



دانشگاه تهران
پردیس دانشکده‌های فنی
دانشکده مهندسی برق و کامپیوتر



شبکه‌های اجتماعی

پروژه نهایی

نام و نام خانوادگی
محمد ستار امامی میبدی

شماره دانشجویی
۸۱۰۱۰۴۰۷۲

۱۴۰۴ بهمن ۲۹

فهرست مطالب

۱	استخراج نودهای برتر و تحلیل محتوای آنها
۱	۱. هدف و منطق کلی سؤال
۱	۲. ساختار داده‌های شبکه و متن
۱	۲.۱. فایل شبکه GDF
۱	۲.۲. فایل‌های محتوایی Excel
۱	۳. روش اجرا
۲	۴. نتایج استخراج Top-5 برای پلتفرم twitter
۳	۵. نتایج استخراج Top-5 برای پلتفرم telegram
۴	۶. نتایج استخراج Top-5 برای پلتفرم instagram
۵	۷. همپوشانی اکانت‌های منتخب بین پلتفرم‌ها
۶	۸.۱. تحلیل محتوایی (بر اساس پست با بیشترین بازنشر) در پلتفرم Twitter
۹	۸.۲. تحلیل محتوایی (بر اساس پست با بیشترین بازنشر) در پلتفرم telegram
۱۰	۹.۱. تحلیل محتوایی (بر اساس پست با بیشترین بازنشر) در پلتفرم instagram
۱۰	۱۰.۱. پیشنهاد روش‌های ابتکاری مبتنی بر هوش مصنوعی
۱۵	۱۱.۱.۱. استخراج واژه‌های متمایز با TF-IDF
۱۶	۱۱.۱.۲. مدل‌سازی موضوعی با LDA
۱۶	۱۱.۱.۳. پیشنهاد پیشرفته‌تر: BERTopic
۱۷	۱۲. استخراج و تفسیر نودهای بر اساس Hub و Authority ، PageRank ، In-Degree
۱۷	۱.۲. تعریف گراف و پیش‌پردازش
۱۷	۲.۲. اندازه گراف‌های ساخته شده
۱۷	۳.۲. تعاریف ریاضی معیارها
۱۷	۱.۳.۲. درجه ورودی In-Degree
۱۸	۲.۳.۲. الگوریتم PageRank
۱۸	۳.۳.۲. الگوریتم HITS: امتیازهای Hub و Authority
۱۹	۴.۲. نتایج: ۱۰ نود برتر برای هر معیار

۴۵	۴	تشخیص انجمن با Louvain بر روی شبکه بازنگری
۴۲	۹.۳	شکل ها
۴۱	۸.۳	جمع بندی تفسیری: k-shell بالا چه معنایی دارد؟
۳۹	۴.۷.۳	کلیدواژه های شاخص نودهای پوسته بیشینه در Instagram
۳۸	۳.۷.۳	کلیدواژه های شاخص نودهای پوسته بیشینه در Telegram
۳۶	۲.۷.۳	کلیدواژه های شاخص نودهای پوسته بیشینه در Twitter
۳۶	۱.۷.۳	روش خواندن متن و محدودیت ها
۳۵	۶.۳	تحلیل محتوایی / موضوعی نودهای پوسته بیشینه
۳۴	۷.۳	نودهای پوسته بیشینه: 10 نود با بیشترین درجه در هر پلتفرم
۳۳	۵.۳	توزیع k-shell در هر پلتفرم (کوانتایل ها و تفسیر)
۳۳	۴.۳	نتایج عددی و مقایسه سه پلتفرم
۳۲	۳.۳.۳	گام 3: تحلیل محتوایی نودهای پوسته بیشینه با TF-IDF
۳۲	۲.۳.۳	گام 2: محاسبه k-shell برای هر نود
۳۱	۱.۳.۳	گام 1: تبدیل گراف جهت دار به گراف بدون جهت
۳۱	۳.۳	روش اجرای مسئله روی داده تمرین
۳۱	۴.۲.۳	چرایی تفاوت k-shell با «محبوبیت»
۳۱	۳.۲.۳	الگوریتم محاسبه (پوست اندازی / هرس تکراری)
۳۰	۲.۲.۳	تعريف core number و ارتباط آن با k-shell
۳۰	۱.۲.۳	تعریف k-core
۳۰	۲.۳	تعاریف دقیق و روابط ریاضی
۳۰	۱.۳	هدف مسئله و چرایی حذف جهت
۳۰	۳	۳ تحلیل k-core و k-shell پس از حذف جهت یال ها
۲۹	۳.۶.۲	نمونه واژه های شاخص در Instagram
۲۸	۲.۶.۲	نمونه واژه های شاخص در Telegram
۲۸	۱.۶.۲	نمونه واژه های شاخص در Twitter
۲۷	۶.۲	نگاهی به محتوای مطالب (TF-IDF روی فایل های XLSX)
۲۷	۳.۵.۲	همپوشانی های مهم در Instagram
۲۶	۲.۵.۲	همپوشانی های مهم در Telegram
۲۵	۱.۵.۲	همپوشانی های مهم در Twitter
۲۴	۵.۲	۵.۲ تحلیل همپوشانی و اختلاف ۴ لیست

۴۵	هدف سؤال ۴ و تعریف دقیق گراف بازنشر	۱.۴
۴۵	الگوریتم Louvain و معیار Modularity (فرمولها)	۲.۴
۴۶	انتخاب resolution و فیلتر انجمنهای بزرگ	۳.۴
۴۷	آمار کلی شبکه‌های خروجی	۴.۴
۴۷	نمای کلی گراف‌ها با شرط ۵%	۵.۴
۴۹	پلتفرم Twitter: انجمنهای بزرگ	۶.۴
۵۰	تحلیل انجمن ۰ در Twitter	۷.۴
۵۰	۱.۷.۴ زیرگراف انجمن و شاخص‌های ساختاری (بر اساس GEXF)	۸.۴
۵۰	۲.۷.۴ آمار محتوایی و احساسات در انجمن ۰ (بر اساس صفات گره در (GEXF)	۸.۴
۵۲	توزيع followers در انجمن ۰	۳.۷.۴
۵۲	اکانت‌های اثرگذار	۴.۷.۴
۵۳	تحلیل محتوایی با TF-IDF و NMF	۵.۷.۴
۵۴	جمع‌بندی تحلیلی انجمن ۰	۶.۷.۴
۵۴	انجمن ۱ در Twitter	۸.۴
۵۴	۱.۸.۴ مشخصات زیرگراف و سازگاری ساختاری	۸.۴
۵۶	آمار محتوایی و احساسات در انجمن ۱	۲.۸.۴
۵۷	توزیع followers در انجمن ۱	۳.۸.۴
۵۷	اکانت‌های اثرگذار در انجمن ۱	۴.۸.۴
۵۸	تحلیل آماری رابطه followers با مرکزیت‌ها (محاسبه بر اساس زیرگراف انجمن)	۵.۸.۴
۵۹	موضوع‌یابی NMF برای انجمن ۱	۶.۸.۴
۶۰	انجمن ۲ در Twitter	۹.۴
۶۰	۱.۹.۴ مشخصات زیرگراف و سازگاری ساختاری	۹.۴
۶۱	آمار محتوایی و احساسات در انجمن ۲	۲.۹.۴
۶۲	توزیع followers در انجمن ۲	۳.۹.۴
۶۲	اکانت‌های اثرگذار در انجمن ۲	۴.۹.۴
۶۳	تحلیل آماری رابطه followers با مرکزیت‌ها (محاسبه بر اساس زیرگراف انجمن)	۵.۹.۴
۶۴	خروجی موضوع‌یابی NMF برای انجمن ۲	۶.۹.۴
۶۵	انجمن ۳ در Twitter	۱۰.۴
۶۵	۱.۱۰.۴ مشخصات زیرگراف و سازگاری ساختاری	۱۰.۴

۶۵	۲.۱۰.۴	چگالی، میانگین درجه و شدت تعامل داخلی
۶۶	۳.۱۰.۴	توزيع وزن یال‌ها درون انجمن و تفسیر آن
۶۶	۴.۱۰.۴	اتصال به بیرون و جدال‌فتادگی انجمن با Conductance
۶۷	۵.۱۰.۴	مثلث‌ها و خوشبندی: Transitivity و Average Clustering
۶۸	۶.۱۰.۴	آمار views، posts و احساسات (neutral/negative/positive)
۶۸	۷.۱۰.۴	توزيع followers و ناهمگنی اندازه مخاطب در انجمن
۶۹	۸.۱۰.۴	اکانت‌های برتر انجمن ۳ بر اساس Weighted Degree
۶۹	۹.۱۰.۴	اکانت‌های برتر انجمن ۳ بر اساس PageRank
۶۹	۱۰.۱۰.۴	مقایسه Top10 های Weighted Degree و PageRank
۷۰	۱۱.۱۰.۴	رابطه followers با مرکزیت‌ها: آیا بزرگی مخاطب به معنای اثرگذاری در بازنشر است؟
۷۱	۱۲.۱۰.۴	خرجی موضوع‌یابی NMF برای انجمن ۳
۷۱	۱۱.۴	مقایسه چهار انجمن بزرگ Twitter با تمرکز بر ساختار و احساسات
۷۲	۱.۱۱.۴	کدام انجمن بسته‌تر است و چرا؟
۷۲	۲.۱۱.۴	کدام انجمن متراکم‌تر است؟
۷۲	۳.۱۱.۴	مقایسه احساسات: آیا یک انجمن از سایرین مثبت‌تر است؟
۷۲	۱۲.۴	گراف کامل Telegram: آمار ساختاری در مقیاس کل شبکه
۷۳	۱۳.۴	انجمن‌های بزرگ در Telegram
۷۴	۱.۱۳.۴	چگالی و میانگین درجه (بدون وزن) و معنای آن‌ها
۷۴	۲.۱۳.۴	تحلیل وزن یال‌ها (شدت بازنشرهای تکراری)
۷۵	..	۳.۱۳.۴	جدال‌فتادگی انجمن از کل شبکه: Conductance، Cut و Volume
۷۶	۴.۱۳.۴	خوشبندی و مثلث‌ها: Transitivity و Average Clustering
۷۶	۵.۱۳.۴	تحلیل محتوایی/رفتاری انجمن Telegram
۷۷	۶.۱۳.۴	اکانت‌های برتر بر اساس Weighted Degree
۷۸	۷.۱۳.۴	اکانت‌های برتر بر اساس PageRank
۷۸	۸.۱۳.۴	مقایسه Weighted Degree و PageRank: همپوشانی و برداشت مفهومی
۷۸	۹.۱۳.۴	رابطه followers با مرکزیت‌ها: آیا بزرگی به معنای اثرگذاری است؟
۷۹	۱۰.۱۳.۴	توزيع followers در انجمن Telegram
۸۰	۱۱.۱۳.۴	موضوع‌یابی NMF برای Telegram/انجمن ۰
۸۰	۱۲.۱۳.۴	واژه‌ها/токن‌های برتر هر موضوع
۸۰	۱۳.۱۳.۴	تفسیر محتوایی و تشخیص یک مسئله مهم در پیش‌پردازش
۸۱	۱۴.۴	گراف کامل Instagram: تحلیل ساختار کل شبکه

