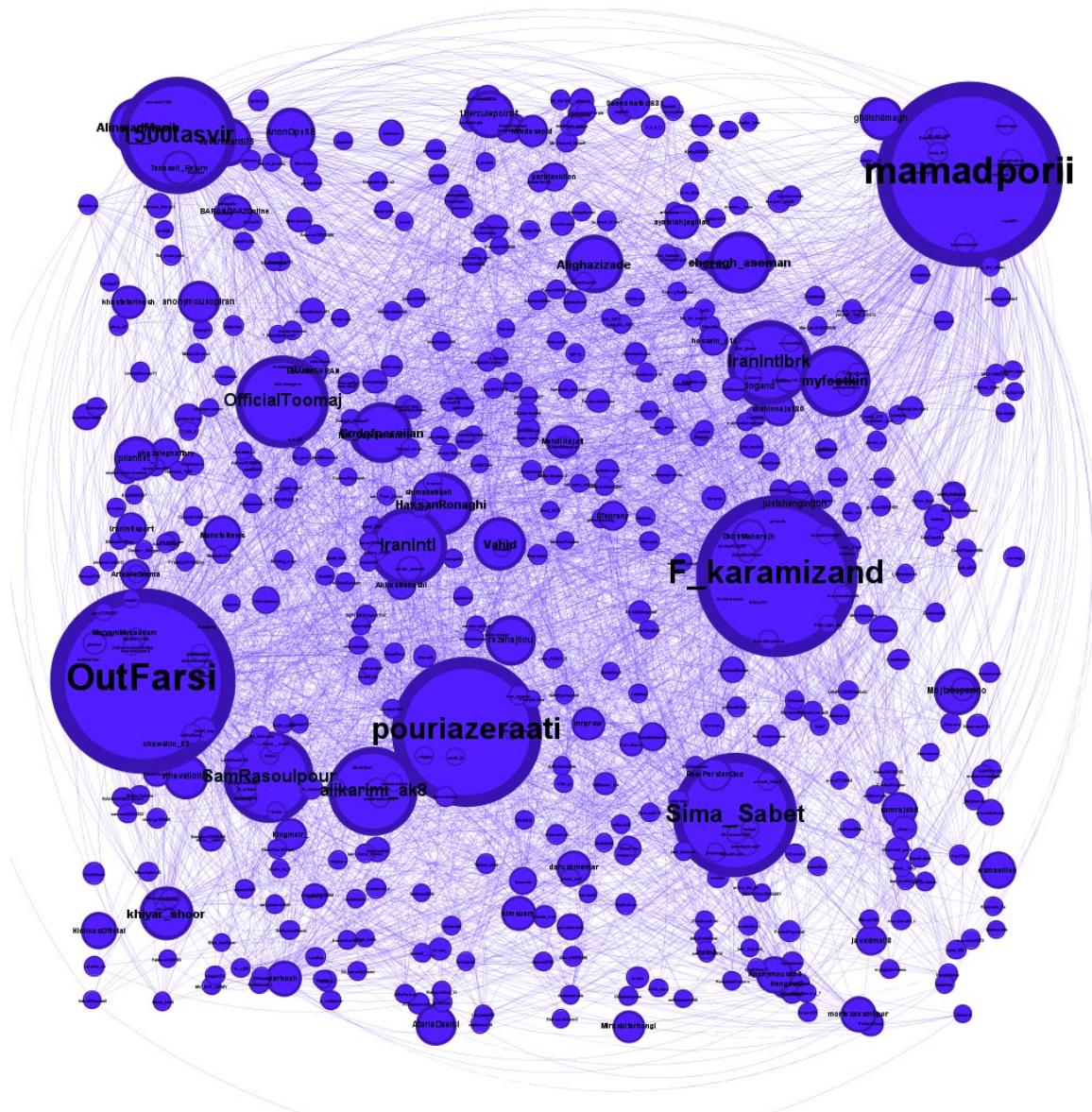
نکته جالب توجه نسبت به مقادیر **K shell** این موضوع است که تعداد مقادیر ۱۲۴ k-shell نسبت به مقادیر کمتر اندازه بسیار بزرگتری دارد در حالیکه روند کلی هیستوگرام یک روند نزولی است.  
لیست اکانت هایی که **k-shell** بالایی دارند ۵۵۰ عدد است که اگر بر اساس **impression** آنها را مرتب کنیم لیست به دست آمده به شکل زیر است:

| Label           | Interval | k_shell | btsexo | robot | activity | reaction | activedays | age      | impressions | followers | quotes | retweets | replies | tweets |
|-----------------|----------|---------|--------|-------|----------|----------|------------|----------|-------------|-----------|--------|----------|---------|--------|
| mamadpori       | 124      | 0       | 0.016  | 3078  | 1223     | 64       | 4146.485   | 39557741 | 266134      | 68        | 532    | 623      | 1855    |        |
| OutFarsi        | 124      | 0       | 0.005  | 1994  | 457      | 64       | 928.857    | 39414082 | 304484      | 3         | 299    | 155      | 1537    |        |
| F_karamizand    | 124      | 0       | 0.058  | 1588  | 364      | 64       | 439.634    | 33918460 | 368572      | 17        | 22     | 325      | 1224    |        |
| pouriazerati    | 124      | 0       | 0.056  | 2356  | 392      | 64       | 3482.478   | 31282587 | 643783      | 15        | 87     | 290      | 1964    |        |
| Sima_Sabet      | 124      | 0       | 0.051  | 2181  | 828      | 64       | 1793.572   | 25123308 | 541929      | 253       | 543    | 32       | 1353    |        |
| 1500taqvir      | 124      | 0       | 0.031  | 2819  | 515      | 64       | 732.672    | 23315324 | 352963      | 29        | 280    | 206      | 2304    |        |
| OfficialToomaj  | 124      | 0       | 0.019  | 683   | 114      | 63       | 421.272    | 17533669 | 466983      | 11        | 42     | 61       | 569     |        |
| SamRasoulpour   | 124      | 0       | 0.02   | 2104  | 575      | 64       | 5073.566   | 17137071 | 221963      | 61        | 443    | 71       | 1529    |        |
| alikarimi_ak8   | 124      | 0       | 0.0    | 75    | 0        | 61       | 1786.974   | 16708897 | 1450184     | 0         | 0      | 0        | 75      |        |
| Iranlightbtk    | 124      | 0       | 0.038  | 4468  | 511      | 64       | 1623.358   | 15570851 | 1119258     | 7         | 1      | 503      | 3957    |        |
| iranlntf        | 124      | 0       | 0.055  | 8393  | 396      | 64       | 2387.548   | 14157855 | 1363341     | 0         | 350    | 46       | 7997    |        |
| myfootkin       | 124      | 0       | 0.008  | 1032  | 542      | 64       | 476.84     | 12391761 | 106971      | 2         | 400    | 140      | 490     |        |
| HassanRonaghi   | 124      | 0       | 0.01   | 603   | 131      | 64       | 2815.309   | 10475919 | 180220      | 67        | 8      | 56       | 472     |        |
| cheragh_aeseman | 124      | 0       | 0.115  | 3560  | 2385     | 62       | 141.768    | 10223671 | 56971       | 13        | 2149   | 223      | 1175    |        |
| Godofpersian    | 124      | 0       | 0.077  | 1880  | 1257     | 64       | 589.335    | 9973434  | 96756       | 188       | 816    | 253      | 623     |        |
| Alighazizade    | 124      | 0       | 0.012  | 457   | 53       | 63       | 3293.354   | 9496464  | 445793      | 4         | 40     | 9        | 404     |        |
| AlinejadMasih   | 124      | 0       | 0.026  | 1844  | 665      | 64       | 1781.37    | 9423225  | 628055      | 72        | 468    | 125      | 1179    |        |
| khiyar_shoor    | 124      | 0       | 0.028  | 747   | 273      | 53       | 3671.577   | 8441075  | 128928      | 17        | 136    | 120      | 474     |        |
| Vahid           | 124      | 0       | 0.016  | 1172  | 213      | 63       | 5695.465   | 8143183  | 447197      | 7         | 23     | 183      | 959     |        |
| rezahajilou     | 124      | 0       | 0.081  | 599   | 227      | 64       | 1041.764   | 7518755  | 168836      | 1         | 178    | 48       | 372     |        |
| AnonOpsSE       | 124      | 0       | 0.029  | 1010  | 124      | 58       | 3626.331   | 7414781  | 460314      | 27        | 38     | 59       | 886     |        |
| Javanvardi75    | 124      | 0       | 0.091  | 4913  | 1513     | 64       | 3882.993   | 7346008  | 140798      | 92        | 605    | 816      | 3400    |        |
| 1Herculepoiro   | 124      | 0       | 0.006  | 1230  | 354      | 64       | 187.843    | 7075373  | 59500       | 8         | 191    | 155      | 876     |        |
| Mojtabapacino   | 124      | 0       | 0.002  | 700   | 35       | 63       | 1977.769   | 6736947  | 42386       | 0         | 0      | 35       | 665     |        |
| iam_papillonnnn | 124      | 0       | 0.191  | 4030  | 2991     | 55       | 964.648    | 6302054  | 54852       | 106       | 1355   | 1530     | 1039    |        |
| piianist        | 124      | 0       | 0.0    | 4013  | 2009     | 55       | 217.712    | 6157032  | 44362       | 180       | 1397   | 432      | 2004    |        |
| sinavalollah    | 124      | 0       | 0.042  | 187   | 26       | 63       | 4978.914   | 6085994  | 1052573     | 3         | 12     | 11       | 161     |        |
| gh0lch0magh     | 124      | 0       | 0.011  | 1031  | 69       | 64       | 2147.65    | 5854537  | 43254       | 10        | 1      | 58       | 962     |        |
| anonymousopiran | 124      | 0       | 0.02   | 803   | 300      | 57       | 58.201     | 5472501  | 147269      | 177       | 21     | 102      | 503     |        |
| AtenaDaemi      | 124      | 0       | 0.006  | 380   | 148      | 64       | 3342.847   | 5395017  | 171003      | 17        | 123    | 8        | 232     |        |
| justchangingun  | 124      | 0       | 0.036  | 916   | 476      | 64       | 4888.518   | 5387159  | 56970       | 54        | 368    | 54       | 440     |        |
| 5ogand          | 124      | 0       | 0.06   | 351   | 139      | 57       | 3582.31    | 5306438  | 451229      | 8         | 129    | 2        | 212     |        |
| ManotoNews      | 124      | 0       | 0.119  | 5970  | 338      | 64       | 4023.649   | 4803780  | 1071549     | 0         | 2      | 336      | 5632    |        |
| Anonymousnn3    | 124      | 0       | 0.033  | 560   | 58       | 55       | 660.774    | 4442537  | 130316      | 23        | 27     | 8        | 502     |        |
| HichkasOfficial | 124      | 0       | 0.028  | 329   | 208      | 63       | 4471.575   | 4410983  | 675705      | 102       | 90     | 16       | 121     |        |
| esmaeilion      | 124      | 0       | 0.064  | 194   | 50       | 62       | 3296.6     | 4393338  | 615671      | 20        | 23     | 7        | 144     |        |

شکل ۱۹: لیست اکانت های با بیشترین **K shell** و **Impression**

این اکانت‌ها همگی اکانت‌های مطرح ترین‌تر هستند که البته در قسمت‌های قبلی نیز به تحلیل آن‌ها پرداختیم.

