## به نام خدا

تمرین چهارم شبکههای اجتماعی

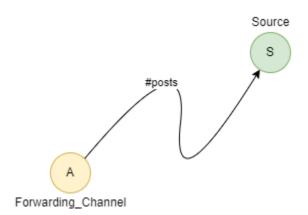
تشخیص و تحلیل انجمن (موضوع برجام در کانالهای تلگرامی)

محمد ناصري

11.1.44

مقدمه

در این تمرین شما باید با استفاده از ابزارهای تحلیل شبکه مثل Inetworkx انجمن های شبکه ضمیمه را تحلیل نمایید. شبکه ضمیمه، شبکه ارتباطی بین کانالهای تلگرامی است که طی یکی دو ماه اخیر درباره برجام مطلب منتشر کرده اند. یالهای این شبکه جهت دار است و جهت آنها از کانال فوروارد کننده مطلب به کانال مرجع رسم شده است. ستونهای مربوط به نودها عبارتند از شناسه یکتای کانال (name) شناسه انگلیسی کانال (label) تعداد مطالب مربوط به برجام(posts) جمع بازدید مطالب اصلی (نه فورواردی) منتشر شده توسط آن کانال (weight) و تعداد فالوئرهای کانال (followers) هستند. ستون وزن (weight) هر یال هم تعداد دفعاتی مطلبی کانال node2 و انشان میدهد.



سوال اول

? لیست ده کانال برتر از منظر معیارهای Authority ، Hub و Authority ، و Degree In را بیابید و تفسیر کنید اشتراکها و اختالفهای این 4 لیست به چه دلیل است؟

بررسی برترینهای In-Degree

Id	Label	posts	views	followers	indegree	outdegree	Degree
1004128493	FarsNewsInt	318	776723	18680	2315	22	2337
1005665139	isna94	195	864839	83263	2122	29	2151
1068487996	irna_1313	193	118005	0	1477	37	1514
1069291200	chandsanieh	209	521493	0	1372	64	1436
1006914438	khabaronline_ir	337	1104816	212358	1281	57	1338
1005280116	entekhab_ir	354	4044199	96667	1265	42	1307
1001976122	Tasnimnews	104	987663	1220150	1240	20	1260
1007884158	MyAsriran	293	984997	108955	1093	71	1164
1445349980	Damunnnn	393	306890	0	1082	66	1148
1002739327	yjcnewschannel	111	959890	778718	1067	28	1095

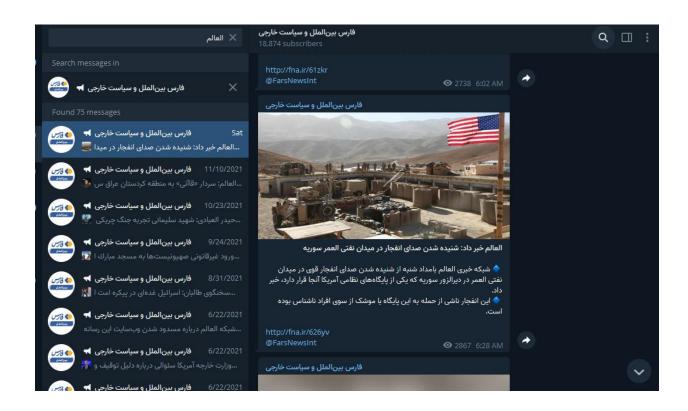
کانالهایی که تعداد درجه ورودی بالایی دارند از لحاظ تئوری کانالهایی هستند که تعداد کانال زیادی از مطالب این کانالها استفاده کرده و آنها را فوروارد میکنند. با بررسی این کانالها متوجه میشویم که کانالهای خبرگزاریهای بزرگ و رسمی کشور هستند مانند: فارس، ایسنا، ایرنا، انتخاب، تسنیم و ... که به طبع اخبار روز سریعتر در این کانالها منتشر میشود و کانالهای دیگر از این مطالب استفاده و آنهارا کپی یا فوروارد میکنند. همچنین اکثریت این کانالها از لحاظ تعداد view جزء ۲۰ کانال برتر از لحاظ view وبازدید نیز میباشند (که این مقدار بازدید مربوط به پستهای غیرفورواردی است). از طرفی با مشاهده مقدار تعداد post مربوط به برجام، این کانالها جزء ptها محسوب نمیشوند. دلیل این امر را با توجه به محتوا میتوان بدین صورت عنوان کرد که از آنجا که این کانالها، مراکز رسمی خبری هستند، تنها اخبار موثق و رسمی و با تعداد کمتری منتشر میکنند.