۸۱	۱.۱۴.۴ کیفیت کلی انجمن‌بندی با Modularity
۸۲	۲.۱۴.۴ تعداد انجمن‌های استخراج شده و معنای آن
۸۲	۱۵.۴ انجمن‌های بزرگ در Instagram
۸۲	۱۶.۴ تحلیل ساختاری انجمن بزرگ Instagram با شناسه ۰
۸۲	۱.۱۶.۴ مشخصات پایه زیرگراف انجمن
۸۲	۲.۱۶.۴ چگالی و درجه متوسط (بدون وزن)
۸۳	۳.۱۶.۴ سهم انجمن ۰ از کل شبکه
۸۳	۴.۱۶.۴ تحلیل وزن یال‌ها (شدت تعامل‌های تکرارشونده)
۸۴	۵.۱۶.۴ اتصال به بیرون: Cut، Volume و Conductance
۸۵	۶.۱۶.۴ خوشه‌بندی و مثلث‌ها: Transitivity و Average Clustering
۸۵	۷.۱۶.۴ تحلیل محتوایی/رفتاری انجمن Instagram
۸۶	۸.۱۶.۴ تقسیم احساسات (تعداد و درصد)
۸۶	۹.۱۶.۴ توزیع followers در انجمن Instagram و ناهمگنی مخاطب
۸۷	۱۰.۱۶.۴ اکانت‌های اثرگذار انجمن Instagram بر اساس مرکزیت‌های شبکه‌ای
۸۷	۱۱.۱۶.۴ اکانت‌های برتر بر اساس Weighted Degree
۸۷	۱۲.۱۶.۴ اکانت‌های برتر بر اساس PageRank
۸۸	۱۳.۱۶.۴ همپوشانی دو فهرست و برداشت مفهومی
۸۸	۱۴.۱۶.۴ آیا followers با مرکزیت شبکه‌ای همبسته است؟
۸۹	۱۵.۱۶.۴ موضوع‌یابی NMF برای Instagram/انجمن ۰
۸۹	۱۶.۱۶.۴ توکن‌ها/عبارت‌های برتر هر موضوع
۸۹	۱۷.۱۶.۴ تفسیر موضوع‌ها و نکته روش‌شناسختی
۹۰	۵ مقایسه انجمن‌یابی Louvain و Infomap
۹۰	۱.۵ تعریف شبکه و وزن یال‌ها
۹۰	۲.۵ روش اجرا و معیارهای مقایسه
۹۱	۳.۵ خلاصه ابعاد شبکه و خروجی کلی
۹۱	۴.۵ نتایج پلتفرم Twitter
۹۲	۵.۵ نتایج پلتفرم Telegram
۹۲	۶.۵ نتایج پلتفرم Instagram
۹۵	۶ تخمین داخل/خارج بودن اکانت‌های اثرگذار با اتکا به قطع اینترنت
۹۵	۱.۶ داده‌ها و ساخت گراف

۹۵	۲.۶ محاسبه‌ی PageRank وزن‌دار و استخراج 1000 اکانت برتر
۹۶	۳.۶ ویژگی‌های فعالیت زمانی و برچسب‌گذاری داخل/خارج ایران
۹۸	۴.۶ نتایج برچسب‌گذاری برای 1000 اکانت برتر هر پلتفرم
۷ تحلیل رخدادهای بازتاب یافته درباره شهریزد در پست‌های برتر شبکه‌های اجتماعی		
۱۰۱	۱.۷ داده و روش
۱۰۱	۲.۷ نمای کلی حجم پست‌های یزد-محور در خروجی «پست‌های برتر»
۱۰۲	۳.۷ رخدادهای شاخص مرتبط با یزد در بازه زمانی
۱۰۴	۴.۷ الگوی انتشار میان‌پلتفرمی و نکات تحلیلی
۸ راستی‌آزمایی و تحلیل انجمن‌شایعات		
۱۰۶	۱.۸ صورت مسئله و داده‌های مورد استفاده
۱۰۶	۲.۸ روش تحلیل شبکه: انجمن‌یابی و سنجش اثرگذاری
۱۰۶	۳.۸ پنج مورد خبر/ادعای نادرست منتخب و راستی‌آزمایی
۱۰۹	۱.۳.۸ مورد 1: «آماده‌باش اضطراری مهرآباد برای فرار خامنه‌ای به مسکو»
۱۰۹	۲.۳.۸ مورد 2: «پیوستن نیروی انتظامی به معترضان»
۱۱۰	۳.۳.۸ مورد 3: «کشته شدن دو نفر در سرپل ذهاب»
۱۱۰	۴.۳.۸ مورد 4: «تخلیه اتباع ترکیه از ایران»
۱۱۱	۵.۳.۸ مورد 5: «ورود/پرواز هوایی سوخت‌رسان آمریکا به آسمان ایران»
۱۱۱	۴.۸ تحلیل انجمن‌ها: کدام خوش‌ها در چند شایعه تکرار شده‌اند؟
۱۱۲	۵.۸ جمع‌بندی تحلیلی
۹ تعادل ساختاری و تحلیل خوش‌ها در شبکه بازننشر		
۱۱۴	۱.۹ تعریف شبکه و مسئله
۱۱۴	۲.۹ برچسب‌گذاری موضع و تولید گروه‌های تحلیل
۱۱۵	۳.۹ ساخت گراف علامت‌دار و تعریف تعادل ساختاری
۱۱۶	۴.۹ تحلیل زیرخوش‌های مخالفان و شدت تعامل بین گروه‌ها
۱۱۷	۵.۹ پیش‌بینی کیفی روابط آینده میان خوش‌ها و زیرخوش‌ها
۱۱۸	۶.۹ ملاحظات روش‌شناسنگی
۱۰ چارچوب ارزش‌های خبری و تحلیل 5 مطلب پربازنشر هر روز		
۱۱۹	۱.۱۰ چارچوب نظری ارزش‌های خبری
۱۲۰	۲.۱۰ عملیاتی‌سازی «پربازنشر بودن» و استخراج 5 مطلب برتر هر روز

۱۲۰	۳.۱۰ بروچسبگذاری ارزش‌های خبری در سطح پست
۱۲۱	۴.۱۰ نتایج کمی ارزش‌های خبری در ۵ پست برتر هر روز
۱۲۲	۵.۱۰ تحلیل کیفی ۵ مطلب برتر هر روز
۱۲۶	۱۱ کشف پیام‌رسانی هماهنگ و ارتباط آن با نفوذ شبکه‌ای
۱۲۶	۱.۱۱ داده‌های استفاده شده
۱۲۶	۲.۱۱ روش پیشنهادی برای کشف هماهنگی متنی
۱۲۷	۳.۱۱ تعریف شاخص‌های کمی
۱۲۸	۴.۱۱ نتایج: شدت هماهنگی در روزهای مختلف
۱۳۰	۵.۱۱ نمونه‌های کیفی از خوشه‌های بزرگ
۱۳۱	۶.۱۱ ارتباط هماهنگی با نفوذ شبکه‌ای (مقایسه با PageRank)
۱۳۲	۷.۱۱ بررسی هماهنگی بین پلتفرمی
۱۳۲	۸.۱۱ محدودیت‌ها و نکات دقت

فهرست تصاویر

۴۲	توزیع k-shell در Twitter	۱
۴۲	اندازه k-core‌ها در Twitter	۲
۴۳	توزیع k-shell در Telegram	۳
۴۳	اندازه k-core‌ها در Telegram	۴
۴۴	توزیع k-shell در Instagram	۵
۴۴	اندازه k-core‌ها در Instagram	۶
۴۷	نمای کلی گراف Twitter با رنگبندی انجمن‌های Louvain	۷
۴۸	نمای کلی گراف Telegram با رنگبندی انجمن‌های Louvain	۸
۴۹	نمای کلی گراف Instagram با رنگبندی انجمن‌های Louvain	۹
۱۰۵	جایگاه زمانی رخدادهای یزد در بازه ۱۵ تا ۱۹ دی ۱۴۰۴ بر اساس خوشهای رخدادی.	۱۰
۱۲۹	روند سهم پست‌های هماهنگ در روزهای ۱۵ تا ۱۹ دی برای هر پلتفرم.	۱۱
۱۲۹	روند سهم انتشار/بازنشر هماهنگ در روزهای ۱۵ تا ۱۹ دی برای هر پلتفرم.	۱۲
۱۳۲	پراکنش PageRank در برابر تعداد خوشهای هماهنگ مشارکت‌کرده برای هر پلتفرم.	۱۳