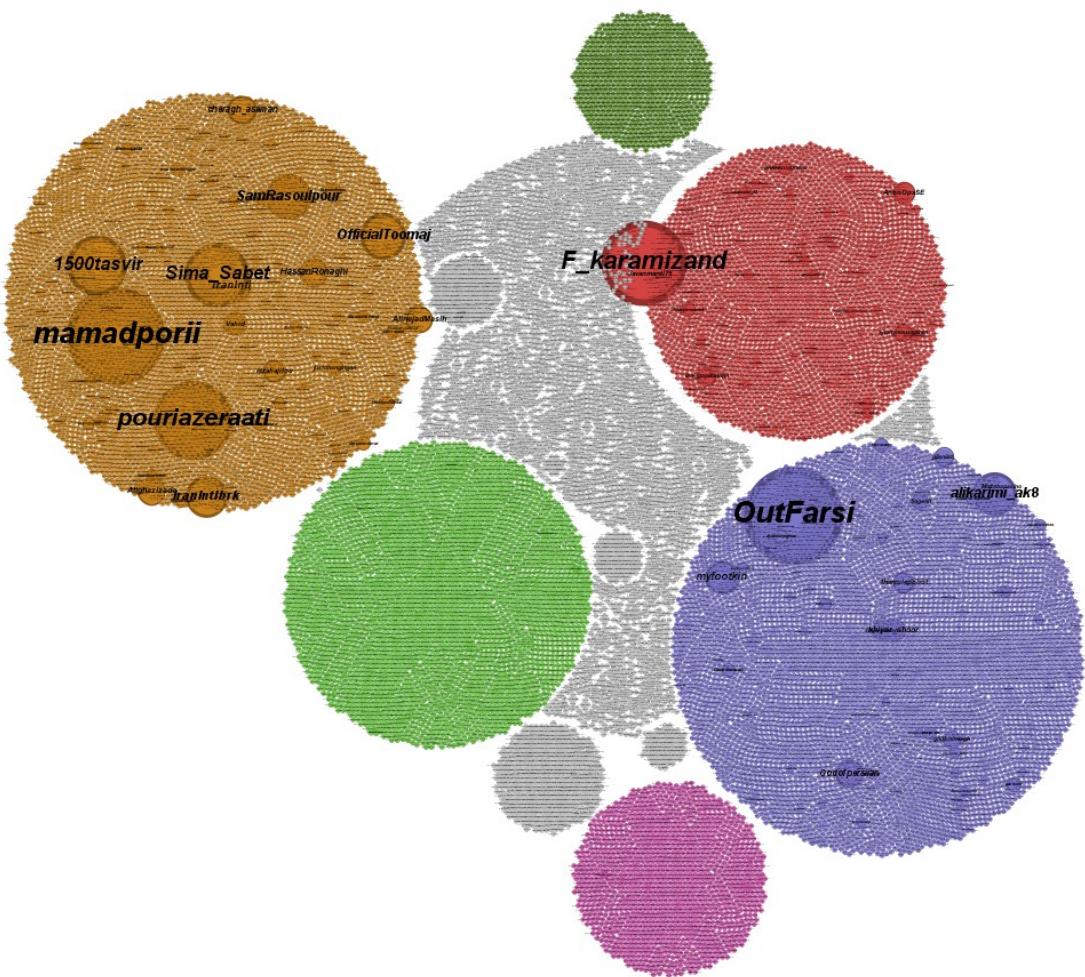
همگی آن‌ها به طور اختصاصی صرفاً مطالب مربوط به اعتراضات و اخبار مربوط به اعتراضات و آموزش و... است، از نظر عقاید این اکانت‌ها لزوماً اشتراکی ندارند به طور مثال ایران اینترنشنال، سام رسول‌پور و سیما ثابت حدوداً در یک جبهه عقیدتی هستند، یا اکانت‌های myfootkin و f\_karamizand، khiyar\_shoor و javanmardi75 ضد سلطنت است بنابرین ویژگی مشترکی که این اکانت‌هارا در یک k-shell قرار داده ربطی به عقیده سیاسی ندارد، بلکه بیشتر عنصر مخالفت با نظام جمهوری اسلامی و پشتیبانی از اعتراضات اشتراک بهتری به نظر می‌رسد.



شکل ۲۰: شبکه گره های با **k-shell** بالا

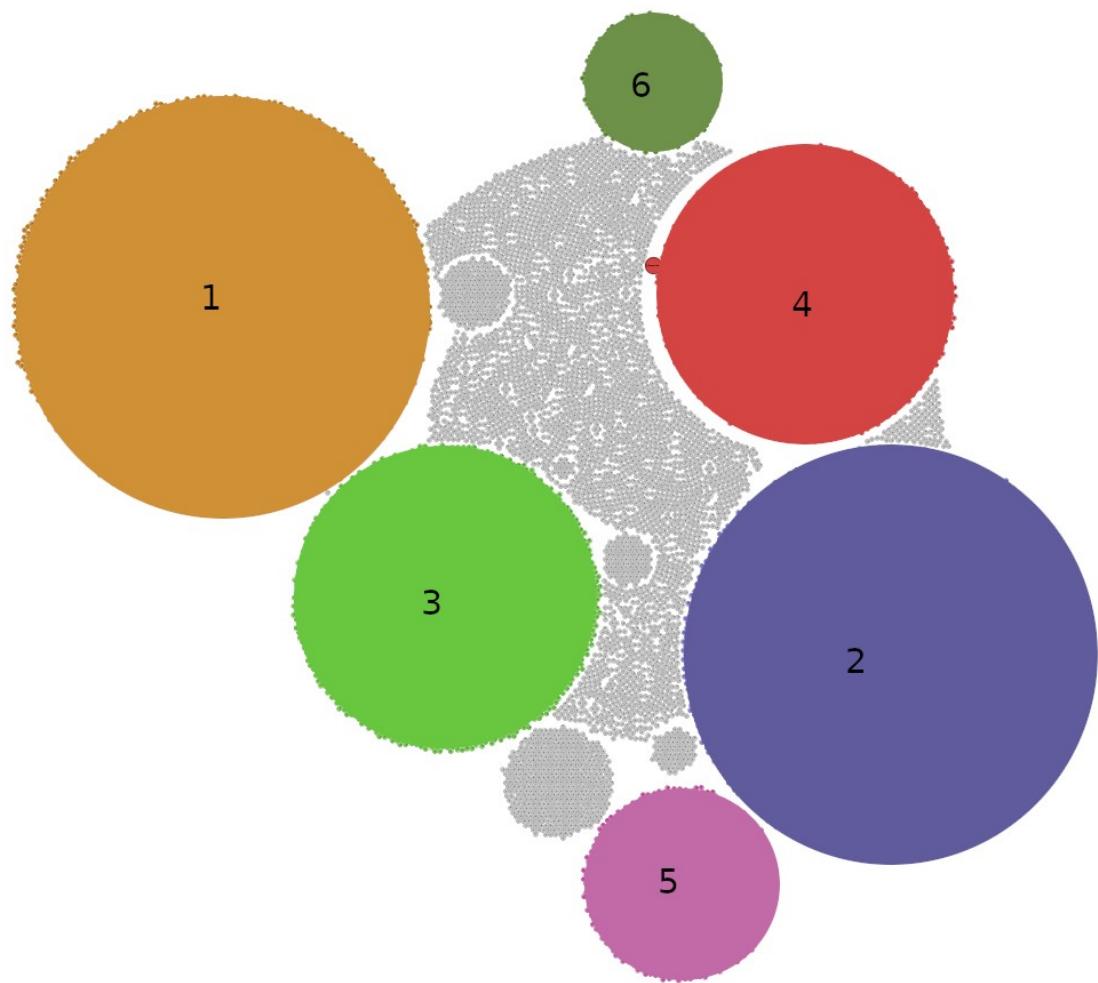
## پاسخ ۴ - خوشه بندی با استفاده از Louvian

پس از انجام Louvian گره‌های گراف به شکل زیر دسته بندی شدند، دقیق داشته باشید که الگوریتم randomize اجرا شده و نتایج ممکن است با اجرای دوباره نتایج متفاوت تولید کند:



شکل ۲۱: نحوه خوشه بندی الگوریتم Louvian

حال به بررسی هر کدام از انجمن‌ها می‌پردازیم:

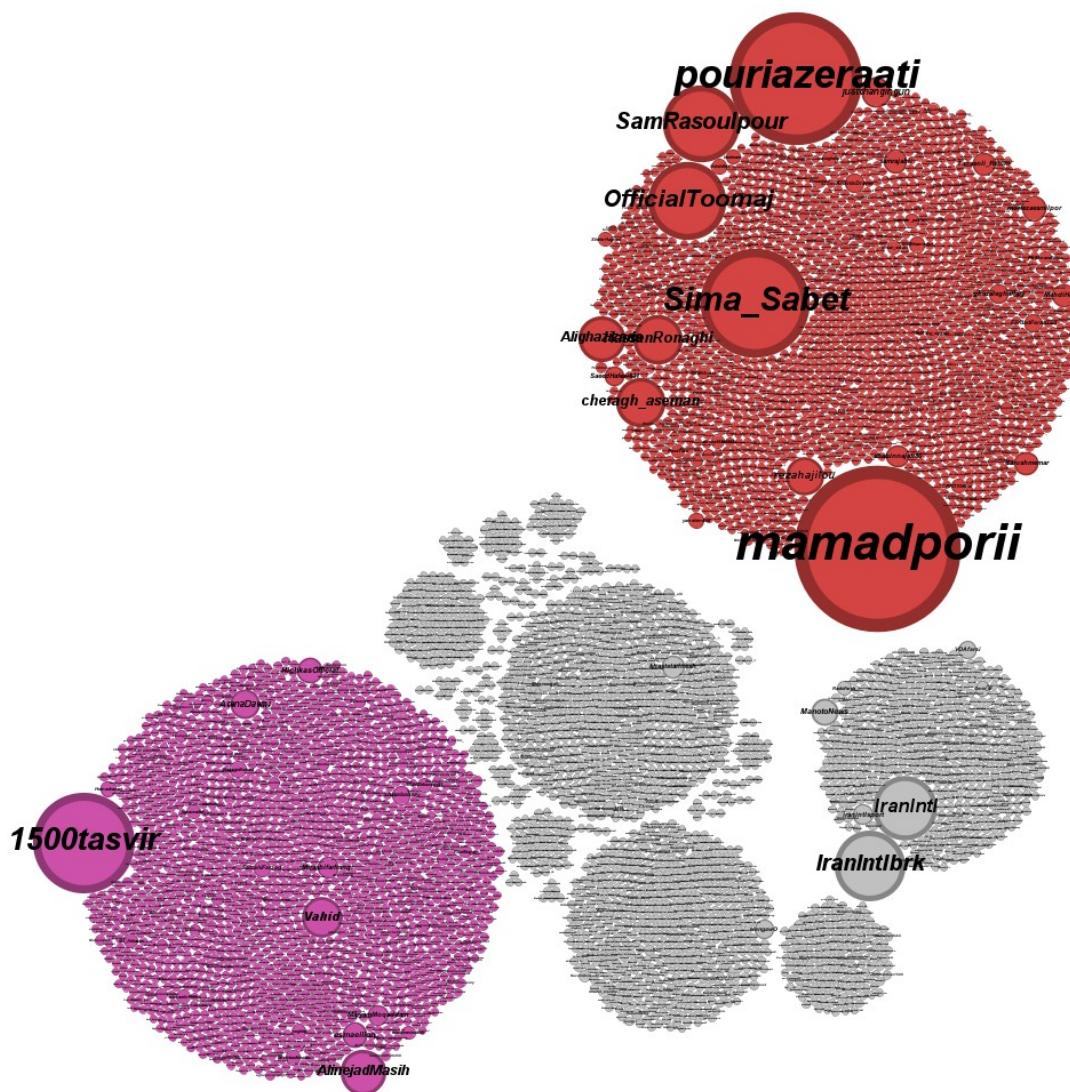


شکل ۲۲: انجمن‌های اصلی گراف

۶ گروه اصلی که سهمی بیشتر از ۲ درصد گره‌های شبکه را شامل می‌شود به شرح زیر هستند:

گروه ۱: شامل اکانت‌های mamadporii سام رسول‌پور، سیما ثابت، ایران اینترنشنال، پوریا زراعتی، cheraghe\_aseman، توماج صالحی، حسین رونقی، مسیح علی‌نژاد، ۱۵۰۰ تصویر، وجه مشترک این اکانت‌ها فعالیت حرفه‌ای و سابقه دار سیاسی و مخالفت با رژیم جمهوری اسلامی است البته این معنا نیست که همگی دارای یک عقیده یکسان هستند بلکه صرفاً به دلیل مخالفت با جمهوری اسلامی میتوان آنها را در یک دسته قرار داد.

حال در صورتی که یکبار دیگر الگوریتم را تنها برروی انجمن ۱ انجام دهیم به شکل زیر می‌رسیم:



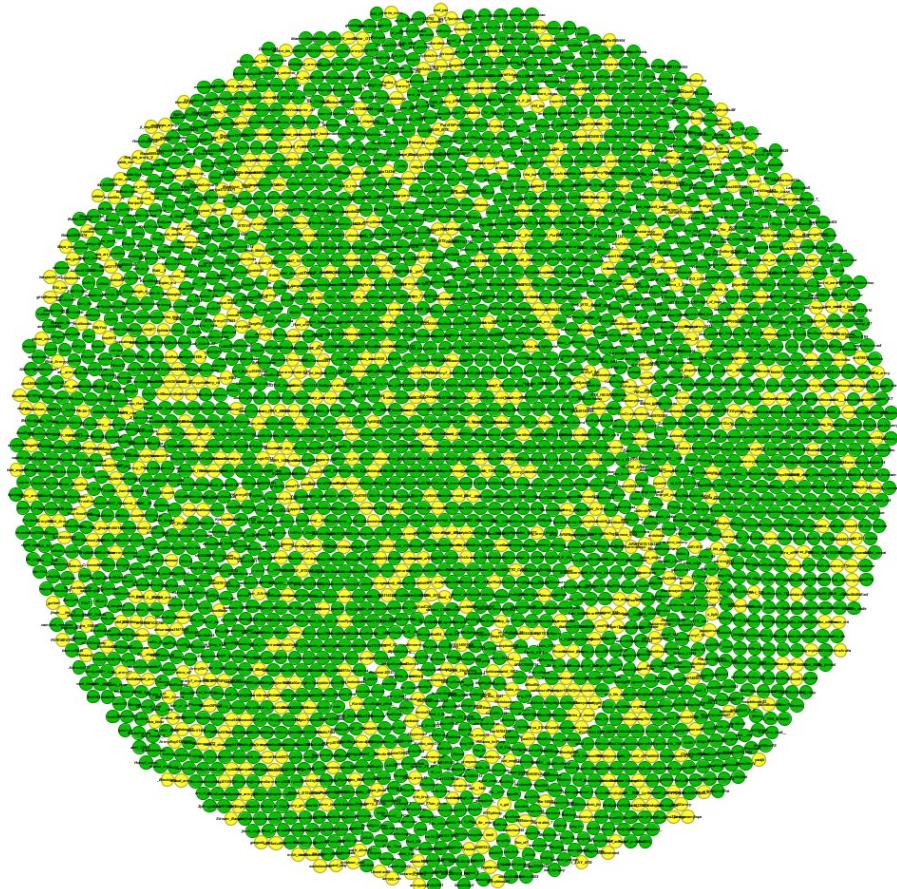
شکل ۲۳: خوشه بندی دوباره خوشه ۱ با الگوریتم Louvian

پس اجرای مجدد الگوریتم اکانت های مسیح علینژاد، وحید آنلاین، هیچکس، ۱۵۰۰ تصویر، آتنا دائمی، حامد اسماعیلیون در یک انجمن، سام رسولپور، سیما ثابت، رضا حاجیلو، توماج صالحی، علی قاضیزاده، حسن رونقی، mamadporii در انجمن دیگر و Iranintl و IranintlBrk در انجمنی جدا قرار می‌گیرند با این دسته بندی جدید اکانت های با عقاید نزدیک تر در یک زیر انجمن قرار گرفته اند.

گروه ۲: شامل اکانت های Impression، OutFarsi، myfootkin، alikarimi\_ak8، Godofpersian، Sogand و Dkhtarmaharaj، Mojtabapacino، 1herculepoirrot، khiyar\_shoor و سینا ولی الله، jz مشترک این اکانت ها این است که به طور حرفه فعالیت سیاسی انجام نمی‌دهند و بعد از شروع اعتراضات شروع به فعالیت سیاسی تمام وقت کردند.

گروه ۳: در این انجمن اکانت های با Impression بالایی وجود ندارد و تقریباً تمامی اکانت ها به شکل یک دست پخش شده اند، با بررسی نام های اکانت به این نکته میتوان پی برد که تعداد زیادی از اکانت

های این انجمن یکی از عبارت های bts,exo و... را در نام خود دارند بنابرین برای بررسی دقیق تر بروی گره های این انجمن گره هایی که ویژگی Btsexو آنها برابر با ۱ بود را رنگ دیگری دادیم:



شکل ۲۴: بررسی گره های با ویژگی Btsexو (گره های زرد رنگ)

با توجه به شکل مشخص است که تعداد بسیار زیادی از گره های موجود در این انجمن دارای ویژگی btsexo هستند، همچنین از لحاظ impression این انجمن کاملاً یک دست هستند.

گروه ۴: شامل اکانت های فریبرز کرمی زند، BARANDAZOnline، AnonopsSE، javanmard75، anonymouseopiran، anonymouseenn3 آن میکنند مانند (فریبرز کریمی)

گروه ۵: این گروه نیز از لحاظ ویژگی Impression بسیار یکدست است و کاربر معروفی داخل آن وجود ندارد اما با بررسی محتوی اکانت ها اکثراً کاربران عادی (غیر معروف) و روزمره نویس هستند که البته بعد از شروع اعترافات محتوی اعتراضی پخش می کند.

گروه ۶: این گروه نیز یک گروه کاملاً یک دست است، با بررسی محتوای اکانت این گروه متوجه می شویم که این افراد پشتیبان حکومت و ارزشی هستند.

## پاسخ ۵ – خوشه بندی با استفاده از Infomap

توضیحات پیاده سازی:

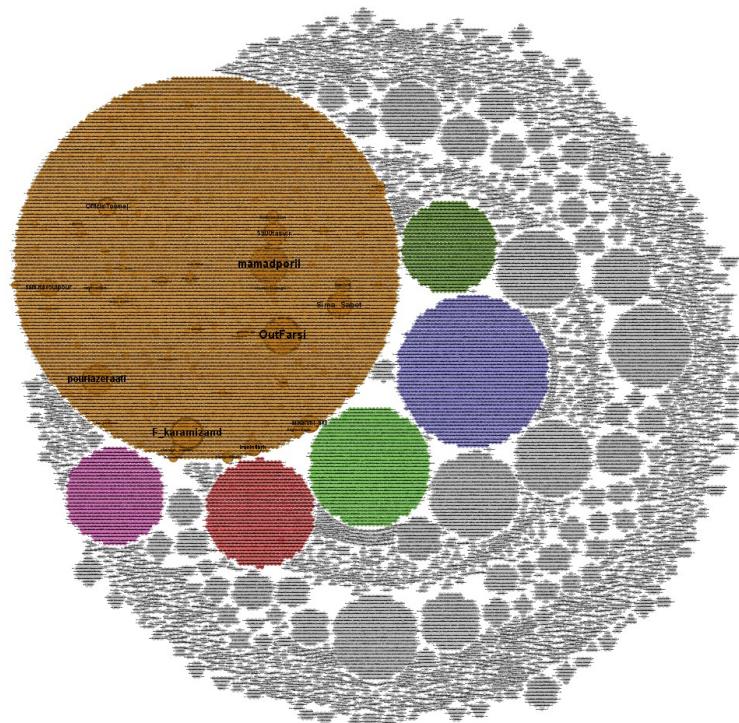
با توجه به اینکه نرم افزار گفی و کتابخانه networkx هیچ کدام از الگوریتم Infomap پشتیبانی نمیکنند برای پیاده سازی این الگوریتم از کتابخانه infomap استفاده کردیم.