بررسی برترینهای page rank

Label	posts	views	followers	indegree	outdegree	Degree	pageranks
FarsNewsInt	318	776723	18680	2315	22	2337	0.022678
isna94	195	864839	83263	2122	29	2151	0.01803
irna_1313	193	118005	0	1477	37	1514	0.01644
Damunnnn	393	306890	0	1082	66	1148	0.015291
alalamfa	201	104708	15073	854	38	892	0.013504
Tasnimnews	104	987663	1220150	1240	20	1260	0.012103
entekhab_ir	354	4044199	96667	1265	42	1307	0.011437
farsna	57	1056889	1309495	1061	9	1070	0.011122
roozArooz_media	55	234884	0	885	5	890	0.01083
MyAsriran	293	984997	108955	1093	71	1164	0.010728

ایده پشت page rank این است که هر گره می تواند با داشتن یک لبه به گره های دیگر، به اهمیت سایر گره ها رأی دهد. گره هایی با درجه (نسبی) بالا مهمتر از گره هایی با درجه پایین یا اصلاً بدون درجه هستند. همچنین، آرائی که از گرههای مهم می آیند، ارزش بیشتری نسبت به رایهای گرههای نه چندان مهم دارند. رتبه بندی صفحات بدین معناست که اگر کاربر از یک صفحه یا کانال شروع به مرور کند و بصورت تصادفی به لینکهای موجود در صفحه برود چقدر احتمال دارد به یک صفحه یا کانال خاص برسد. بصورت کلی اهمیت صفحات (کانال) را مشخص میکند.

دلیل تفاوت page-rank نسبت به معیار فوق،تعریف متفاوت رتبهبندی صفحات است. همانطور که گفته شد Page-rank الگوریتم Page-rank به صفحاتی امتیاز بالاتر میدهد که به نودهای با امتیاز بالا متصل باشد. از این رو یک گره با درجه ورودی کمتر ممکن است امتیاز بالاتری نسبت به یک گره با درجه ورودی بالاتر داشته باشد. برای مثال chandsanie دارای درجه ورودی بالاتری نسبت به alamfa میباشد ولی از امتیاز page-rank پایین تری برخوردار است. زیرا برای مثال از مطالب العالم در کانال farsnewsint استفاده شده که امتیاز PR بسیار بالایی دارد(تصویر صفحه بعد)



به همین نحو باقی تفاوتهای بین رتبهبندیهای درجه ورودی و PR قابل توجیه میباشند.

## بررسی برترینهای Authority

Label	posts	views	followers	indegree	outdegree	Degree	pageranks	Authority
FarsNewsInt	318	776723	18680	2315	22	2337	0.022678	0.307559
isna94	195	864839	83263	2122	29	2151	0.01803	0.28586
irna_1313	193	118005	0	1477	37	1514	0.01644	0.234016
chandsanieh	209	521493	0	1372	64	1436	0.009262	0.196198
entekhab_ir	354	4044199	96667	1265	42	1307	0.011437	0.194525
Tasnimnews	104	987663	1220150	1240	20	1260	0.012103	0.184922
yjcnewschannel	111	959890	778718	1067	28	1095	0.008109	0.181419
Damunnnn	393	306890	0	1082	66	1148	0.015291	0.177085
khabaronline_ir	337	1104816	212358	1281	57	1338	0.009128	0.176122
MyAsriran	293	984997	108955	1093	71	1164	0.010728	0.168541

یک مرجع خوب نشان دهنده صفحه ای است که توسط هاب های مختلف لینک شده است. با توجه به تعاریف جهت یالها که برای این گراف داریم برای تفسیر مراجع میتوان آنها را کانالهایی در نظر گرفت که مطالب را بین دیگر کانالها پخش و توزیع میکنند و باقی کانالها مطالب را از این کانالها فوروارد میکنند. همانطور که مشاهده میشود جدل بدست آمده دقیقا با جدول In\_degree مطابقت دارد که این امر با توجه به تعاریف دو معیار مربوطه قابل پیشبینی بود. همانطور که گفته شد این کانالها، مراکز رسمی خبری هستند که خصوصیات آنها شامل تعدد بالای استفاده دیگر کانالها از مطالب آنها زیاد است، تعداد پست مربوط به برجام پایین تری نسبت به تعداد زیادی از کانالها دارند و اخبار رسمی تر و دقیق تری منتشر میکنند و با این حال بازدید بالایی دارند. در نهایت بعنوان مراجع خبری در میان کانالهای مورد بررسی در نظر گرفته میشوند.