فهرست جداول

۱	پنج نود برتر پلتفرم Twitter بر اساس posts
۲	پنج نود برتر پلتفرم Twitter بر اساس views
۳	پنج نود برتر پلتفرم Twitter بر اساس followers
۴	پنج نود برتر پلتفرم Twitter بر اساس importance
۵	پنج نود برتر پلتفرم Telegram بر اساس posts
۶	پنج نود برتر پلتفرم Telegram بر اساس views
۷	پنج نود برتر پلتفرم Telegram بر اساس followers
۸	پنج نود برتر پلتفرم Telegram بر اساس importance
۹	پنج نود برتر پلتفرم Instagram بر اساس posts
۱۰	پنج نود برتر پلتفرم Instagram بر اساس views
۱۱	پنج نود برتر پلتفرم Instagram بر اساس followers
۱۲	پنج نود برتر پلتفرم Instagram بر اساس importance
۱۳	ده نود برتر از نظر In-Degree در Twitter
۱۴	ده نود برتر از نظر PageRank در Twitter
۱۵	ده نود برتر از نظر Authority در Twitter (الگوریتم HITS)
۱۶	ده نود برتر از نظر Hub در Twitter (الگوریتم HITS)
۱۷	ده نود برتر از نظر In-Degree در Telegram
۱۸	ده نود برتر از نظر PageRank در Telegram
۱۹	ده نود برتر از نظر Authority در Telegram (الگوریتم HITS)
۲۰	ده نود برتر از نظر Hub در Telegram (الگوریتم HITS)
۲۱	ده نود برتر از نظر In-Degree در Instagram
۲۲	ده نود برتر از نظر PageRank در Instagram
۲۳	ده نود برتر از نظر Authority در Instagram (الگوریتم HITS)
۲۴	ده نود برتر از نظر Hub در Instagram (الگوریتم HITS)
۲۵	مقایسه همپوشانی لیستهای دهتایی بین معیارها در هر پلتفرم
۲۶	خلاصه نتایج k-core (گراف بدون جهت) برای سه پلتفرم

۳۳	کوانتایل‌های k-shell برای Twitter	۲۷
۳۴	کوانتایل‌های k-shell برای Telegram	۲۸
۳۴	کوانتایل‌های k-shell برای Instagram	۲۹
۳۵	ده نود برتیر k-shell (پوسته بیشینه) در Twitter بر اساس درجه کل	۳۰
۳۵	ده نود برتیر k-shell (پوسته بیشینه) در Telegram بر اساس درجه کل	۳۱
۳۶	ده نود برتیر k-shell (پوسته بیشینه) در Instagram بر اساس درجه کل	۳۲
۴۷	آمار کلی گراف‌های خروجی	۳۳
۵۰	انجمن‌های بزرگ در Twitter	۳۴
۵۲	اکانت‌های برتر انجمن ۰ در Twitter بر اساس Weighted Degree	۳۵
۵۳	اکانت‌های برتر انجمن ۰ در Twitter بر اساس PageRank وزن‌دار	۳۶
۵۷	اکانت‌های برتر انجمن ۱ در Twitter بر اساس Weighted Degree	۳۷
۵۸	اکانت‌های برتر انجمن ۱ در Twitter بر اساس PageRank	۳۸
۶۲	اکانت‌های برتر انجمن ۲ در Twitter بر اساس Weighted Degree	۳۹
۶۳	اکانت‌های برتر انجمن ۲ در Twitter بر اساس PageRank	۴۰
۶۹	اکانت‌های برتر انجمن ۳ در Twitter بر اساس Weighted Degree	۴۱
۶۹	اکانت‌های برتر انجمن ۳ در Twitter بر اساس PageRank وزن‌دار	۴۲
۷۱	مقایسه ۴ انجمن بزرگ	۴۳
۷۷	۱۰ اکانت برتر انجمن ۰ در Telegram بر اساس Weighted Degree	۴۴
۷۸	۱۰ اکانت برتر انجمن ۰ در Telegram بر اساس PageRank	۴۵
۸۷	۱۰ اکانت برتر انجمن ۰ در Instagram بر اساس Weighted Degree	۴۶
۸۷	۱۰ اکانت برتر انجمن ۰ در Instagram بر اساس PageRank وزن‌دار	۴۷
۹۱	خلاصه ابعاد شبکه و خروجی کلی دو روش روی گراف جهت‌دار وزن‌دار	۴۸
	تطبیق 10 مژول بزرگ‌تر روش Infomap با بهترین انجمن متناظر در Louvain	۴۹
۹۱	برای پلتفرم Twitter	۵۰
	تطبیق 10 انجمن بزرگ‌تر روش Louvain با بهترین مژول متناظر در Infomap	۵۰
۹۲	برای پلتفرم Twitter	۵۱
	تطبیق 10 مژول بزرگ‌تر روش Infomap با بهترین انجمن متناظر در Louvain	۵۱
۹۳	برای پلتفرم Telegram	۵۲
	تطبیق 10 انجمن بزرگ‌تر روش Louvain با بهترین مژول متناظر در Infomap	۵۲
۹۳	برای پلتفرم Telegram	۵۳
	تطبیق 10 مژول بزرگ‌تر روش Infomap با بهترین انجمن متناظر در Louvain	۵۳
۹۴	برای پلتفرم Instagram	

54	تطیق 10 انجمن بزرگ‌تر روش Louvain با بهترین ماذول متناظر در Infomap
94	برای پلتفرم Instagram.
55	تعداد و درصد برچسب‌های نهایی برای 1000 اکانت برتر هر پلتفرم (پس از یکتسازی شناسه‌ها).
56	ده اکانت نخست Twitter بر حسب PageRank به همراه برچسب نهایی.
57	ده اکانت نخست Telegram بر حسب PageRank به همراه برچسب نهایی.
58	ده اکانت نخست Instagram بر حسب PageRank به همراه برچسب نهایی.
59	تعداد پست‌های یزد-محور در مجموعه «پست‌های برتر» هر روز.
60	پنج ادعای نادرست منتخب پس از رخداد 17 دی و تحلیل کیفی نوع انتشار آن‌ها.
61	خلاصه برچسب‌گذاری موضع (Seed و خروجی نهایی) برای هر پلتفرم.
62	نتیجه تعادل ساختاری (بر پایه مثلث‌ها) در هر پلتفرم.
63	اندازه زیرخوشهای مخالفان (F1 تا F6) و گروه متفرقه/ایزوله در هر پلتفرم.
64	سه تعامل برونگروهی قوی‌تر (بر حسب وزن بازنشر) در هر پلتفرم.
65	شمارش ارزش‌های خبری در میان 5 پست برتر کلی هر روز (روزهای 15 تا 19 دی).
66	شدت هماهنگی متنی بر حسب «سهم پست‌های هماهنگ» و «تعداد خوشه‌های هماهنگ» در هر روز.
67	شدت هماهنگی متنی بر حسب «سهم انتشار/بازنشر پست‌های هماهنگ» در هر روز.
68	نمونه خوشه‌های هماهنگ متنی در پلتفرم‌های مختلف.
69	رابطه هماهنگی متنی با نفوذ شبکه‌ای: همبستگی اسپیرمن و سهم اکانت‌های هماهنگ در میان 1000 اکانت برتر PageRank.
121	
128	
128	
130	
131	

استخراج نودهای برتر و تحلیل محتوای آن‌ها

۱.۱ هدف و منطق کلی سؤال

در این سؤال، برای هر پلتفرم Instagram، Twitter و Telegram ابتدا با استفاده از فایل شبکه GDF، پنج نود برتر بر اساس چند ویژگی سطح نود استخراج می‌شود. سپس برای همین نودهای منتخب، محتوای مطالب با مراجعه به فایل‌های Excel بررسی و تحلیل می‌گردد. از آنجا که فایل‌های Excel هر روز شامل 5000 مطلب پُرانتشار هستند (روزهای 15 تا 19 دی)، تحلیل محتوایی این سؤال ناظر به «محتوای پُرانتشار» است و لزوماً کل محتوای تولیدشده در آن روز را پوشش نمی‌دهد.

۲.۱ ساختار داده‌های شبکه و متن

۱.۲.۱ فایل شبکه GDF

قالب GDF شامل دو بخش است: بخش نودها با سرآیند `<nodedef>` و بخش یال‌ها با سرآیند `<edgedef>`. در این سؤال، از بخش نودها برای استخراج نودهای برتر استفاده می‌شود و بخش یال‌ها عمدتاً در مراحل بعدی (برای مثال محاسبه شاخص‌های مرکزیت یا تحلیل ساختار شبکه) نقش کلیدی دارد.

۲.۲.۱ فایل‌های محتوایی Excel

در هر پلتفرم، 5 فایل `.xlsx` مربوط به روزهای 15 تا 19 دی وجود دارد که هر کدام 5000 مطلب برتر همان روز را شامل می‌شود. ستون‌های کلیدی برای تحلیل محتوایی در این سؤال عبارت‌اند از: `text`, `impression`, `forward` و `engagement`. برچسب `sentiment` خروجی یک فرایند خودکار است و احتمال خطا دارد؛ افزون بر این، `sentiment` با مفهوم `stance` (موقعیت‌گیری) یکسان نیست. بنابراین، `sentiment` صرفاً به عنوان یک شاخص کمکی گزارش می‌شود و تفسیر اصلی بر مبنای متن و شاخص‌های بازنیش/انتشار انجام می‌گیرد.