برای پیاده اجرای الگوریتم ابتدا گراف را در کتابخانه networkx بارگذاری کردیم و سپس با استفاده از کتابخانه infomap خوشه بندی کردیم، در نهایت شماره خوشه هر نود را به عنوان یک ویژگی در گراف ذخیره کردیم و گراف را برای بررسی دقیق تر به نرم افزار گفی منتقل کردیم.

```
1 G = nx.read_graphml('/content/complete.graphml')
2 info_map = Infomap(silent=True,two_level=True)
3 node_to_id = info_map.add_networkx_graph(G)
4 communities = {}
5 info_map.run(num_trials=10)
6 for node in info_map.nodes:
7     communities[node_to_id[node.node_id]] = node.module_id
8 nx.set_node_attributes(G,communities,'infomap')
9 nx.write_graphml(G,'./infomap.graphml')
```

شکل ۲۵: کد اجرای الگوریتم Infomap

نحوه خوشه بندی الگوریتم Infomap به شکل زیر است:



شکل ۲۶: خوشه بندی گره های شبکه با الگوریتم Infomap

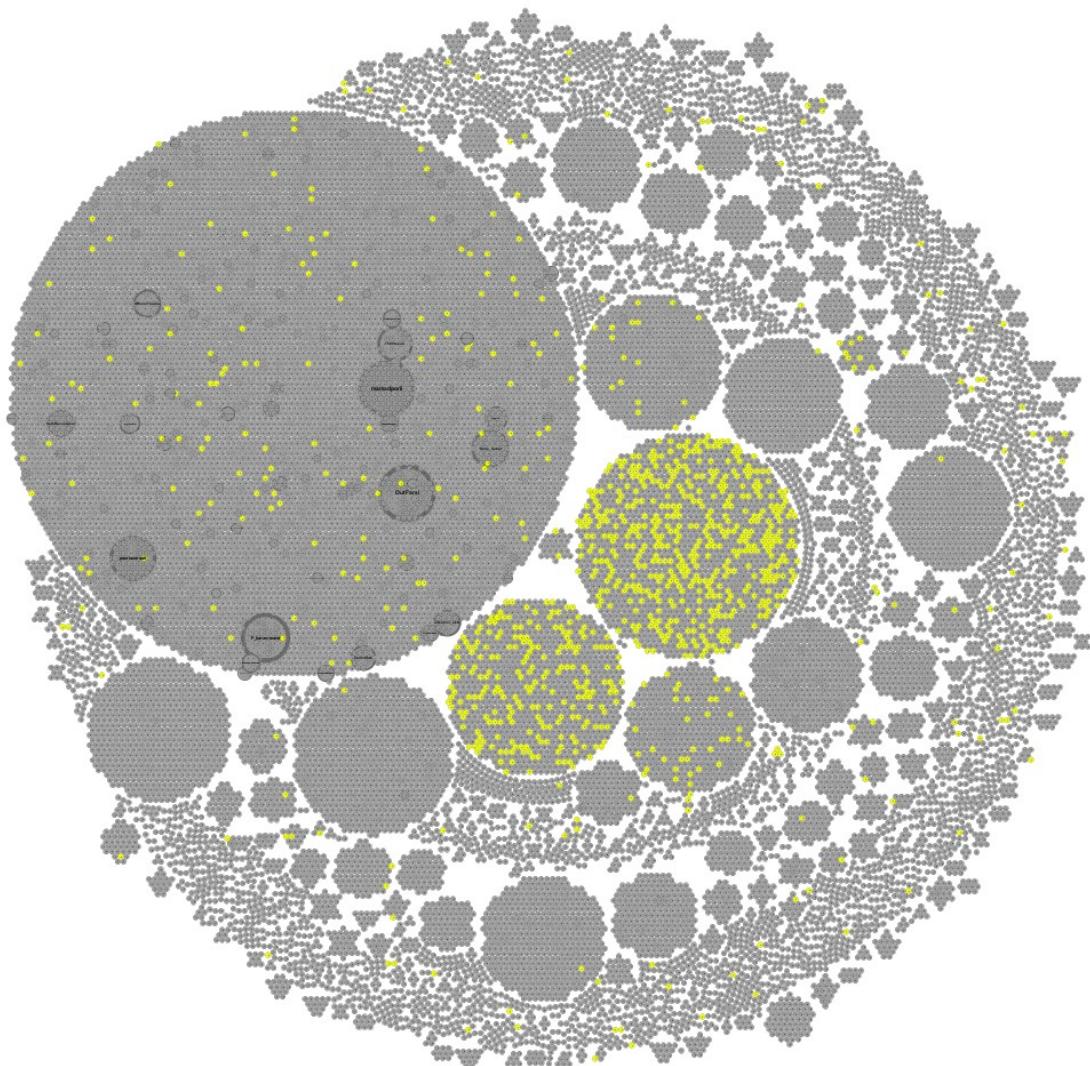
گروه ۱ - گروه آجری رنگ: بزرگترین گروه شبکه شامل گره های اصلی مخالف حکومت مثل ایران اینترنشنال، کرمی زند، علی کریمی، outfitarsi، سیما ثابت، mamadporii، 1500tasvir، Godofpersian، tomaj salhi و... است این گروه بیشترین تعداد جمیعت را دارد و ماهیت اصلی آن براندازی رژیم جمهوری اسلامی است.

گروه ۲ - گروه صورتی رنگ: این گروه حامیان حکومت از جمله khamenei\_fa، raefipour و... هستند Farsnews\_agency.

گروه ۳ - گروه قمز رنگ: حامیان سلطنت طلب از قبیل رضا پهلوی و اکانت های سلطنت طلب

گروه ۴ - گروه سبز روش رنگ: طرفداران BTS و Exo

گروه ۵ - گروه بنفش رنگ: این گروه هم اکثراً از طرفداران BTS و Exo تشکیل شده است بنابرین ترغیب شدیم تا تعداد کاربران btsexو را در خوشه بندی بررسی کنیم:

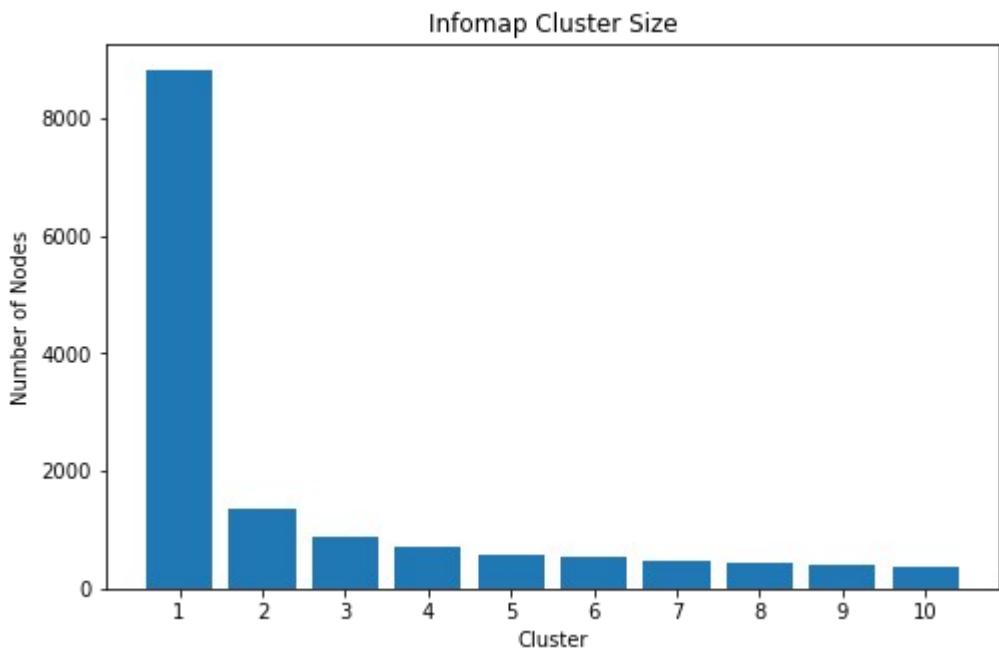


شکل ۲۷: توزیع کاربران با ویژگی **btsexo** در خوشه بندی **Infomap**

همانطور که انتظار می‌رفت پس از خوشه بندی Infomap کاربران **btsexo** در دو گروه ۴ و ۵ پخش شده‌اند اما دلیل اینکه این دو خوشه از هم‌دیگر جدا شده‌اند به وضوح مشخص نیست با بررسی دقیق‌تر این دو خوشه نیز به تفاوت معنا داری در محتوی و ماهیت اکانت‌های این دو خوشه بر نخوردیم.

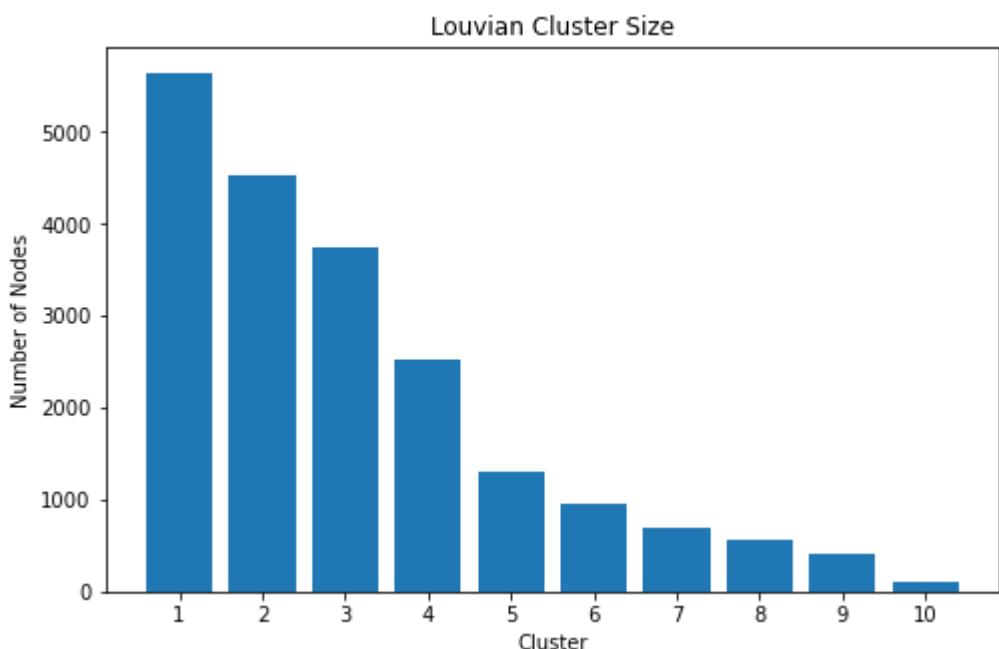
گروه ۶ - گروه سبز پرنگ: با بررسی کاربران با **Impression** بالای این گروه مشاهده می‌شود که اکثراً عکس پروفایل مربوط به انیمه و مانگا‌های ژاپنی دارند گرچه شاخص **btsexo** در خوشه خیلی بالا نیست، هر دو الگوریتم **Infomap** و **Louvian** اتفاق نظر مشترکی برروی این کاربران داشته‌اند و کاربران این خوشه در خوشه بندی **Louvian** هم اکثراً در یک خوشه قرار دارند، محتوی پست‌های این کاربران تشویق به اعتراضات، آموزش اعتراض و... است و اکثراً به نظر می‌رسد جزو **genZ** باشند.