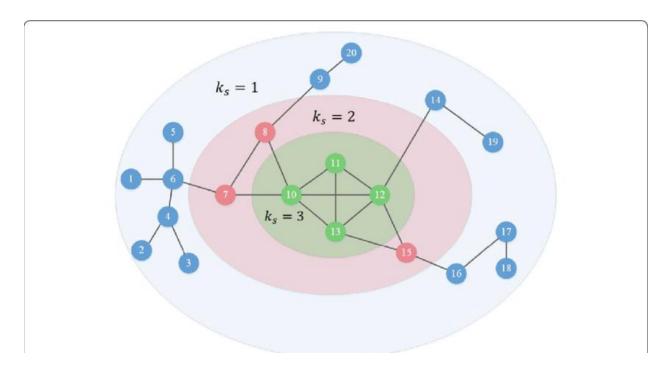
بررسی برترینهای Hub

Label	posts	views	followers	indegree	outdegree	Degree	pageranks	Authority	Hub
BigTrader24	754	1314	0	48	168	216	0.000311	0.008677	0.065962
khalerize2	395	698	0	41	127	168	0.000236	0.010182	0.064212
goldteknik1201400	553	24851	0	229	180	409	0.002634	0.04443	0.063748
GOLDBEIGLOO	481	22086	0	258	140	398	0.001554	0.050649	0.061316
Khabardun1	338	383	0	376	108	484	0.003921	0.074939	0.060799
Myasrkhabar	377	906	2090	94	109	203	0.000984	0.021406	0.059448
khabarhayemohem14	653	17	0	49	156	205	0.000158	0.00636	0.058727
twiier_news	435	19181	0	188	123	311	0.001218	0.038656	0.058431
economicsdayfacts	255	6	0	8	102	110	0.000194	0.001617	0.056548
koolakmag	238	0	0	0	99	99	0.000017	0	0.055652

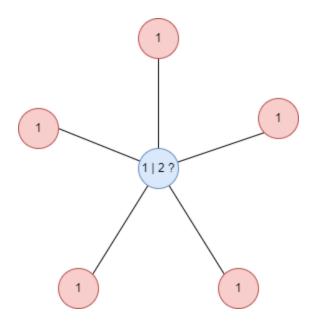
یک hub خوب نمایانگر صفحه ای است که به بسیاری از صفحات دیگر اشاره می کند، با توجه به تعاریف جهت یالها که برای این گراف داریم میتوان تعریف هاب را اینگونه تفسیر کرد که کانالهایی هستند که بیشتر از باقی، پیام از کانالهای دیگر فوروارد کرده و بیشتر از بقیه از مطالب کانالهای دیگر استفاده میکنند. از خصوصیات این کانالها با توجه به مقادیر جدول میتوان به تعداد پست بالای آنها اشاره کرد. این کانالها محتوای بیشتری نسبت به کانالهای دیگر در حوزه برجام منتشر کردهاند و از طرفی از لحاظ out-degree نیز جزء کانالهای برتر محسوب میشوند. از دو مورد فوق و بررسی محتوایی مشخص میشود که این کانالها، اخبار را از کانالهای دیگر کپی و منتشر میکنند و بعلت استفاده از کانالها و منابع متفاوت تعداد اخبار منتشر شده در این حوزه توسط این کانالها بیشتر از بقیه است.

سوال دوم

یکی از راههای درک گراف این است که بررسی گرهها را با معیار کمترین اتصال شروع کنیم. K-core یک گراف حداکثر زیرگراف است به طوری که هر راس دارای درجه حداقل k باشد. پوسته k مجموعه ای از رئوس است که بخشی از هسته k هستند اما بخشی از هسته k نیستند.

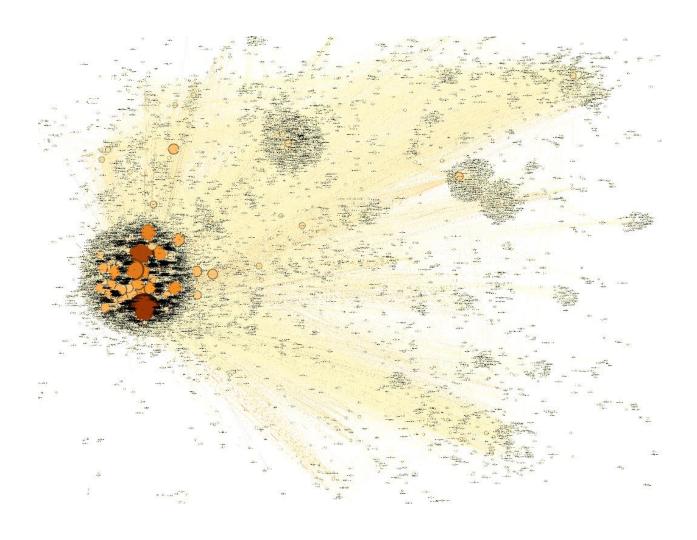


به عنوان مثال، هسته 0 یک گراف به سادگی کل نمودار است زیرا هر راس حداقل دارای صفر یال است. یک راس نمی تواند تعداد یال های منفی داشته باشد. هسته 1 گراف شامل رئوس هایی است که به رئوس دیگر متصل هستند. شما با دور انداختن پوسته 0، مجموعه ای از رئوس جدا شده، هسته 1 را تشکیل می دهید. با حذف گره های درجه 1 از 1 هسته به 2 هسته می روید تا زمانی که هر چیزی که باقی مانده است حداقل درجه 2 داشته باشد. شما به سادگی با دور انداختن گره های درجه کمتر از 1 هسته 1 را تشکیل نمی دهید. برای مثال، نمودار ستاره ای زیر را در نظر بگیرید