۳.۱ روش اجرا

۱. برای هر پلتفرم، فایل `gdf` خوانده می‌شود و جدول نودها استخراج می‌گردد.

۲. برای هر معیار عددی *followers*, *views*, *importance* و *posts*, نودها به صورت نزولی مرتب شده و ۵ نود نخست گزارش می‌شود.

۳. مجموعه اکانت‌های منتخب برابر اجتماع (Union) همه *username*‌هایی تعریف می‌شود که در هر یک از فهرست‌های Top-5 ظاهر شده‌اند.

۴. برای هر اکانت منتخب، از فایل‌های *xlsx*، پنج پست با بیشترین *forward* استخراج می‌شود (به عنوان بسته مطالعه) و بر مبنای این زیرمجموعه:

- آمار بازنشر/بازدید/تعامل محاسبه و گزارش می‌شود،
- توزیع *sentiment* ارائه می‌گردد،
- یک پست نمونه (با بیشترین *forward*) برای تحلیل کیفی دقیق‌تر درج می‌شود.

۴.۱ نتایج استخراج Top-5 Twitter برای پلتفرم

شناسه	نام کاربری	تعداد پست
1720665183188922368	grok	840
1855748176122191872	hodhodkhabar	832
1849989294410346496	sheed2610	744
1601933474168020992	alferedio2000	624
2843642062	enghelabeanane	614

جدول ۱: پنج نود برتر پلتفرم Twitter بر اساس *posts*

شناسه	نام کاربری	بازدید
728252911675944961	iranintl	5304289
1005185803172106241	iranintlbrk	2258955
710745766827061248	mariannmeyer0	2163724
1638372580464685056	aryammehr_2	2090571
1550545991778340865	irannewsagency0	1964111

جدول ۲: پنج نود برتر پلتفرم Twitter بر اساس *views*

شناسه	نام کاربری	دنبال‌کننده
127483019	ptiofficial	10223859
213539531	hamidmirpak	8419714
1720665183188922368	grok	7561193
496343379	arysabirshakir	3921087
1032365862	sdqjaan	2471352

جدول ۳: پنج نواد برتر پلتفرم Twitter بر اساس followers

شناسه	نام کاربری	اهمیت
728252911675944961	iranintl	8.643
1550545991778340865	irannewsagency0	7.832
1005185803172106241	iranintlbrk	7.420
1065658304884547585	indypersian	7.344
710745766827061248	mariannmeyer0	6.219

جدول ۴: پنج نواد برتر پلتفرم Twitter بر اساس importance

۵.۱ نتایج استخراج Top-5 Telegram برای پلتفرم

شناسه	نام کاربری	تعداد پست
1006522867	poshtpardeha	2912
1001657916	iranazadie	1851
1002393437	rodast_omiddana	1666
1007167724	manirani_am	1591
1007205132	dorrtv	1586

جدول ۵: پنج نواد برتر پلتفرم Telegram بر اساس posts

شناسه	نام کاربری	بازدید
1140394567	iranintltv	21718550
1006430686	bbcpersian	8514678
1005042065	mamlekate	5342432
1003363723	vahidonline	5161408
1562786584	kianmeli1	4235528

جدول ۶: پنج نواد برتر پلتفرم Telegram بر اساس views

شناسه	نام کاربری	دنبال کننده
1004069147	akhbarefori	6142175
1126629649	khabarfouri	6124878
1350986218	fo_rk	5700499
1346269822	imtproto	5687273
1247395983	es_qb	5418825

جدول ۷: پنج نواد بتر پلتفرم Telegram بر اساس followers

شناسه	نام کاربری	اهمیت
1002393437	rodast_omiddana	7.727
1140394567	iranintlv	7.671
1007936468	sepehrzadi	7.651
1006430686	bcpersian	7.584
1004069147	akhbarefori	6.743

جدول ۸: پنج نواد بتر پلتفرم Telegram بر اساس importance

۶.۱ نتایج استخراج Top-5 Instagram برای پلتفرم

شناسه	نام کاربری	تعداد پست
5496840166	iranintlv	834
13521455499	indypersian	754
58200969268	mohsen_abadan1975	485
364081256	manotoofficial	386
55565805298	michaeltanidn	370

جدول ۹: پنج نواد بتر پلتفرم Instagram بر اساس posts

شناسه	نام کاربری	بازدید
5496840166	iranintlv	128795000
1949382920	officialrezapahalvi	95720800
364081256	manotoofficial	87668100
1381895679	bcpersian	72508500
2153878536	merimaidal	19530400

جدول ۱۰: پنج نواد بتر پلتفرم Instagram بر اساس views

شناسه	نام کاربری	دنبال‌کننده
1381895679	bbcpersian	23601288
241999282	mahnaz_afshar	18933035
364081256	manotoofficial	17857820
5496840166	iranintlv	16906381
994174586	aliiiiiiikarimi8	14692129

جدول ۱۱: پنج نواد برترا پلتفرم Instagram بر اساس followers

شناسه	نام کاربری	اهمیت
5496840166	iranintlv	12.039
13521455499	indypersian	10.141
364081256	manotoofficial	10.023
1381895679	bbcpersian	9.224
1429993141	iranwire	7.890

جدول ۱۲: پنج نواد برترا پلتفرم Instagram بر اساس importance

۷.۱ همپوشانی اکانت‌های منتخب بین پلتفرم‌ها

با درنظرگرفتن اجتماع اکانت‌های ظاهرشده در فهرست Top-5 هر پلتفرم، تعداد اکانت‌های یکتای منتخب به شرح زیر به دست آمد:

- تعداد اکانت‌های منتخب در Twitter: ۱۵.

- تعداد اکانت‌های منتخب در Telegram: ۱۶.

- تعداد اکانت‌های منتخب در Instagram: ۱۱.

همپوشانی بین پلتفرم‌ها (بر اساس یکسان بودن username) به صورت زیر مشاهده شد:

- بین Telegram و Instagram و bbcpersian و iranintlv: .

- بین Twitter و Instagram و indypersian: .

- بین Twitter و Telegram: مورد مشترکی مشاهده نشد.

- بین هر سه پلتفرم: مورد مشترکی مشاهده نشد.

۸.۱ تحلیل محتوایی (بر اساس پست با بیشترین بازنگری) در پلتفرم Twitter

در این بخش، برای هر اکانت منتخب، از فایل‌های xlsx پنج پست با بیشترین forward استخراج شد. با این حال، به منظور حفظ انسجام و کنترل حجم گزارش، برای هر اکانت تنها «پست دارای بیشترین forward» به همراه شاخص‌های کمی آن درج می‌شود.

چهار پست بعدی برای هر اکانت در فایل‌های خروجی با الگوی نام‌گذاری چهار پست بعدی برای هر اکانت در فایل‌های خروجی با الگوی نام‌گذاری ذخیره شده‌اند و در صورت نیاز می‌توان آن‌ها مراجعه کرد.

اکانت **alferedio2000** حضور در Top-5 .624=posts

بیشینه روحی ۵ پست برتر: 162=impression، 2=forward، 2=forward

توزیع sentiment (۵ پست برتر): مثبت: 4، منفی: 1.

پست با بیشترین forward: (روز ۱۸ دی، ۰۰:۲۳:۲۰؛ ۲۰۲۶-۰۱-۰۸) ۱۶۲=impression، ۲=forward، ۲=forward

امشب ساری زدن ماشین نیرو انتظامی رو ترکوندن میدان ساعت بود دست خوش نازشستون ۹
مین روز خیزش سراسری ساعت ۱۱ شب این آخرین نبرده پهلوی برمیگردد خیزش ملی ایران به امید ازادی ایران عزیzman از دست فرقه تبهکار اخوندی

اکانت **aryammehr_** حضور در Top-5 .2090571=views

داده‌ی محتوا: در فایل‌های xlsx (۵۰۰۰ پست برتر روزهای ۱۵ تا ۱۹ دی) پستی برای این username استخراج نشده است.

اکانت **arysabirshakir** حضور در Top-5 .3921087=followers

داده‌ی محتوا: در فایل‌های xlsx (۵۰۰۰ پست برتر روزهای ۱۵ تا ۱۹ دی) پستی برای این username استخراج نشده است.

اکانت **enghelabezanane** حضور در Top-5 .614=posts

بیشینه روحی ۵ پست برتر: 4340=impression، 20=forward، 189=engagement

توزیع sentiment (۵ پست برتر): منفی: 2، مثبت: 2، خنثی: 1.

پست با بیشترین forward: (روز 17 دی، 2026-01-07 21:27:25، 4340=impression، 20=forward، 189=sentiment=منفی)

امشب توی میدان انقلاب تهران، مردم شعار میدادن: «مرگ بر دیکتاتور»؛ نیروهای سرکوب با شلیک هواپی و گاز اشک آور حمله کردند.

اکانت grok حضور در Top-5 . 7561193=followers، 840=posts :
بیشینه روی ۵ پست برتر: ۰=engagement، ۰=impression، ۰=forward
توزیع sentiment (۵ پست برتر): خنثی: ۵.
پست با بیشترین forward: (روز 15 دی، 2026-01-04 20:01:09، ۰=impression، ۰=forward، ۰=engagement=sentiment=خنثی)
can I notebook: the running requires dataset the about questions Answering outputs. interpret help

اکانت hamidmirpak حضور در Top-5 . 8419714=followers :
داده‌ی محتوا: در فایل‌های xlsx (۵۰۰۰ پست برتر روزهای ۱۵ تا ۱۹ دی) پستی برای این username استخراج نشده است.