برای بررسی ۱۰ انجمن هر روش خوشه بندی ابتدا به بررسی آمار خوشه ها می پردازیم:  
در ابتدا اندازه خوشه ها را بررسی می کنیم:



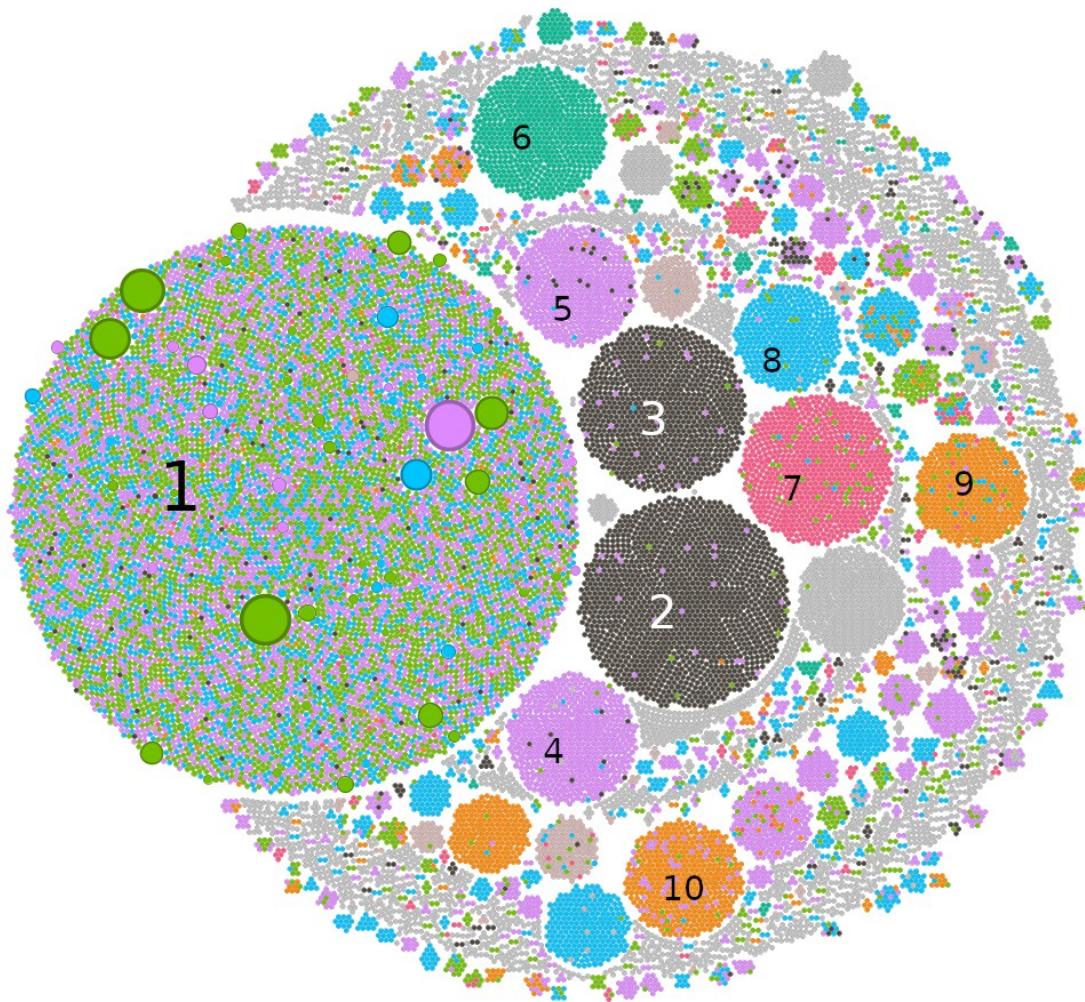
شکل ۲۸: سایز خوشه های الگوریتم Infomap

الگوریتم Infomap یک خوشه با اندازه بسیار بزرگ و مابقی خوشه هارا به نسبت بسیار کوچک و نسبتاً هم اندازه ساخته است.



شکل ۲۹: اندازه خوشه بندی ها با الگوریتم Louvian  
الگوریتم Louvian اما خوشه هارا به شکل کاملاً یک دست نزولی تقسیم کرده و تفاوت خوشه اول با مابقی خوشه ها مانند الگوریتم قبلی زیاد نیست.

در قسمت اول سؤال اما متوجه شدیم که در برخی از خوشه ها بین دو الگوریتم اتفاق نظر بسیار زیادی وجود داشته و گره ها در هر دو الگوریتم در یک خوشه قرار گرفته اند اینبار برای بررسی دقیق تر یک نمودار جدید رسم می کنیم، خوشه بندی گره ها را با استفاده از الگوریتم Infomap رسم می کنیم اما رنگ بندی گره ها با استفاده از شماره هی خوشه Louvian است.

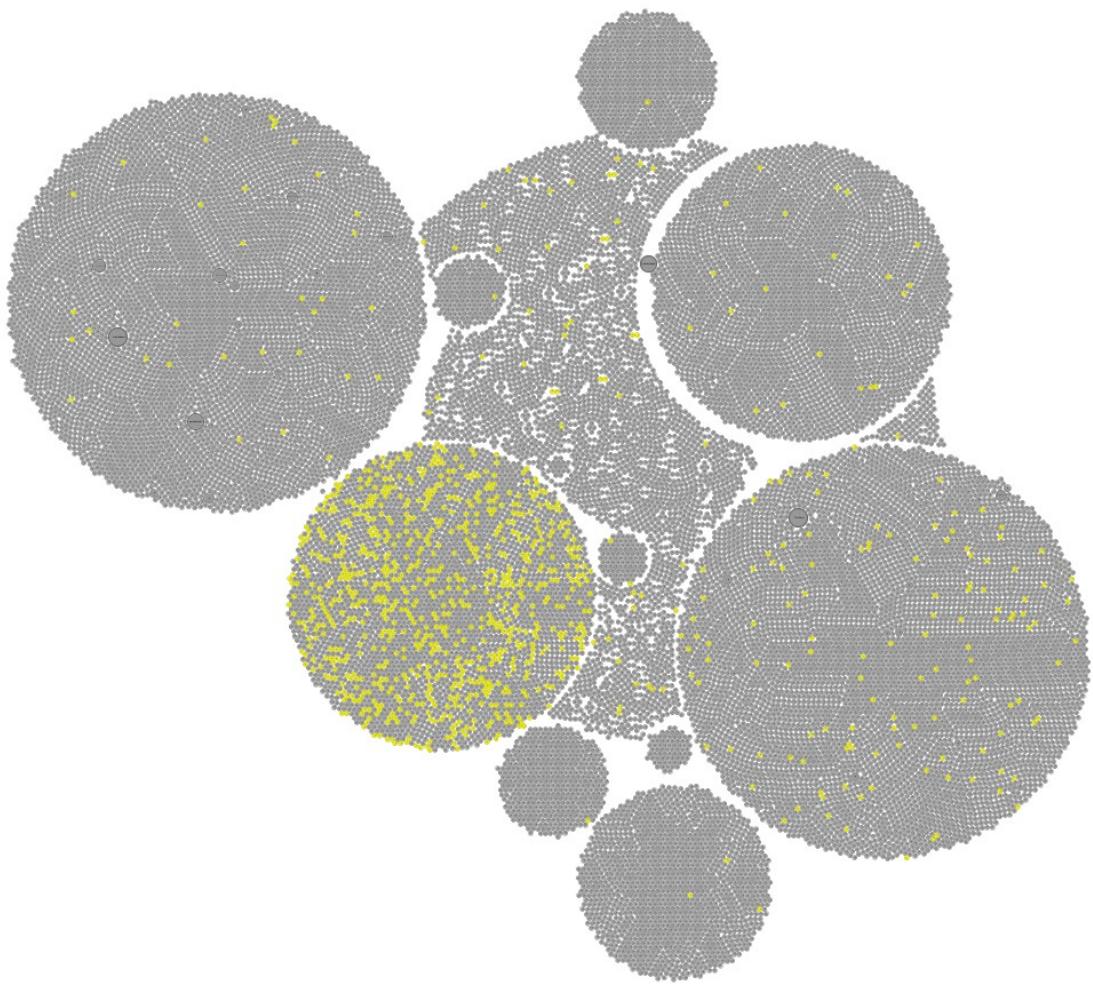


شکل ۳۰: ترکیب خوشه بندی Louvian با رنگ بندی Infomap

با توجه به شکل خوشه ۱ به طور کلی ترکیبی از خوشه بنش و سبز و آبی در Louvian است پس الگوریتم Louvian این ۳ خوشه را جدا از هم در نظر گرفته اما آنها را یکی در نظر گرفته است، از طرفی خوشه های ۲ و ۳ که در Infomap از یکدیگر جدا شده اند در Louvian یک خوشه در نظر گرفته شده است که از لحاظ محتوایی نیز این دو خوشه بسیار نزدیک به هم بودند، گروه ۴ و ۵ نیز در Infomap جدا اما در Louvian یکی در نظر گرفته شده اند، با بررسی نود های با Impression بالا گروه ۵ به این نکته جالب برخور迪م که اکثراً طرفداران گروه Onedirection بودند اما این موضوع در گروه ۴ دیده نمیشد. خوشه شماره ۶ حامیان حکومت، شماره ۷ سلطنت طلبان و نزدیکان رضا پهلوی، شماره ۸ طرفداران عرفان حلقة! و محمدعلی طاهری هستند که هر دو الگوریتم در این گروه ها اتفاق نظر دارند، خوشه ۹ مجاهدین خلق و نزدیکان مریم رجوی بودند، در گروه ۱۰ اما جهت گیری خاصی دیده نشد.