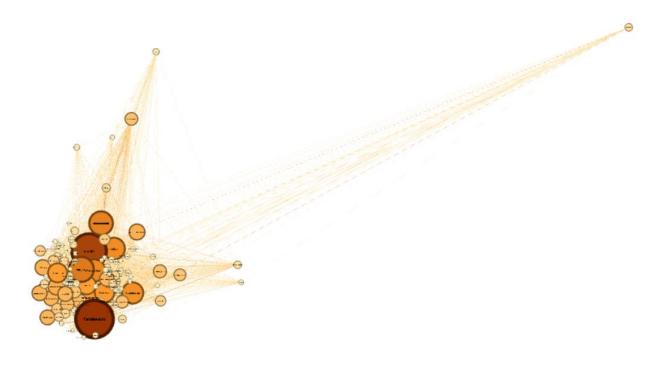


اگر گره های درجه 1 را جدا کنیم، در نهایت چیزی برای ما باقی نمی ماند. هیچ 2 هسته ای وجود ندارد، حتی اگر گراف اصلی دارای یک گره با درجه 5 بود هم به همین صورت بود. یا اگر به تعریف برگردیم، هیچ زیرگرافی وجود ندارد که هر راس دارای درجه 2 باشد.

حال به بررسی گراف مورد نظر خود میپردازیم. پس از حذف جهت از روی یالهای گراف شکل بدست آمده به صورت زیر است:



با حذف جهت از گراف، الگوریتم k-coreرا روی این دیتاست اجرا نموده و k-shell الگوریتم k-coreرا روی این دیتاست اجرا نموده و k-shellرا در کل گراف دارند چه ویژگیهای موضوعی یا محتوایی داشتهاند پس از اعمال فیلترینگ روی رئوس گراف به نتایج زیر دست میابیم:



Id	Label	timeset	posts	views	followers
1002062924	Khabar_Fouri		382	13240810	992777
1005280116	entekhab_ir		354	4044199	96667
1007007698	akharinkhabar		220	3201202	2801299
1005866359	Eslahatnews		167	1200565	118346
1006914438	khabaronline_ir		337	1104816	212358
1002738390	Eghtesadnews_com		453	1068352	119442
1006939659	farsna		57	1056889	1309495
1001976122	Tasnimnews		104	987663	1220150
1007884158	MyAsriran		293	984997	108955
1002739327	yjcnewschannel		111	959890	778718
1005665139	isna94		195	864839	83263
1013336404	shafaqna_farsi		126	840166	210696
1004128493	FarsNewsInt		318	776723	18680
1001170974	serfan_jahate_ettela		147	702107	1275796
1003329464	tsepress		358	668535	100028
1126629649	khabarfouri		206	523925	0
1069291200	chandsanieh		209	521493	0
1002249912	Khabare_vije		302	477432	2003317
1496273363	bourseonshows		339	468305	0
1154252621	Nournews_ir		137	424043	0
1426783083	namehnewss		147	423781	0
1007883087	eghtesadonline		438	382970	117595

1116567042	akhbar_montakhab	149	330242	0
1027676885	eslahaat press	169	308149	116854
1445349980	Damunnnn	393	306890	0
1050921798	ibenanews	186	249490	0
		55		0
1086482465	roozArooz_media		234884	
1006263686	rajanews_com	156	217921	98512
1004793780	iran_times	168	194237	118919
1074482831	betasahm1	196	173938	0
1006227643	jamarannews	270	166470	25009
1006212930	www_snn_ir	136	153990	28023
1233226069	mazene	95	138775	0
1068487996	irna_1313	193	118005	0
1008702726	alalamfa	201	104708	15073
1074370071	rasadkhabar	128	103694	0
1443574986	poshtpardeaslli	176	96604	0
1002483751	irannewspaper	221	92388	13447
1002765601	mehrnews	106	86649	70158
1221666621	JamejamDaily	48	77940	0
1138354631	BourseKhabar_IR	184	77791	0
1062062882	simasiasi	68	76784	0
1377594379	AJPersian	24	74165	0
1061186078	radinnews	293	73227	0
1497446976	farzadpp22	124	67849	0
1006005588	smart_investor	176	66874	92530
1131008271	Avayyeiran	97	59805	0
1001156199	alef_news	345	59473	11124
1497319226	snaptal1399pik	175	58375	0
1367035239	Indypersian	62	58049	0
1251576339	irinn_channel	35	53570	0
1143390513	Jahan_Fouri	226	49701	0
1064973232	perancees	216	48785	0
1004576149	iribnews	49	41662	71200
1346187186	Ustadzarifvip	258	34586	0
1324689269	etemadsaz_ir	215	30341	0
1007714672	bisto4news	158	27432	282775
1107633871	mostaghelnewspaper	59	26616	0
1048391437	sobhe_no	123	26206	0
1004899453	irankhabar	137	24870	37172
1491259826	goldteknik1201400	553	24851	0
1058667053	farhikhteganonline2	103	22711	0
1489832130	GOLDBEIGLOO	481	22086	0