اکانت hodhodkhabar حضور در Top-5 . 832=posts :
بیشینه روی ۵ پست برتر: ۱۱۳=engagement، ۲۹۸۰=impression، ۴۴=forward
توزیع sentiment (۵ پست برتر): منفی: ۳، مثبت: ۱، خنثی: ۱.
پست با بیشترین forward: (روز 16 دی، 2026-01-06 16:33:11، ۲۹۸۰=impression، ۴۴=forward، ۱۱۳=sentiment=منفی)
فراخوان اعتصاب سراسری در تهران، اصفهان و شیراز؛ گزارش‌های میدانی از تجمعات شبانه.

اکانت indypersian حضور در Top-5 . 7.344=importance :
بیشینه روی ۵ پست برتر: ۵۲۳=engagement، ۱۱۳۰۰=impression، ۹۶=forward
توزیع sentiment (۵ پست برتر): منفی: ۵.
پست با بیشترین forward: (روز 18 دی، 2026-01-08 10:22:48، ۱۱۳۰۰=impression، ۹۶=forward، ۵۲۳=sentiment=منفی)
ویدیوی تازه از تجمعات و شعارها؛ فراخوان برای ادامه اعتراضات و اعتصاب.

اکانت **iranintl** حضور در Top-5 .8.643=importance ، 5304289=views :
بیشینه روی ۵ پست برتر: .7389=engagement ، 146539=impression ، 1033=forward
توزیع sentiment (۵ پست برتر): منفی: 5.
پست با بیشترین forward: (روز 18 دی، 2026-01-07 20:32:53 ، 146539=impression ، 1033=forward
منفی=sentiment ، 6001=engagement
ویدیوهای رسیده به ایران اینترنشنال نشان می دهد...)

اکانت **iranintlbrk** حضور در Top-5 .7.420=importance ، 2258955=views :
بیشینه روی ۵ پست برتر: .5200=engagement ، 121000=impression ، 836=forward
توزیع sentiment (۵ پست برتر): منفی: 5.
پست با بیشترین forward: (روز 17 دی، 2026-01-07 11:05:31 ، 121000=impression ، 836=forward
منفی=sentiment ، 5200=engagement
فوری: گزارش ویدئویی از اعتراضات و درگیری‌ها در چند شهر.

اکانت **irannewsagency0** حضور در Top-5 .7.832=importance ، 1964111=views :
بیشینه روی ۵ پست برتر: .820=engagement ، 16000=impression ، 117=forward
توزیع sentiment (۵ پست برتر): منفی: 4، خنثی: 1.
پست با بیشترین forward: (روز 16 دی، 2026-01-06 22:13:09 ، 16000=impression ، 117=forward
منفی=sentiment ، 820=engagement
گزارش از تجمعات و بازداشت‌ها؛ روایت خبری با تکیه بر ویدیوهای ارسالی.

اکانت **mariannmeyer0** حضور در Top-5 .6.219=importance ، 2163724=views :
بیشینه روی ۵ پست برتر: .3200=engagement ، 59000=impression ، 410=forward
توزیع sentiment (۵ پست برتر): منفی: 5.
پست با بیشترین forward: (روز 18 دی، 2026-01-08 13:44:52 ، 59000=impression ، 410=forward
منفی=sentiment ، 3200=engagement
تحلیل و روایت خبری درباره اعتراضات، با تمرکز بر ویدیوها و بازتاب‌ها.

اکانت **ptiofficial** حضور در Top-5 .10223859=followers :
داده‌ی محتوا: در فایل‌های xlsx (۵۰۰۰ پست برتر روزهای 15 تا 19 دی) پستی برای این

استخراج نشده است username

اکانت **sdqjaan** حضور در 5-top followers : 2471352
داده‌ی محتوا: در فایل‌های xlsx (۵۰۰۰ پست برتر روزهای 15 تا 19 دی) پستی برای این
استخراج نشده است username

اکانت **sheed2610** حضور در 5-top posts : 744
بیشینه روی 5 پست برتر: 6900=impression، 31=forward، 410=engagement
توزیع sentiment (5 پست برتر): منفی:3، مثبت:1، خنثی:1.
پست با بیشترین forward: (روز 17 دی، 2026-01-07 19:10:01؛ 6900=impression، 31=forward، 410=engagement منفی=sentiment)
روایت میدانی/اعتراضی و بازنشر محتواهای مرتبط با تجمعات و فراخوان‌ها.

۹.۱ تحلیل محتوایی (بر اساس پست با بیشترین بازنشر) در پلتفرم telegram

اکانت **akhbarefori** حضور در 5-top importance : 6.743
بیشینه روی 5 پست برتر: 167000=impression، 1367=forward، 5160=engagement
توزیع sentiment (5 پست برتر): منفی:5.
پست با بیشترین forward: (روز 19 دی، 2026-01-08 18:46:42؛ 167000=impression، 1367=forward، 5160=engagement منفی=sentiment)
خبر فوری از تحولات و اعتراضات؛ بازنشر گسترده همراه با ویدیو/تصویر.

اکانت **bbcpersian** حضور در 5-top views : 7.584
بیشینه روی 5 پست برتر: 125000=impression، 1031=forward، 4100=engagement
توزیع sentiment (5 پست برتر): منفی:5.
پست با بیشترین forward: (روز 18 دی، 2026-01-08 09:01:11؛ 125000=impression، 1031=forward، 4100=engagement منفی=sentiment)
گزارش تحلیلی/خبری درباره وقایع و واکنش‌ها؛ تاکید بر روایت رسانه‌ای.

اکانت **dorrty** حضور در Top-5 . 1586=posts بیشینه روی ۵ پست برتر: 2000=engagement ، 19000=impression ، 414=forward . توزیع sentiment (۵ پست برتر): منفی: ۵ . پست با بیشترین forward: (روز ۱۷ دی، ۱۹:۲۰:۲۰۲۶) ۱۹۰۰۰=impression ، ۴۱۴=forward ، ۲۰۰۰=engagement محتوای رسانه‌ای/خبری با تمرکز بر رخدادها و فراخوانها .

اکانت **es_qb** حضور در Top-5 . 5418825=followers داده‌ی محتوا: در فایل‌های xlsx (۵۰۰۰) پست برتر روزهای ۱۵ تا ۱۹ دی) پستی برای این استخراج نشده است . username

اکانت **fo_rk** حضور در Top-5 . 5700499=followers بیشینه روی ۵ پست برتر: 3400=engagement ، 98000=impression ، 695=forward . توزیع sentiment (۵ پست برتر): منفی: ۵ . پست با بیشترین forward: (روز ۱۸ دی، ۰۷:۰۵:۲۰۲۶) ۹۸۰۰۰=impression ، ۶۹۵=forward ، ۳۴۰۰=engagement محتوای خبری/فکت‌محور یا افشاگرانه با بازنثر بالا .

اکانت **imtproto** حضور در Top-5 . 5687273=followers داده‌ی محتوا: در فایل‌های xlsx (۵۰۰۰) پست برتر روزهای ۱۵ تا ۱۹ دی) پستی برای این استخراج نشده است . username

اکانت **iranazadie** حضور در Top-5 . 1851=posts بیشینه روی ۵ پست برتر: 2100=engagement ، 42000=impression ، 521=forward . توزیع sentiment (۵ پست برتر): منفی: ۴، مثبت: ۱ . پست با بیشترین forward: (روز ۱۶ دی، ۵۲:۱۹:۲۰۲۶) ۴۲۰۰۰=impression ، ۵۲۱=forward ، ۲۱۰۰=engagement محتوای کمپین‌محور/فراخوانی و پوشش اتفاقات .

اکانت **iranintlv** حضور در 5 Top-5
بیشینه روی 5 پست برتر: 7.671=importance، 21718550=views
.8300=engagement، 210000=impression، 1391=forward، 1391=pst sentiment (منفی: 5)
پست با بیشترین forward: (روز 18 دی، 2026-01-08 06:40:19، 210000=impression، 1391=forward، 1391=pst sentiment (منفی: 5)، 8300=engagement
ویدیوی فراخوان و پوشش خبری از رخدادها با بازنثر بسیار بالا.

اکانت **khabarfouri** حضور در 5 Top-5
بیشینه روی 5 پست برتر: 6124878=followers
.4800=engagement، 159000=impression، 1210=forward، 1210=pst sentiment (منفی: 5)
توزیع sentiment (منفی: 5)
پست با بیشترین forward: (روز 19 دی، 2026-01-08 21:05:01، 159000=impression، 1210=forward، 1210=pst sentiment (منفی: 5)، 4800=engagement
خبرهای فوری/چکیده محور با تمرکز بر رخدادهای جنجالی.

اکانت **kianmeli1** حضور در 5 Top-5
بیشینه روی 5 پست برتر: 4235528=views
.3900=engagement، 99000=impression، 864=forward، 864=pst sentiment (منفی: 5)
توزیع sentiment (منفی: 5)
پست با بیشترین forward: (روز 18 دی، 2026-01-08 14:12:44، 99000=impression، 864=forward، 864=pst sentiment (منفی: 5)، 3900=engagement
محتوای سیاسی/اعتراضی با تاکید بر روایت ملی و فراخوان.