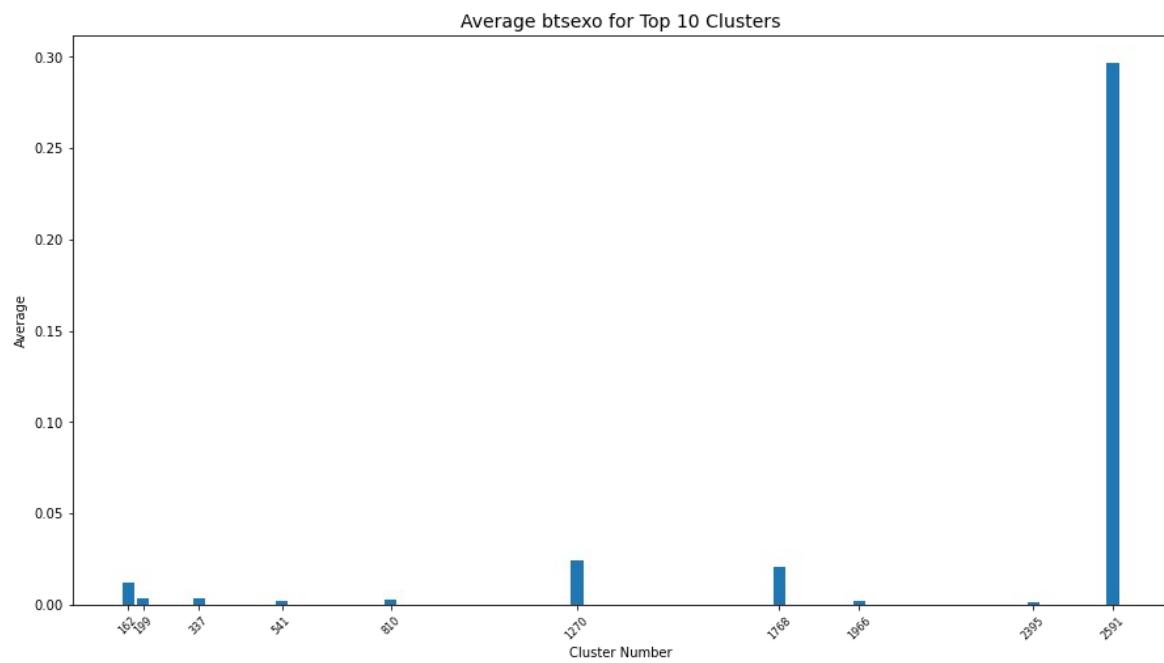
## پاسخ ۶ - بررسی ویژگی و BTSexo

در ابتدا به بررسی ویژگی Louvian در خوشه بندی btsexو می پردازیم:



شکل ۳۱: توزیع طرفداران **BTSexo** در خوشه بندی **Louvian**

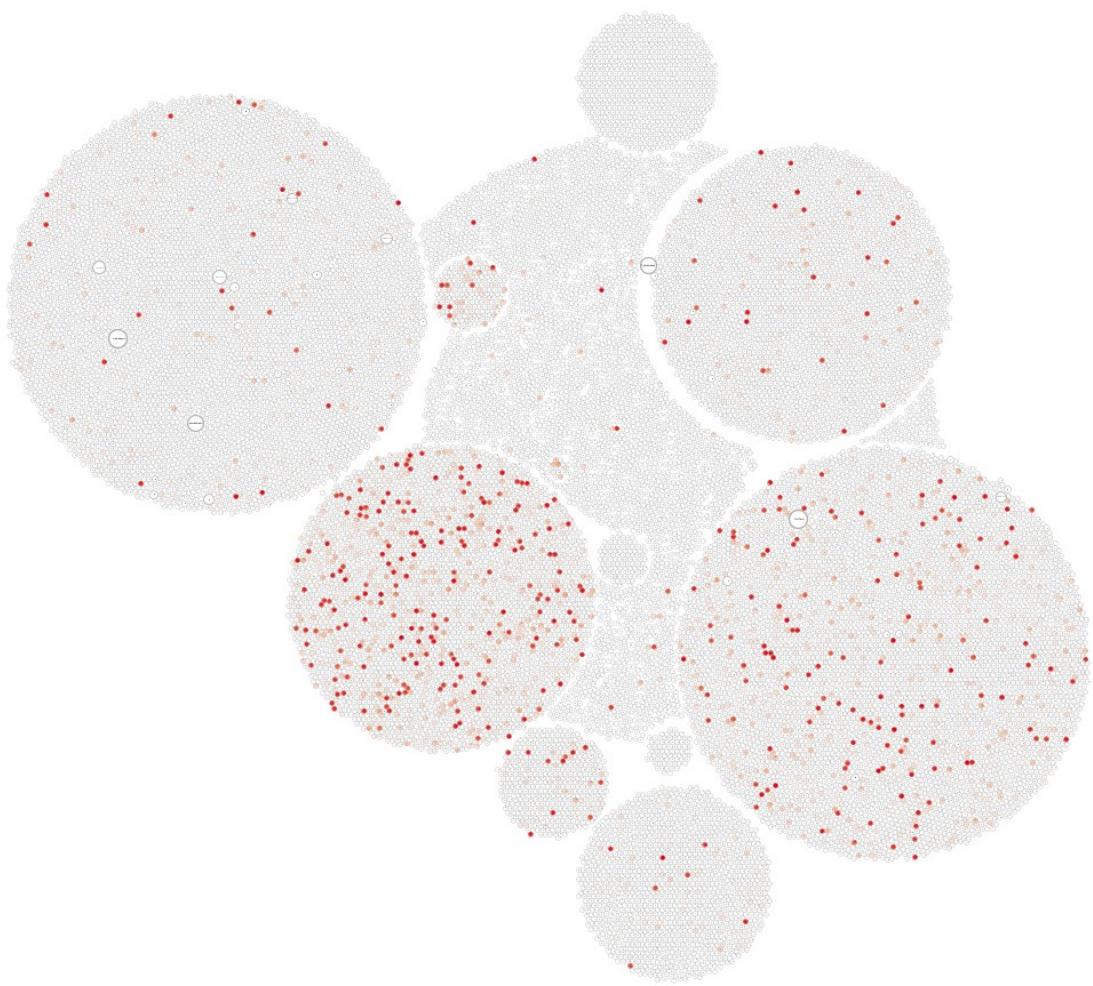
همانطور که از شکل مشخص است کاملاً واضح است که طرفداران این گروهها در یک گروه بسیار بیشتر هستند، حال به بررسی میانگین این ویژگی در ۱۰ انجمان بزرگ می پردازیم:



شکل ۳۲: میانگین مقدار ویژگی **btsexo** در هر انجمن

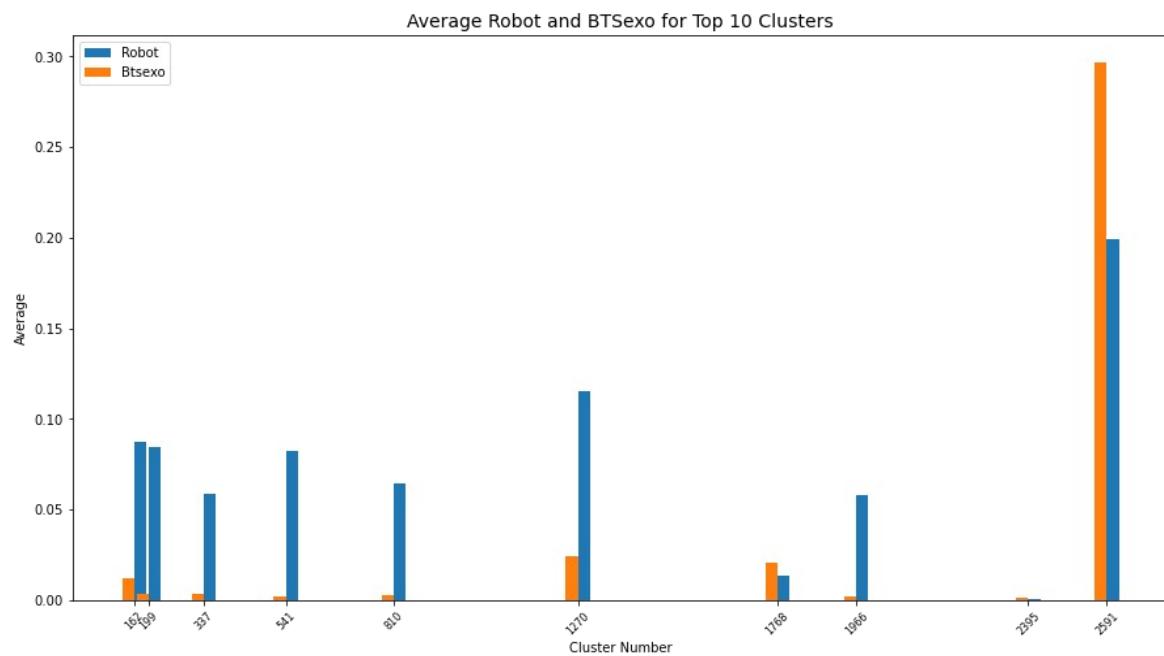
مقدار میانگین **btsexo** در انجمن ۲۵۹۱ از مابقی کلاسترها به شکل معناداری بیشتر است این کلاستر مربوط به انجمن معتبرضانی بود که اکثراً طرفداران bts بودند، انجمن بعدی ۱۲۷۰ که اکانت معروفی چون Godofpersian و outfarsi، myfootkin نمیداده اند و صرفاً پس از شروع اعتراض‌ها شروع به فعالیت سیاسی کرده اند، انجمن بعدی ۱۷۶۸ مربوط به گروه ریستارتی‌ها و محمد حسینی مجری سابق تلویزیون هستند، انجمن بعدی ۱۶۲ که مربوط به ایران اینternشناال و مجریان و نزدیک آن هست که به شکل حرfe ای فعالیت سیاسی انجام می‌دهند.

با استفاده از رنگ دهی Ranking بر اساس ویژگی **robot** و تابع spline سیگموئید گره های با احتمال بالاتر **robot** را رنگ قرمز تر می دهیم.