1285183655	twiier_news	435	19181	0
1127390421	EtemadOnline	245	17989	0
1561020337	MFAIran	245	16141	
1576927172	moshaveremal	311	15278	0
1005338558	jahan_24	129	13935	2392
1331110775	purson_ir	154	11757	0
1191729070	goldteknik1201399	276	11559	0
1005036683	movassagh SaMassoumi	128	10824	5687
1130196213		79	8917	77010
1016512199	Telnews	236	8653	77818
1192060328	khodro_irani_021	167	7977	0
1007542449	dana_news	85	7819	108033
1196112447	harattahlil	157	7503	0
1349835325	iran_usa_Akber	205	7271	0
1215355989	akhbaarejahan	193	6916	0
1099300086	jahantasnim	56	6834	0
1012687880	oitahlil	151	6222	5292
1006058128	tabanweekly	124	6054	2826
1009750873	siasimashhad	233	5003	10281
1524703348	segnalharat	197	3343	0
1229976663	polmarket	198	3188	0
1002904936	khabaredagh	72	2529	41336
1286006168	revajnews	477	2072	0
	hamshahriazarbayjansharghi	218	1863	0
1018944837	basiji_313	441	1489	757
1386452125	signaalsabz	342	1329	0
1082356308	boushehriha	294	1326	0
1299980473	BigTrader24	754	1314	0
1168763012	hemmati_fan	172	1287	0
1205628632	dolaat	88	1264	0
1030399326	hamiyannevelayat	213	1062	9911
1033556721	ABG1340	281	1055	1141
1088270190	Minoudarweekly	193	984	0
1030252762	irsputnik	128	939	279
1006273170	Myasrkhabar	377	906	2090
1427040320	khalerize2	395	698	0
1162268257	SNIPER_IRANII	205	589	0
1281811805	IranOnline_News	369	450	0
1304849023	toolou021	152	447	0
1000445007				
1006445097	newsroom1	91	438	3779

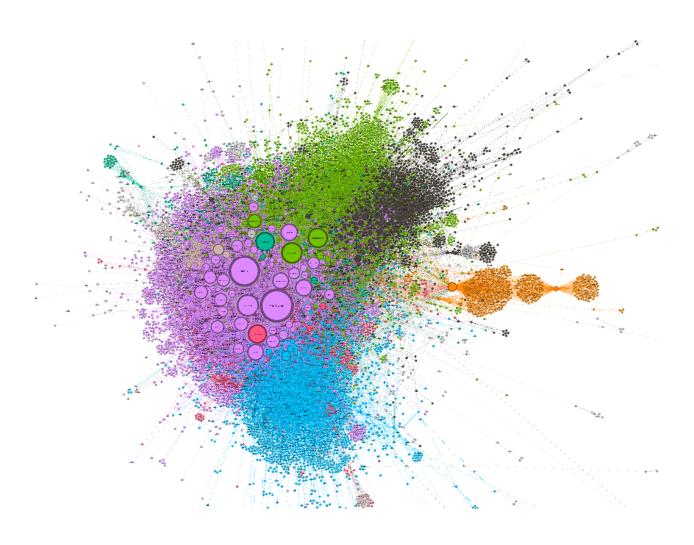
1090542605	iranresa	171	405	0
1266511682	Khabardun1	338	383	0
1592823356	akhbarefuri	245	347	0
1258109136	moghavemat_roozplus	119	314	0
1416744500	aryaee115	165	295	0
1053269707	ertebatpr2azwest	151	279	0
1485731336	fardaenews	155	263	0
1386985395	vipheratgorup	50	239	0
1481578235	JBCnews	379	238	0
1493948974	elmiokhabari	160	151	0
1237236083	Behzad_News_Network	143	131	0
1102047090	gsozlarei	187	111	0
1093216501	farsbourse	191	87	0
1251741394	mehrnema	166	74	0
1009194190	kresaneh	277	38	336
1035099694	sorooshekhabar	51	27	187
1158949652	ToseemeliKhorasanRazavi	287	26	0
1462815219	khabarhayemohem14	653	17	0
1192370795	IRAAN_Bourse	322	9	0
1300140040	economicsdayfacts	255	6	0
1005926885	eranico_com	311	0	66427
1006845767	irantopnews	201	0	622241
1149847754	koolakmag	238	0	0
1382953364	farhikhtegan_vahdat	153	0	0
1406220734	khabare8_ir	173	0	0
1478116506	MrStockExchange_Group	198	0	0

پس از فیلترینگ با مشاهده شبکه و گرهها و توضیحات داده شده پیرامون k-core مشاهده و نتیجه میشود که گرههای باقیمانده، بالاترین درجه را بین دیگر گرهها دارند و از همین رو در لیست برترینهای ذکر شده در قسمت اول تمرین نیز جای دارند. یعنی این گرهها خصوصیاتی مانند درجه بالا(in & out)، Authority یا Hub بالا و PageRank بالا دارند.