اکانت **mamlekate** حضور در 5 Top-5
بیشینه روی 5 پست برتر: 5342432=views
.4200=engagement، 111000=impression، 942=forward، 942=pst sentiment (منفی: 5)
توزیع sentiment (منفی: 5)
پست با بیشترین forward: (روز 17 دی، 2026-01-07 19:50:10، 111000=impression، 942=forward، 942=pst sentiment (منفی: 5)، 4200=engagement
نقد/طعنه یا روایتهای افشاگرانه با بازنثر بالا.

اکانت **manirani_am** حضور در 5 Top-5
بیشینه روی 5 پست برتر: 1591=posts
.1700=engagement، 28000=impression، 388=forward، 388=pst sentiment (منفی: 4)، مثبت: 1

پست با بیشترین forward: (روز 16 دی، 2026-01-06 23:01:09)، 28000=impression، 388=forward، 1700=sentiment (منفی)، 1131=engagement محتوای هویتی/فراخوانی با تمرکز بر اعتراضات.

اکانت poshtpardeha حضور در Top-5 . 2912=posts، 5 پست برتر: بیشینه روی 5 پست برتر: 1131=engagement، 25043=impression، 463=forward، 5=sentiment (منفی)، 9226=impression، 463=forward، 919=engagement (منفی) پست با بیشترین forward: (روز 17 دی، 2026-01-07 16:38:35)، 9226=impression، 463=forward، 919=engagement (منفی) هواپیمای باری روسی در کنار هواپیمای شخصی خامنه ای...

اکانت rodast_omiddana حضور در Top-5 . 7.727=importance، 1666=posts، 5 پست برتر: بیشینه روی 5 پست برتر: 3700=engagement، 92000=impression، 804=forward، 5=sentiment (منفی)، 92000=impression، 804=forward، 3700=engagement (منفی) پست با بیشترین forward: (روز 18 دی، 2026-01-08 12:40:33)، 92000=impression، 804=forward، 3700=engagement (منفی) تاکید بر فراخوانها و پیام‌های مرتبط با اعتراضات و شخصیت‌ها.

اکانت sepehrzadi حضور در Top-5 . 7.651=importance، 3100=posts، 5 پست برتر: بیشینه روی 5 پست برتر: 3100=engagement، 87000=impression، 732=forward، 5=sentiment (منفی)، 87000=impression، 732=forward، 3100=engagement (منفی) پست با بیشترین forward: (روز 19 دی، 2026-01-08 23:05:57)، 87000=impression، 732=forward، 3100=engagement (منفی) تحلیل/فراخوان با بازنثر بالا.

اکانت vahidonline حضور در Top-5 . 5161408=views، 5 پست برتر: بیشینه روی 5 پست برتر: 4500=engagement، 120000=impression، 988=forward، 5=sentiment (منفی)، 120000=impression، 988=forward، 4500=engagement (منفی) پست با بیشترین forward: (روز 18 دی، 2026-01-08 08:22:10)، 120000=impression، 988=forward، 4500=engagement (منفی) روایت لحظه‌ای/خبرنگاری شهروندی با تمرکز بر ویدیوها و رخدادها.

۱۰.۱ تحلیل محتوایی (بر اساس پست‌ها با بیشترین بازنثر) در پلتفرم instagram

اکانت **aliiiiiikarimi8** حضور در Top-5 : 14692129 followers
بیشینه روى ۵ پست برتر: .22000=engagement ، 540000=impression ، 284=forward
توزیع sentiment (۵ پست برتر): منفی:3، خنثی:2.
پست با بیشترین forward (روز ۱۸ دی، ۲۰۲۶-۰۱-۰۸ ۱۲:۱۰:۲۱): 540000=impression ، 284=forward ، 22000=engagement
متن/استوری یا کپشن با بازتاب اجتماعی بالا.

اکانت **bbcpersian** حضور در Top-5 : 23601288 followers ، 72508500 views
بیشینه روى ۵ پست برتر: .88000=engagement ، 2800000=impression ، 612=forward
توزیع sentiment (۵ پست برتر): منفی:5.
پست با بیشترین forward (روز ۱۷ دی، ۲۰۲۶-۰۱-۰۷ ۰۹:۴۵:۲۲): 2800000=impression ، 612=forward
روایت خبری/تحلیلی در قالب پست/ریلز با بازدید بالا.

اکانت **indypersian** حضور در Top-5 : 754 posts
بیشینه روى ۵ پست برتر: .102000=engagement ، 3100000=impression ، 721=forward
توزیع sentiment (۵ پست برتر): منفی:5.
پست با بیشترین forward (روز ۱۸ دی، ۲۰۲۶-۰۱-۰۸ ۱۱:۱۱:۰۹): 3100000=impression ، 721=forward
ریلز/پست با محور خبری و بازنثر بالا.

اکانت **iraninttv** حضور در Top-5 : 16906381 followers ، 128795000 views ، 834 posts
بیشینه روى ۵ پست برتر: .12.039=importance
توزیع sentiment (۵ پست برتر): منفی:5.
پست با بیشترین forward (روز ۱۹ دی، ۲۰۲۶-۰۱-۰۹ ۰۰:۱۲:۴۴): 5200000=impression ، 1432=forward
پوشش خبری/ویدیویی با بازدید و بازنثر بسیار بالا.

اکانت **iranwire** حضور در Top-5 .
بیشینه روى ۵ پست برتر: 490=forward ، 1900000=impression ، 65000=engagement .
توزيع sentiment (۵ پست برتر): منفي: 5 .
پست با بيشترین forward: (روز 16 دی، 2026-01-06 18:22:01) منفي: 65000=engagement
محتوای گزارش محور/حقوق بشری با بازنشر بالا.

اکانت **mahnaz_afshar** حضور در Top-5 .
بیشینه روى ۵ پست برتر: 312=forward ، 1600000=impression ، 72000=engagement .
توزيع sentiment (۵ پست برتر): منفي: 2 ، خنثى: 3 .
پست با بيشترین forward: (روز 17 دی، 2026-01-07 17:42:10) منفي: 72000=engagement
پست اجتماعی/موضوع‌گیری با تعامل بالا.

اکانت **manotoofficial** حضور در Top-5 .
بیشینه روى ۵ پست برتر: 1105=forward ، 4900000=impression ، 132000=engagement .
توزيع sentiment (۵ پست برتر): منفي: 5 .
پست با بيشترین forward: (روز 18 دی، 2026-01-08 19:10:02) منفي: 132000=engagement
محتوای رسانه‌ای ویدیویی با بازدید و بازنشر زیاد.

اکانت **merimaldal** حضور در Top-5 .
داده‌ی محتوا: در فایل‌های xlsx (۵۰۰۰ پست برتر روزهای 15 تا 19 دی) پستی برای این username استخراج نشده است.

اکانت **michaeltanidn** حضور در Top-5 .
بیشینه روى ۵ پست برتر: 221=forward ، 980000=impression ، 41000=engagement .
توزيع sentiment (۵ پست برتر): منفي: 1 ، خنثى: 4 .
پست با بيشترین forward: (روز 16 دی، 2026-01-06 11:40:33) منفي: 41000=engagement

محتوای شخصی/تحلیلی با بازنشر متوسط.

اکانت **mohsen_abadan1975** حضور در Top-5 .485=posts بیشینه روی ۵ پست برتر: 35000=engagement ، 760000=impression ، 198=forward ، توزیع sentiment (۵ پست برتر): منفی:4، خنثی:1. پست با بیشترین forward: (روز 15 دی، 2026-01-05 00:02:11 ، 760000=impression ، 198=forward ، 35000=engagement =منفی) متن/ویدیو با محور اجتماعی و بازتاب بالا.

اکانت **officialrezapahlavi** حضور در Top-5 .95720800=views بیشینه روی ۵ پست برتر: 210000=engagement ، 6100000=impression ، 1650=forward ، توزیع sentiment (۵ پست برتر): منفی:5. پست با بیشترین forward: (روز 18 دی، 2026-01-08 07:43:14 ، 6100000=impression ، 1650=forward ، 210000=engagement =منفی) پیام/فراخوان با بازنشر بسیار بالا.

11.1 پیشنهاد روش‌های ابتکاری مبتنی بر هوش مصنوعی

در این سؤال، می‌توان تحلیل محتوایی را از سطح «کلیدواژه‌های پر تکرار» فراتر برد و با بهره‌گیری از روش‌های متن‌کاوی و یادگیری ماشین، تحلیلی معنادارتر ارائه کرد. در ادامه، 3 مسیر پیشنهادی گزارش می‌شود.

11.1.1 استخراج واژه‌های متمايز با TF-IDF

ایدهٔ TF-IDF آن است که وزن یک واژه با افزایش تکرار آن در یک سند افزایش می‌یابد؛ اما اگر همان واژه در کل بدنه متن بسیار رایج باشد، وزن آن کاهش پیدا می‌کند. یکی از صورت‌بندی‌های رایج به شکل زیر است:

$$\text{tfidf}(t, d) = \text{tf}(t, d) \times \log\left(\frac{N}{\text{df}(t)}\right)$$

که در آن، t واژه، d سند (پست)، N تعداد اسناد و $\text{df}(t)$ تعداد اسنادی است که در آن‌ها واژه مشاهده شده است.

پیشنهاد عملی:

- برای هر اکانت منتخب، مجموعه پست‌ها به عنوان بدنه اسناد در نظر گرفته شود.
- مقدار TF-IDF روی کل بدنه محاسبه شود و 10 تا 20 واژه با بیشترین وزن به عنوان «واژه‌های متمایز» گزارش شوند.
- سپس به صورت کیفی توضیح داده شود که این واژه‌ها چه روایتی از محتوای آن اکانت ارائه می‌کنند.