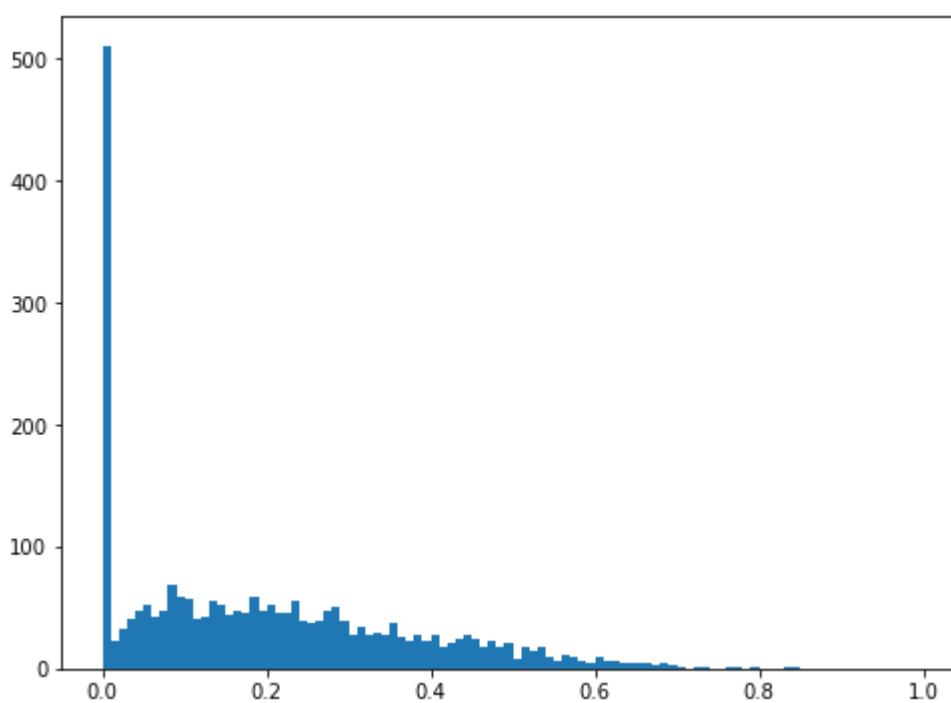
شکل ۳۳: توزیع مقدار ویژگی **Robot** در گره های هر انجمن

با توجه به شکل به نظر می رسد که انجمن ۲۵۹۱ که بیشترین میزان btsexo را داشت از نظر احتمال ربات هم مقادیر بالاتری دارند، برای بررسی دقیق تر نمودار میانگین را برای ۱۰ انجمن بزرگ بر اساس احتمال ربات بودن رسم می کنیم:



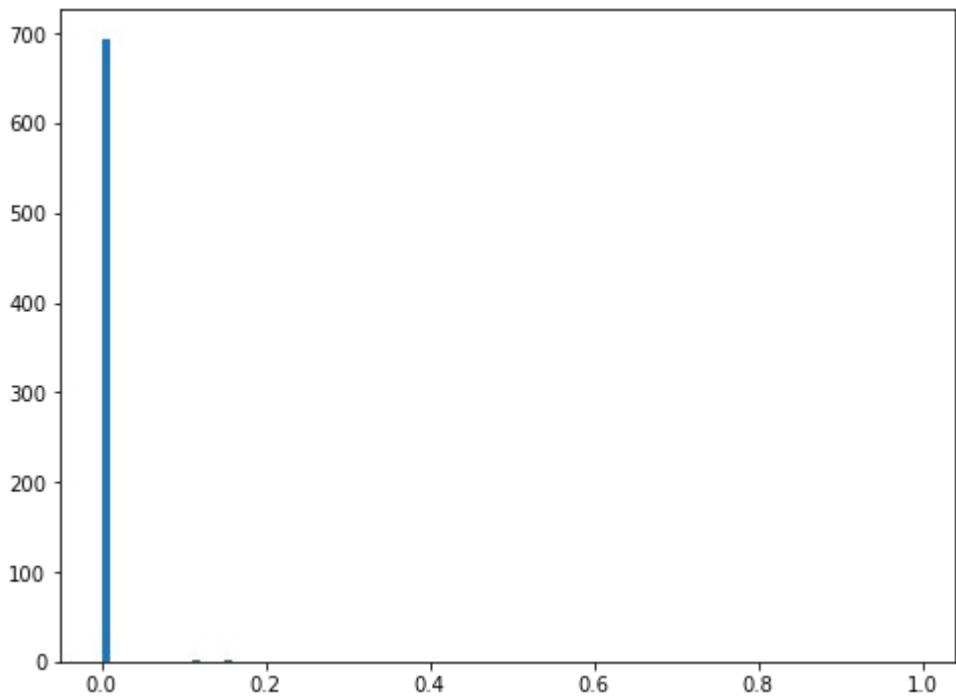
شکل ۳۴: میانگین احتمال ربات بودن و **btsexo** در ۱۰ انجمن برتر

به نظر می‌رسد که در انجمن ۲۵۹۱ که مربوط به انجمن طرفداران گروه bts است میانگین احتمال ربات بودن نیز به شکل غیر طبیعی بالاتر است اما به طور کلی نمیتوان رابطه‌ی مستقیمی میان bts و احتمال ربات بودن در همه انجمن‌ها مشاهده کرد. انجمن ۱۲۷۰ که در مقام دوم از لحاظ میانگین احتمال ربات بودن قرار گرفته مربوط به انجمنی بود که کاربران myfootkin و outfarsi و godofpersian و ... در آن قرار داشتند و به طور کلی مربوط به کاربرانی بود که قبل از اعتراضات به طور حرفه‌ای کار سیاسی انجام نمی‌دادند، انجمن ۱۶۲ مربوط به ایران اینترنشنال و نزدیکانش، انجمن ۱۹۹ مربوط به javanmardi75 و تجزیه طلبان بود، انجمن ۵۴۱ نیز مربوط به مجاهدین خلق و نزدیک مریم رجوی است.



شکل ۳۵: هیستوگرام توزیع مقادیر احتمال ربات بودن در کلاستر ۲۵۹۱

با توجه به این نمودار تعداد زیادی از کاربران احتمال ربات بودنشان صفر است اما مقدار زیادی از کاربران نیز مقادیر بیشتر از صفر و بیشتر از  $0.5$  دارند برای درک زیاد بودن این مقدار هیستوگرام کلاستر ۲۳۹۵ که کمترین میزان btsexo را داشت را نیز رسم می‌کنیم:

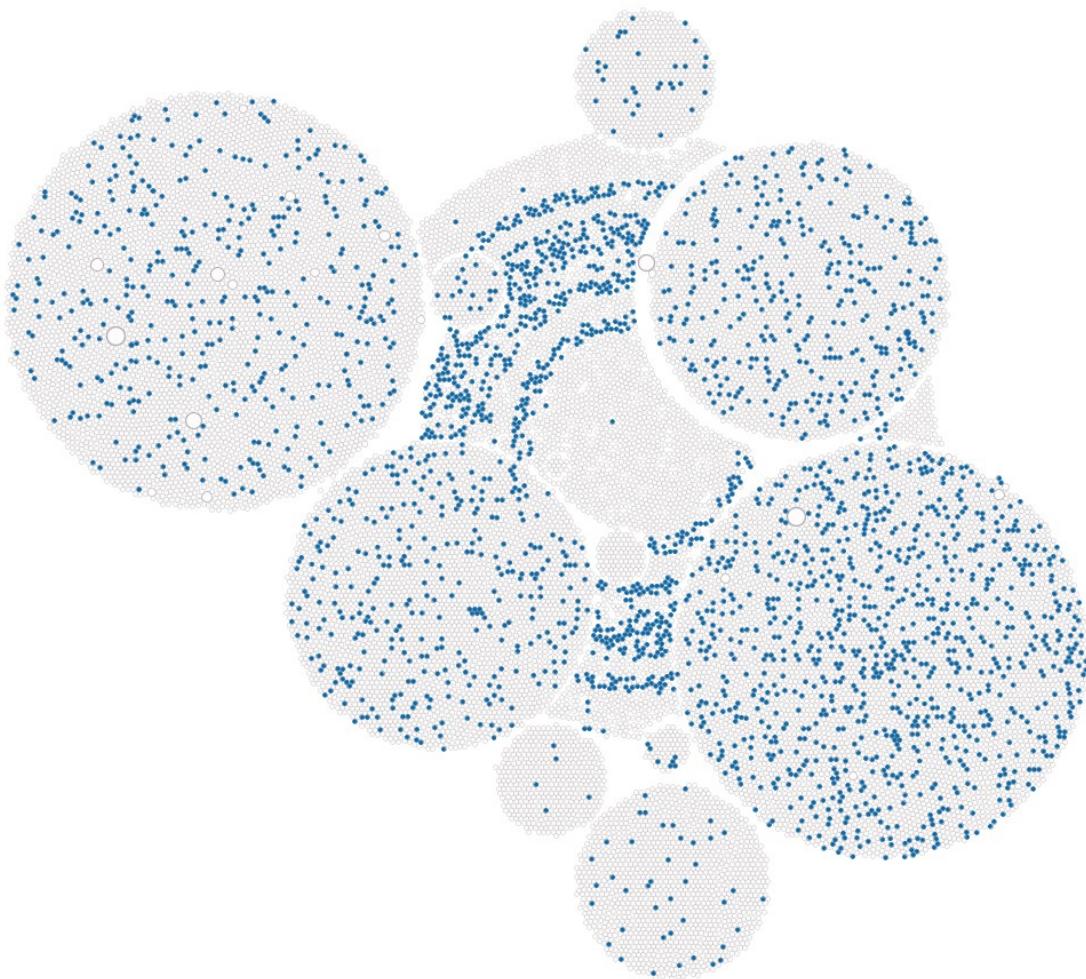


شکل ۳۶: توزیع مقادیر احتمال ربات در کلاستر ۲۳۹۵

همانطور که مشخص است احتمال تقریباً تمام کاربران صفر است بنابرین میتوان گفت این تفاوت در کلاستر های معنا دار است.