این بدین معنیست که با بررسی گرههای باقیمانده نتیجه میشود این گرهها نمایانگر کانالهایی هستند که یا خود از اعتبار و رسمیت بالایی برخوردار هستند و جزء مراکز خبری محسوب میشوند و یا مطالب خود را از کانالهای معتبر کپی و فوروارد کردهاند. به صورت کلی با رجوع به کانالهای این مجموعه، احتمال دریافت خبر رسمی و موثق نسبت به بقیه بالاتر است.

سوال سوم

الگوریتم Louvainرا برای استخراج انجمنها روی گراف اجرا نمایید. پس از اجرای الگوریتم، انجمنهای بدست آمده را با layout مناسب بازنمایی و نتیجه را تحلیل کنید (احتمالا نام کانالهای هر انجمن بتواند دلیل تشکیل آن انجمن را برای شما مشخص نمایید.



modularity یکی از معیارهای ساختار شبکه ها یا گراف است. این معیار برای اندازه گیری قدرت تقسیم یک شبکه به ماژول ها (که گروه ها، خوشه ها یا جوامع نیز نامیده می شوند) طراحی شده است. شبکههای با

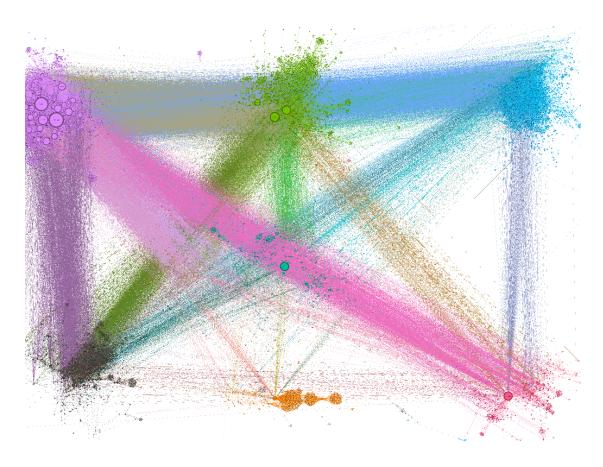
modularity بالا دارای اتصالات متراکم بین گرههای درون ماژولها هستند. modularity اغلب در روش های بهینه سازی برای تشخیص ساختار جامعه در شبکه ها استفاده می شود.

با این حال، نشان داده شده است که modularity دارای محدودیت هایی است و بنابراین، قادر به تشخیص جوامع کوچک نیست. برای مثال شبکه های بیولوژیکی، از جمله مغز حیوانات، درجه بالایی از modular بودن را نشان می دهند.

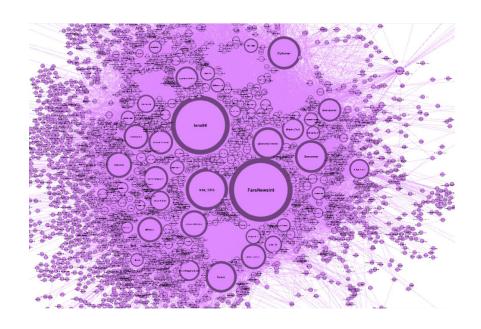
Gephi از روش Louvain برای محاسبه ماژولار استفاده می کند و یک نکته قابل توجه این است که modularity بالا به معنای پارتیشن "خوب" نیست.

در مرحله اول به بررسی انجمنها بزرگ شبکه میپردازیم. این انجمن ها در شکل زیر به صورت تفکیک شده نمایش داده شدهاند.

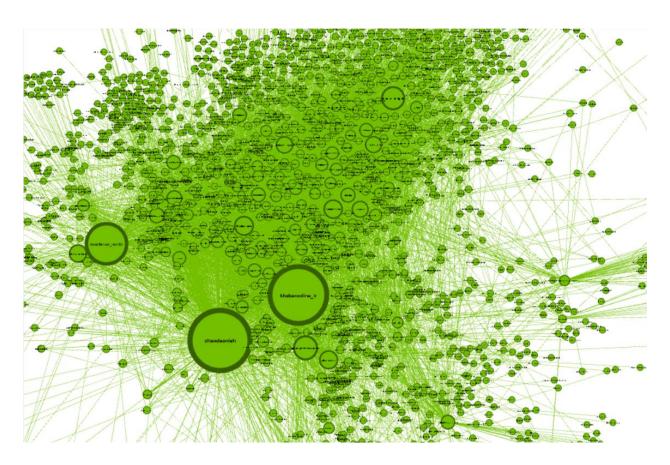
برای نمایش این داده، ابتدا ابزار مربوط به محاسبه modularity در گفی را اجرا میکنیم و سپس توسط فیلترهای مربوط به attribute و قسمت partition آنها را مرتب میکنیم.