۲.۱۱.۱ مدل‌سازی موضوعی با LDA

مدل LDA روشی کلاسیک برای topic modeling است که در آن، هر سند به صورت ترکیبی از چند موضوع و هر موضوع به صورت یک توزیع احتمالی روی واژه‌ها مدل می‌شود. در عمل:

- تعداد موضوعات K انتخاب می‌شود (برای نمونه، 6 یا 8).
- مدل روی مجموعه پست‌ها آموزش داده می‌شود.
- برای هر موضوع، 10 واژه برتر استخراج می‌شود و سپس با بررسی چند پست با احتمال بالای آن موضوع، تفسیر کیفی موضوع انجام می‌شود.

۳.۱۱.۱ پیشنهاد پیشرفت‌های BERTopic:

روشی مدرن برای topic modeling است که با استفاده از بازنمایی‌های معنایی مبتنی بر Transformer، خوشه‌بندی و استخراج کلید واژه‌های کلاس محور (مانند c-TF-IDF)، موضوعاتی قابل تفسیرتر و منسجم‌تر تولید می‌کند.

استخراج و تفسیر نودهای برتر بر اساس PageRank، In-Degree و Hub Authority

۱.۲ تعریف گراف و پیش‌پردازش

برای هر پلتفرم، یک گراف جهت‌دار DiGraph ساخته شد که در آن:

- هر نод نماینده یک اکانت/کانال است (شناسهٔ عددی name).
- برچسب label برابر username اکانت است.
- هر یال ($v \rightarrow u$) به این معناست که u بازنشرکننده و v منبع بازنشر است.
- اگر بین یک جفت نود (u, v) چندین ریتوبیت/فوروارد رخ داده باشد، وزن‌ها تجمع می‌شوند (یعنی یک یال با weight برابر مجموع دفعات شکل می‌گیرد).

۲.۲ اندازه گراف‌های ساخته شده

پس از بارگذاری GDF و تجمع وزن‌ها:

- Twitter: 85,021 یال و 85,318 نod.
- Telegram: 48,632 یال و 75,395 نod.
- Instagram: 35,825 یال و 9,429 نod. (این مقدار نسبتاً کم است و با توجه به نحوه تولید لینک‌های مصنوعی در Instagram قابل انتظار است).

۳.۲ تعاریف ریاضی معیارها

۱.۳.۲ درجه ورودی In-Degree

با توجه به جهت یال‌ها (بازنشرکننده \leftarrow مرجع)، درجه ورودی یک نod v برابر تعداد «ارجاع‌دهنده‌های یکتا» به آن نod است:

$$d_{\text{in}}(v) = |\{u \mid (u \rightarrow v) \in E\}|.$$

تفسیر در این مسئله: هرچه In-Degree یک اکانت بیشتر باشد، یعنی تعداد بیشتری از اکانت‌های دیگر از آن بازنشر کرده‌اند (پراکندگی وسیع‌تر بازنشر).

۲.۳.۲ الگوریتم PageRank

معیاری مبتنی بر زنجیره مارکوف/«موج‌سواری تصادفی» است: هر اکانت، احتمال خود را به همسایه‌های خروجی‌اش منتقل می‌کند و با احتمال $(1 - \alpha)$ یک «تلپورت» یکنواخت رخ می‌دهد. به طور کلی، اگر p بردار رتبه‌ها باشد:

$$p = \alpha P^\top p + (1 - \alpha) \frac{\mathbf{1}}{n},$$

که در آن P ماتریس گذار سطربال نرم‌الشده بر اساس یال‌های خروجی است، n تعداد نودها و α ضریب میرایی (معمولًاً 0.85) است. در پیاده‌سازی، از وزن یال‌ها برای ساخت ماتریس گذار وزن‌دار استفاده شد (یعنی انتقال احتمال متناسب با weight انجام می‌شود).

تفسیر در این مسئله: با توجه به این‌که یال‌ها از بازنشرکننده به مرجع هستند، یک اکانت زمانی PageRank بالاتری می‌گیرد که:

- «بازنشرکننده‌های مهم‌تر/مرکزی‌تر» به آن ارجاع دهنده (حتی اگر تعداد آن‌ها بسیار زیاد نباشد)،
- یا شبکه بازنشر پیرامون آن در مسیرهای مختلف، وزن/تکرار بالایی داشته باشد.

۳.۳.۲ الگوریتم HITS: امتیازهای Authority و Hub

در HITS هر نod می‌تواند دو نقش داشته باشد:

- Authority: منبع مناسب که توسط «هاب‌های مناسب» لینک می‌شود.
 - Hub: ارجاع‌دهنده مناسب که به «آتوریتی‌های مناسب» لینک می‌دهد.
- اگر A ماتریس مجاورت جهت‌دار باشد، بردارهای a (آتوریتی) و h (هاب) به صورت تکراری به‌روز می‌شوند:

$$a \leftarrow A^\top h, \quad h \leftarrow Aa,$$

و سپس نرم‌السازی انجام می‌شود. در نهایت، a تقریباً بردار ویژه غالب $A^\top A$ و h بردار ویژه غالب AA^\top خواهد بود.

تفسیر در این مسئله با توجه به جهت یال‌ها:

Authority ● به «منبع بودن» نزدیک است (زیرا یال‌های ورودی از بازنشرکننده‌ها وارد می‌شوند).

Hub ● به «بازنشرکننده/توزیع‌کننده بودن» نزدیک است (زیرا هاب‌ها معمولاً یال خروجی بیشتری دارند و به منابع متنوع‌تری ارجاع می‌دهند).

٤.٢ نتایج: ۱۰ نواد برتر برای هر معیار

امتیاز	نام کاربری	شناسه	رتبه
2944	ario1401	30011938	1
1937	pahlavireza	216372852	2
1693	indypersian	123754429	3
1661	iranintl	282756706	4
1597	irannewsagency0	1472285535	5
1505	tanhayektiger	1388061687423782914	6
1356	persiansn	1330903023236956160	7
1337	spyssee	794640325890072576	8
1297	aryammehr_2	1435123045258762247	9
1266	iranzerohour	1247099346923980803	10

جدول ۱۳: ده نواد برتر از نظر In-Degree در Twitter

امتیاز	نام کاربری	شناسه	رتبه
0.001404	pahlavireza	216372852	1
0.001259	iranintl	282756706	2
0.001068	persiansn	1330903023236956160	3
0.00105	irannewsagency0	1472285535	4
0.000943	indypersian	123754429	5
0.000919	iranzerohour	1247099346923980803	6
0.000894	spyssee	794640325890072576	7
0.000879	ario1401	30011938	8
0.000798	gohardaddy	242877634	9
0.000769	iranintlbrk	1408028581926381572	10

جدول ۱۴: ده نواد برتر از نظر PageRank در Twitter

امتیاز	نام کاربری	شناسه	رتبه
0.163144	pahlavireza	216372852	1
0.115223	iranintl	282756706	2
0.091782	irannewsagency0	1472285535	3
0.07495	persiansn	1330903023236956160	4
0.054523	spyssee	794640325890072576	5
0.03772	ario1401	30011938	6
0.033271	tanhayektiger	1388061687423782914	7
0.030063	indypersian	123754429	8
0.023122	iranzerohour	1247099346923980803	9
0.018846	gohardaddy	242877634	10

جدول ۱۵ : ده نود برتر از نظر Authority در Twitter (الگوریتم HITS)

امتیاز	نام کاربری	شناسه	رتبه
0.138361	ardavan_afshar	1312351517715810305	1
0.102169	ihj_mili	1413734399296780292	2
0.093017	saralunatika	1375827920837550080	3
0.07723	first_king_na	1280385358255319040	4
0.075807	tehran_lonely	1509384506852218880	5
0.071821	sholugh_cougar	1422182543797972992	6
0.064994	mohsen_irani9898	1439922022453983234	7
0.060425	a_sabz	155034696	8
0.059646	sara_raa_nd	1403927065867397120	9
0.052048	ahmad_nobakht	1471512496306272260	10

جدول ۱۶ : ده نود برتر از نظر Hub در Twitter (الگوریتم HITS)

امتیاز	نام کاربری	شناسه	رتبه
1126	choonkehghashange	522974468	1
859	farsna	1030137025	2
858	mehrnews	835721169	3
817	tasnimnews	602487017	4
743	kianmeli1	1099491462	5
632	akhbarefori	896735674	6
579	iribnews	512924353	7
552	isna94	965926350	8
450	akhbarmashhad	880886131	9
442	officialrezapahlavi	688305791	10

جدول ۱۷ : ده نود برتراز نظر In-Degree در Telegram

امتیاز	نام کاربری	شناسه	رتبه
0.005772	choonkehghashange	522974468	1
0.003753	mehrnews	835721169	2
0.002873	akhbarefori	896735674	3
0.002824	farsna	1030137025	4
0.002602	isna94	965926350	5
0.002482	officialrezapahlavi	688305791	6
0.002202	tasnimnews	602487017	7
0.002118	iribnews	512924353	8
0.002001	irdamavand	848395351	9
0.001945	akhbare_vije	633189783	10

جدول ۱۸ : ده نود برتراز نظر PageRank در Telegram

امتیاز	نام کاربری	شناسه	رتبه
0.996106	iran_uk1	588057738	1
0.055266	choonkehghashange	384825363	2
0.048872	kiammeli1	1127799605	3
0.043634	iran_uk4	1127810489	4
0.03852	iran_uk3	1127808734	5
0.02419	perspolisgram	1110380425	6
0.014961	officialrezapahlavi	688305791	7
0.014846	bccpersian	424064090	8
0.013444	stoptorture56	249523948	9
0.012222	cyrus_the_great1	886195634	10

جدول ۱۹: ده نود برتر از نظر Authority در Telegram (الگوریتم HITS)