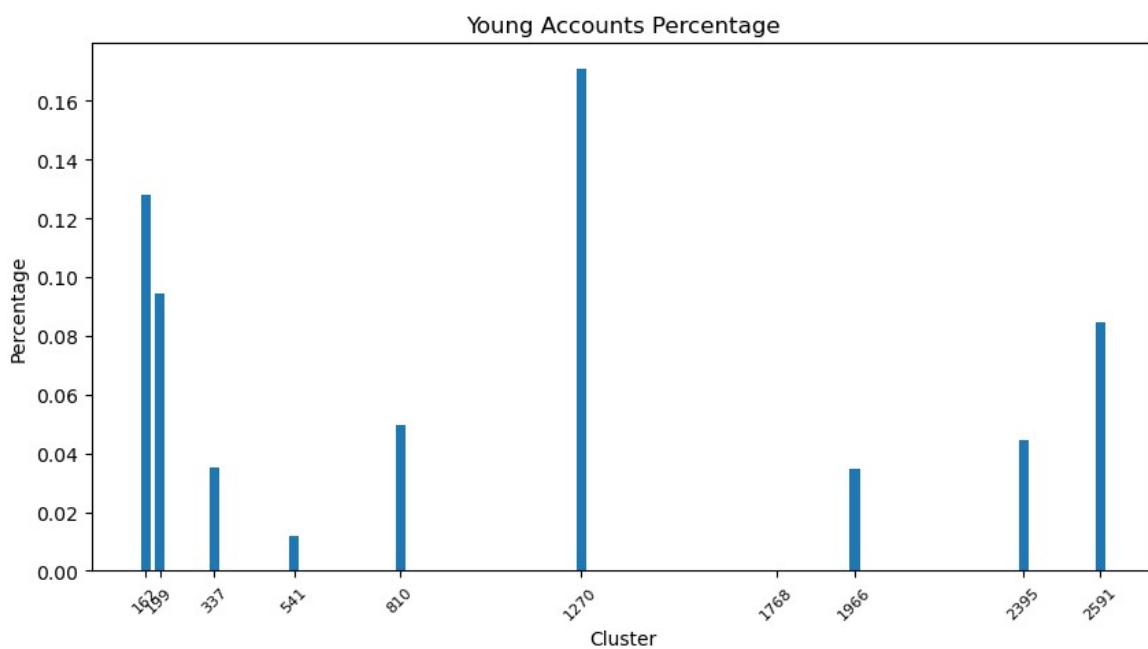
## پاسخ ۷ - بررسی سن اکانت‌ها

برای مشاهده اکانت‌هایی که سن پایینی دارند تمام گره‌های شبکه را سفید کردیم و سپس کاربرانی که از ۶۴ روز کمتر سن داشتند به نسبت هرچه کم سن‌تر بودند آبی پررنگ‌تر کردیم نتیجه را در عکس ۳۳ مشاهده می‌کنید.



شکل ۳۷: توزیع کاربران کم سن در گره‌های شبکه

با توجه به شکل مشخص است که توزیع کاربران کم سن به طور کلی در انجمان‌های بزرگ حدوداً یکسان است اما نکته متفاوت در این باره این است که کاربرانی که جزء انجمان به خصوصی نبودند تعداد کاربر کم سن زیادی دارند. برای بررسی دقیق‌تر بهتر است نسبت کاربران کم سن (کمتر یا مساوی با ۶۴) به کل تعداد افراد را در ۱۰ انجمان اول پیدا کنیم:



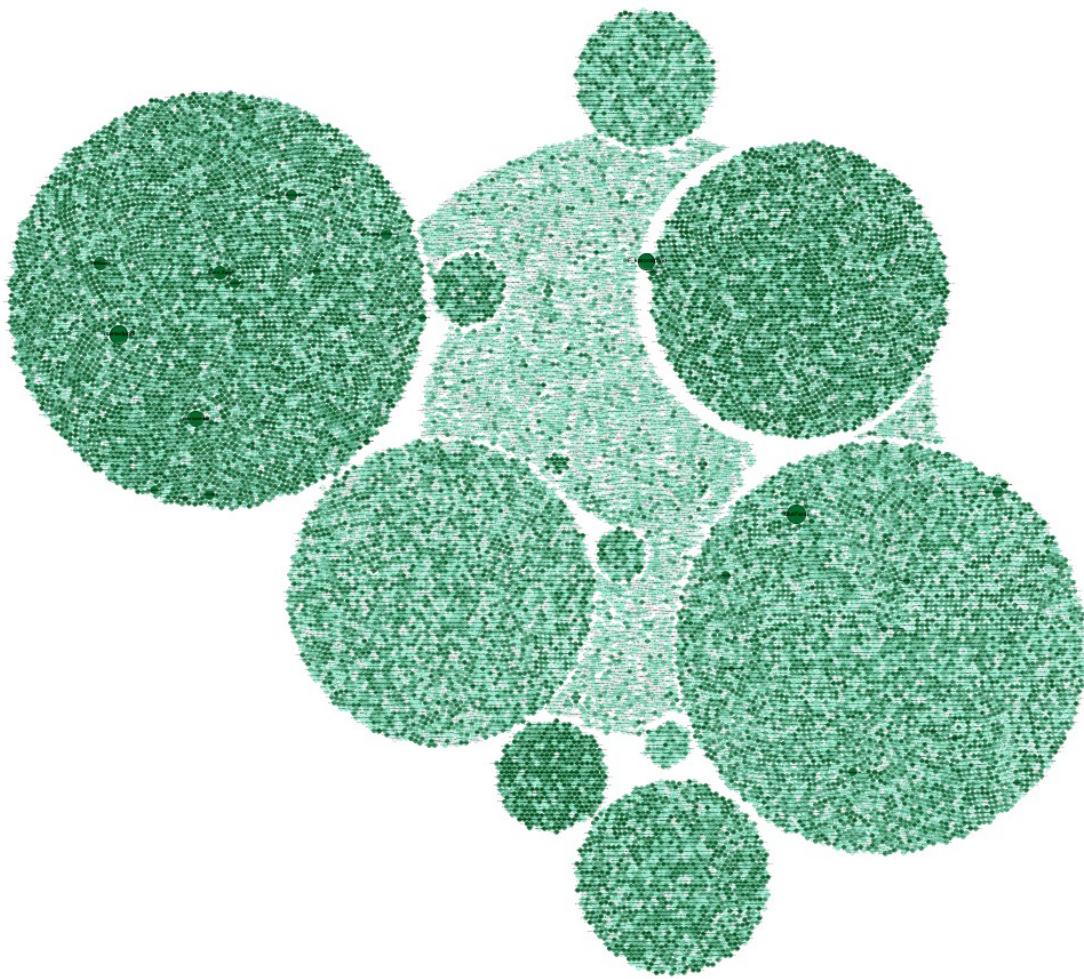
شکل ۳۸: نسبت اکانت های با عمر کم نسبت به کل اکانت ها در ده انجمن برتر

با توجه به نمودار بیشترین نسبت در انجمن ۱۲۷۰ وجود دارد که مربوط به انجمنی است که کاربرانی که فعالیت غیر حرفه ای سیاسی انجام می دهند و اکثراً بعد از اعتراضات شروع به فعالیت سیاسی کرده اند است، در این انجمن اغلب خوانندگان، یوتیوبرها و چهره های سلبریتی نسل جدید و روزمره نویسان وجود دارند.

بعد از آن انجمن ۱۶۲ که مربوط به فعالان حرفه ای سیاسی است مانند اکانت های ایران اینترنشنال و مجریان آن مثل سیما ثابت و سام رسول پور و پوریا زراعتی است بیشترین نسبت کاربران جوان را دارد و سپس انجمن ۱۹۹ که از اعضای آن علی جوانمردی، محمدعلی طاهری، کانون حقوق بشر ایران و می باشند، انجمن ۲۵۹۱ که مربوط به طرفداران bts بود نیز در ردی قرار دارد.

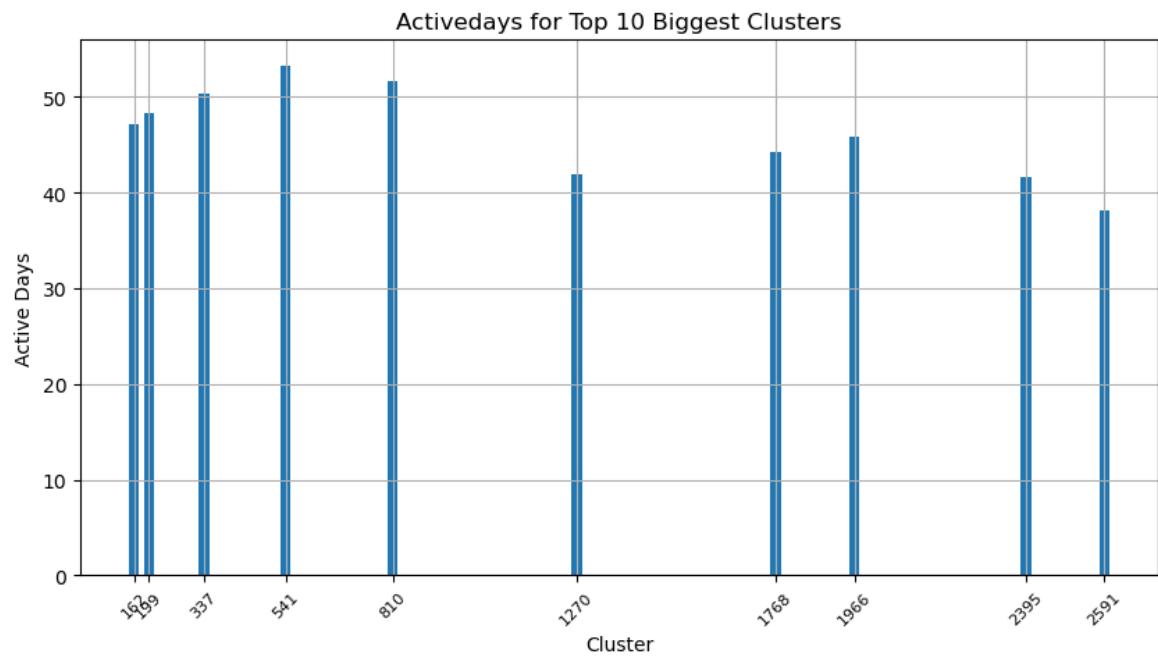
## پاسخ ۸ - بررسی میزان فعالیت کاربران

برای نشان دادن کاربران با ranking بالا رنگ گره هارا به شکل activedays بر اساس مقدار ستون در آوردهیم:



شکل ۳۹: میزان فعالیت گره های گراف

با توجه به شکل قابل مشاهده است که انجمن های اصلی شکل گرفته همگی activedays نسبتاً بالایی دارند همچنین گره های با Impression بالاتر (گره های بزرگ تر) همگی فعالیت بسیار بالایی داشته اند. برای بررسی دقیق تر میانگین فعالیت هر کدام از انجمن هارا برای ۱۰ انجمن برتر رسم می کنیم:



شکل ۴۰: میزان فعالیت هر کدام از انجمن های بزرگ

با توجه به نمودار بیشتر میزان فعالیت در انجمن ۵۴۱ است که مربوط به انجمن مجاهدین خلق و نزدیکان مریم رجوی است، بعد از آن انجمن ۸۱۰ مربوط به سلطنت طلبان و رضا پهلوی بیشترین فعالیت را داشته‌اند، انجمن بعدی ۳۳۷ مربوط به ایران اینternشناو و اکانت های نزدیک به آن است.