در انجمن صورتی رنگ در شکل زیر شاهد کانالهای خبرگزاریهای رسمی(اغلبا) داخلی(شامل کانال شبکههای تلویزیون، روزنامهها و پایگاههای خبری آنلاین) هستیم که دلیل interaction آنها با یکدیگر به طور واضح این است که اخبار را با همکاری همدیگر پوشش میدهند.

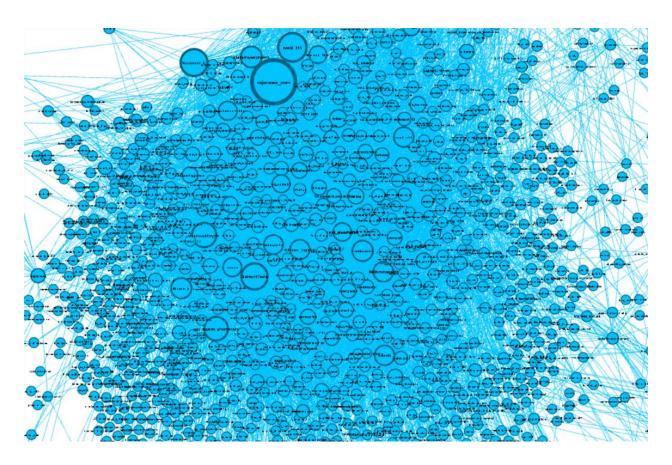


با بررسی شبکه و محتوای کانالهای انجمن بزرگ دوم که با رنگ سبز مشخص شده است، نتیجه میشود که این مجموعه شامل کانالهای خبری مستقل و بیطرف میباشند که هرگونه محتوای خبری را منتشر میکنند.



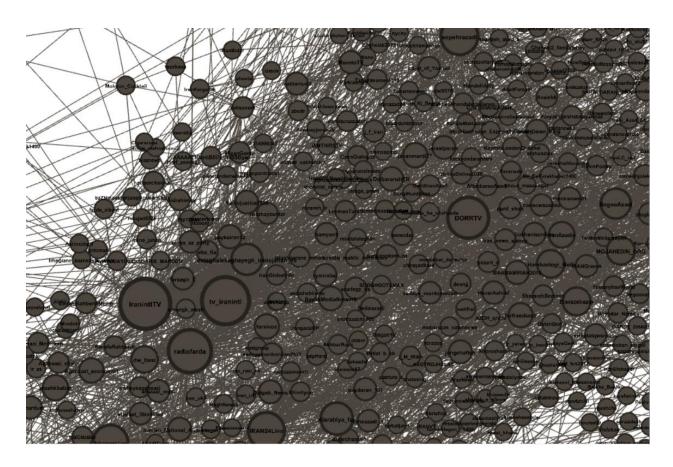
از این کانالها و رسانههای خبری مستقل میتوان به chandsanie, khabaronline\_ir, و ستقل میتوان به mostaghelnewspaper, tabnanweekly

با بررسی شبکه و محتوای کانالهای انجمن بعدی که با رنگ آبی مشخص شده متوجه میشویم این دسته (اغلبا) شامل کانالهای خبری و یا غیرخبری و با موضوع طرفداری و متعلق به طرفداران نظام هستند که با یکدیگر برای نشر محتواهای خود interaction دارند.



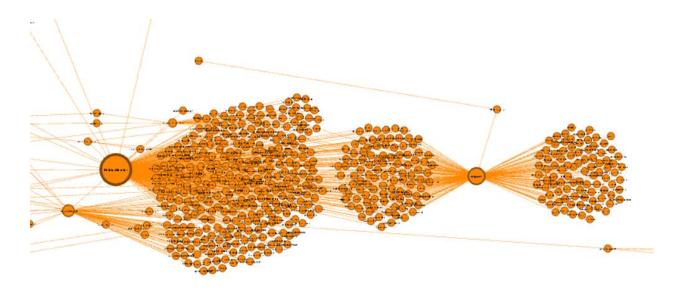
rajanews\_com, nournews\_ir, basij\_313, bisimchimediam, از این کانالها میتوان به jahanaraofoghtv

با بررسی شبکه و محتوای کانالهای انجمن مشکی رنگ موجود در گراف حاصل،به دست می آید که شامل کانالهای مخالف نظام و وابسته به شبکهها و پایگاههای خبری بیگانه خارجی میشود که محتوای مخالف و مغایر با انجمنهای پیشین دارند.

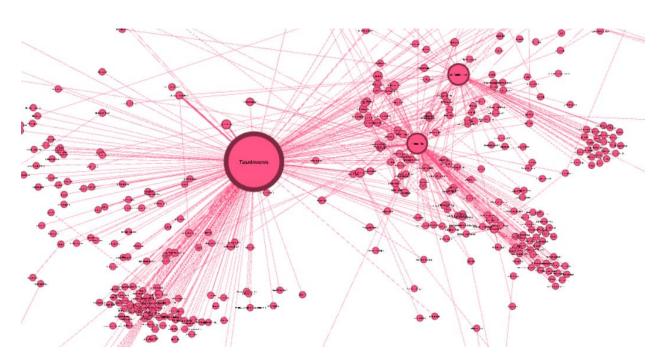


radiofarda, iranintITV, tv\_iranintl, iran24live, ddrtv, از این دسته کانالها میتوان به bbcpersian و ... اشاره کرد.

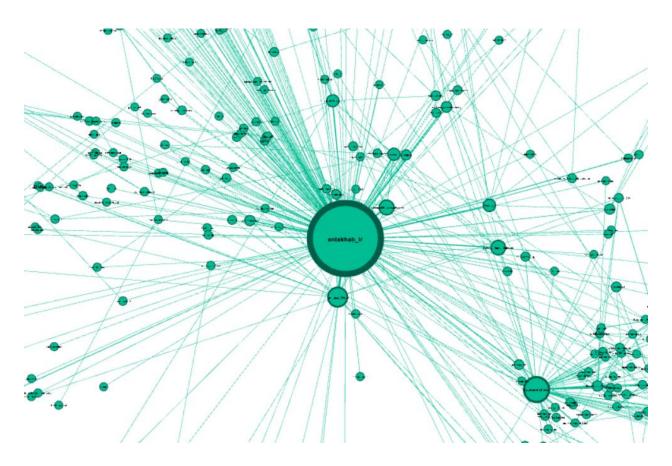
در این میان چند انجمن کوچک که کمتر از ۵ درصد داده را شامل میشوند وجود دارد که در ادامه به بررسی آنها میپردازیم. دسته بندی اول گروهی از کانالها هستند که محتوای خود را از کانالهایی که دسترسی مشاهده forwarded\_message



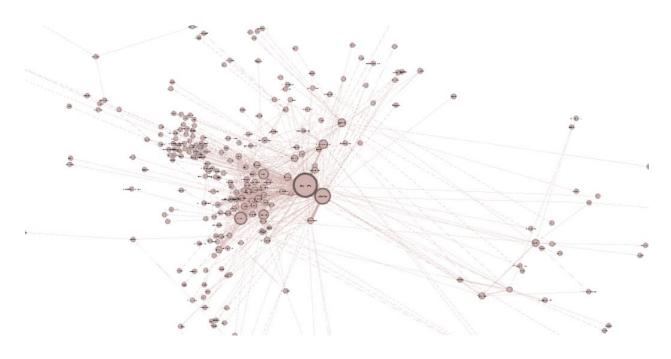
دستهبندی بعدی گروهی از کانالها هستند که با خبرگزاری تسنیم در ارتباط هستند و مطالب خود را از این کانال یا کانالهای وابسته مانند اخبارفوری و الفباخبر، تهیه میکنند



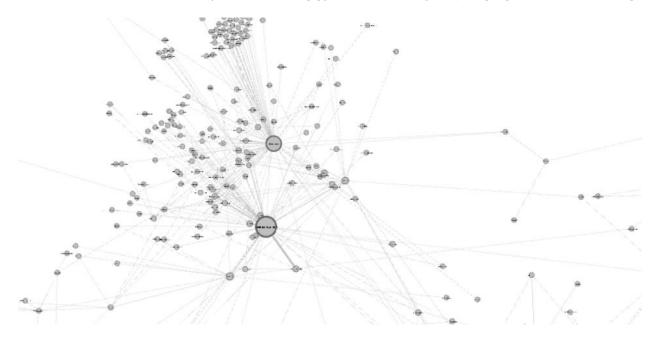
انجمن بعدی نیز مانند دسته بندی گذشته میباشد و کانالهایی را شامل میشود که مطالب از کانال پایگاه خبری انتخاب و یا کانالهای وابسته آن کپی برداری و فوروارد کرده اند.



انجمن بعدی نیز مانند دستهبندیهای گذشته میباشد و کانالهایی را شامل میشود که مطالب از کانال پایگاه خبری و اطلاع رسانی رواج و یا کانالهای وابسته آن کپی برداری و فوروارد کرده اند. همچنین این دسته شامل کانالهای مجموعه پُرسون نیز میشود



در یک دسته بندی دیگر نیز کانالهای وابسته به اقتصادنیوز را مشاهده میکنیم



انجمنهای دیگر بسیار کوچک و دارای درجات کم هستند و به بررسی آنها نمیپردازیم.

مورد بعدی که میتوان از شبکه تشکیل شده دریافت کرد ارتباطات انجمنهای مختلف خبری با یکدیگر است.

برای مثال بیشترین حجم ارتباط مربوط به کانالهای خبرگزاریهای رسمی داخلی میباشد و کمترین interaction میان انجمن کانالهای طرفدار جمهوری است. یا برای مثال interaction میان انجمن کانالهای طرفدار جمهوری اسلامی و مخالفان و شبکه های بیگانه نسبت به باقی ارتباطات ضعیفتر است.

پایان