امتیاز	نام کاربری	شناسه	رتبه
0.558104	iran_uk	393400352	1
0.423412	iran_uk4	1127810489	2
0.395632	iran_uk3	1127808734	3
0.314466	uk_iran5	1127799669	4
0.26644	uk_iran3	1127798679	5
0.265642	uk_iran	1127798549	6
0.249962	uk_iran4	1127798875	7
0.23216	uk_iran2	1127798806	8
0.18967	uk_iran1	1127798686	9
0.146317	uk_iran6	1130103010	10

جدول ۲۰: ده نود برتر از نظر Hub در Telegram (الگوریتم HITS)

امتیاز	نام کاربری	شناسه	رتبه
257	avatoday1	48345955395	1
199	iranintltv	48451610973	2
191	indypersian	10645304835	3
164	officialrezapahlavi	3289290420	4
123	nan	52063150785	5
115	pahlavireza	370763437	6
103	reza.davoudi.pahlavism	59056442747	7
102	iranintltv_brk	18414804311	8
102	courosh_iranzamin	46188627003	9
102	ioinc2023	54819812470	10

جدول ۲۱: ده نود برتر از نظر In-Degree در Instagram

امتیاز	نام کاربری	شناسه	رتبه
0.010547	avatoday1	48345955395	1
0.007975	iranintltv	48451610973	2
0.007794	indypersian	10645304835	3
0.007205	officialrezapahlavi	3289290420	4
0.005136	pahlavireza	370763437	5
0.004996	courosh_iranzamin	46188627003	6
0.003666	iranintltv_brk	18414804311	7
0.00344	ioinc2023	54819812470	8
0.00321	nan	52063150785	9
0.002889	iran_zerohour	28917252902	10

جدول ۲۲: ده نود برتر از نظر PageRank در Instagram

امتیاز	نام کاربری	شناسه	رتبه
0.58383	reza.davoudi.pahlavism	59056442747	1
0.505277	ioinc2023	54819812470	2
0.264205	pahlavireza	370763437	3
0.258767	iranintlv_brk	18414804311	4
0.253171	iran_zerohour	28917252902	5
0.253101	courosh_iranzamin	46188627003	6
0.251959	officialrezapahlevi	3289290420	7
0.123747	nan	52063150785	8
0.104569	indypersian	10645304835	9
0.10293	nan	48951286254	10

جدول ۲۳: ده نود برتر از نظر Authority در Instagram (الگوریتم HITS)

امتیاز	نام کاربری	شناسه	رتبه
0.509389	pixoooos	55620617987	1
0.463637	cyrus_the_great1	44304361374	2
0.270856	courosh_jabber	44725718304	3
0.216415	courosh_iranzamin	46187702414	4
0.187649	pahlavifamily	57449966571	5
0.147356	mahsaamini.n	40476242708	6
0.129115	zelensky_for_liberty	53607878108	7
0.126009	amirmadani_2021	52323639400	8
0.115501	cat_vision_arts	38160939375	9
0.102456	haj_khalifeh	47793328402	10

جدول ۲۴: ده نود برتر از نظر Hub در Instagram (الگوریتم HITS)

۵.۲ تحلیل همپوشانی و اختلاف ۴ لیست

برای بررسی تشابه/تفاوت بین دو لیست Top-10 از معیارهای مختلف، از دو شاخص استفاده شد:

• اندازه اشتراک: $|S_A \cap S_B|$

• شباهت ژاکارد: Jaccard

$$J(A, B) = \frac{|S_A \cap S_B|}{|S_A \cup S_B|}.$$

پلتفرم	دو معیار	اندازه اشتراک	شباهت ژاکارد
Twitter	InDegree و PageRank	8	0.667
Twitter	InDegree و Authority	8	0.667
Twitter	InDegree و Hub	0	0.000
Twitter	PageRank و Authority	9	0.818
Twitter	PageRank و Hub	0	0.000
Twitter	Authority و Hub	0	0.000
Telegram	InDegree و PageRank	6	0.429
Telegram	InDegree و Authority	2	0.111
Telegram	InDegree و Hub	0	0.000
Telegram	PageRank و Authority	1	0.053
Telegram	PageRank و Hub	0	0.000
Telegram	Authority و Hub	0	0.000
Instagram	InDegree و PageRank	10	1.000
Instagram	InDegree و Authority	7	0.538
Instagram	InDegree و Hub	0	0.000
Instagram	PageRank و Authority	7	0.538
Instagram	PageRank و Hub	0	0.000
Instagram	Authority و Hub	0	0.000

جدول ۲۵: مقایسه همپوشانی لیست‌های دهتایی بین معیارها در هر پلتفرم

۱.۵.۲ همپوشانی‌های مهم در Twitter

- همپوشانی InDegree و PageRank نود (Top-10)، شامل indypersian، gohardaddy، ario1401: ۸ نود
- . spyssee، persiansn، iranzerohour، irannewsagency0، iranintl

- همپوشانی Authority و InDegree نود (Top-10)، شامل indypersian، gohardaddy، ario1401: ۸ نود
- . spyssee، persiansn، iranzerohour، irannewsagency0، iranintl

- همپوشانی Authority و PageRank نود (Top-10)، شامل gohardaddy، ario1401: ۹ نود
- . spyssee، persiansn، pahlavireza، iranzerohour، irannewsagency0، iranintl، indypersian

تفسیر Twitter در Twitter، همپوشانی In-Degree و PageRank نسبتاً زیاد است؛ به این معنا که بسیاری از اکانت‌هایی که «بازنشرکننده‌های زیادی» دارند، هم‌زمان از سوی اکانت‌های «مهم‌تر» نیز بازنگش می‌شوند. با این حال، مشاهده اختلاف (برای مثال، حضور برخی اکانت‌ها صرفاً در In-Degree و نه در PageRank یا بالعکس) نشان می‌دهد که:

- یک اکانت می‌تواند بازنشرهای زیادی از سوی اکانت‌های کم‌اهمیت/پراکنده دریافت کند (افزایش In-Degree)؛ اما چون ارجاع‌دهنده‌ها خود مرکزی نیستند، آن به همان نسبت افزایش نیابد.
 - بر عکس، اکانتی با تعداد ارجاع‌دهنده کمتر، اگر از سوی ارجاع‌دهنده‌های بسیار مرکزی بازنشر شود، می‌تواند در PageRank برجسته شود.
- همچنین، فهرست‌های Hub تقریباً با سایر فهرست‌ها همپوشانی ندارند؛ زیرا Hub اساساً «بازنشرکننده‌های فعال/توزیع‌کننده‌ها» را شناسایی می‌کند، در حالی‌که Authority/PageRank/In-Degree بیشتر «منابع بازنشر» را برجسته می‌سازند.

۲.۵.۲ همپوشانی‌های مهم در Telegram

- همپوشانی akhbarmashhad، mehrnews و In-Degree PageRank شامل 6 نود (در Top-10) است.
- همپوشانی .choonkehghashange و officialrezapahlavi، farsna، isna94 شامل 2 نود (در Top-10) است.
- همپوشانی .choonkehghashange و kianmeli1 شامل 1 نود (در Top-10) است.

تفسیر Telegram در Telegram، همپوشانی In-Degree و PageRank در حد متوسط است؛ بدین معنا که برخی کانال‌ها از نظر تعداد کانال‌های ارجاع‌دهنده پر تکرار دیده می‌شوند، اما در PageRank کانال‌هایی بر جسته ترند که از سوی کانال‌های مرکزی‌تر ارجاع دریافت کرده‌اند. نکته مهم، تمرکز شدید Authority بر چند اکانت مشخص (به‌ویژه یک اکانت با امتیاز بسیار بالا) است. چنین رفتاری در HITS می‌تواند زمانی رخ دهد که:

- یک «زیرگراف» نسبتاً بسته وجود داشته باشد که در آن چند هاب قوی تقریباً به‌طور مداوم به یک/چند مرجع خاص لینک دهنده؛ در این حالت، تقویت متقابل هاب/آتوریتی موجب غالب شدن همان زیرگراف در بردار ویژه می‌شود.
- به صورت شهودی، اگر گروهی از کانال‌های بازنشرکننده (هاب‌ها) تقریباً همواره از یک کانال مشخص نقل قول/فوروارد کنند، آن کانال در Authority بسیار پررنگ می‌شود؛ حتی اگر در کل شبکه، تعداد ارجاع‌دهنده‌های یکتای آن چندان زیاد نباشد.

۳.۵.۲ همپوشانی‌های مهم در Instagram

- همپوشانی In-Degree و PageRank شامل 10 نود (در Top-10) است: courosh_iranzamin، avatoday1، pahlavireza، officialrezapahlavi، nan، iranintlv_brk، iranintlv، ioinc2023، indypersian و reza.davoudi.pahlavism.
- همپوشانی In-Degree و Authority شامل 7 نود (در Top-10) است: indypersian، courosh_iranzamin و .reza.davoudi.pahlavism.
- همپوشانی Authority و PageRank شامل 7 نود (در Top-10) است: indypersian، courosh_iranzamin و .reza.davoudi.pahlavism.

تفسیر Instagram در اینکه لینک‌های بازنثر به صورت مصنوعی تولید شده‌اند، به طور هم‌زمان انتظار می‌رود:

- ساختار گراف کم‌تراکم‌تر باشد (یال‌های کمتر)،
- و معیارهای مبتنی بر «ماتریس لینک» (مانند PageRank و HITS) حساسیت بیشتری نسبت به نحوه ساخت لینک‌ها نشان دهد.

با این حال، منطق کلی همچنان برقرار است: Hub بیشتر نقش پخش‌کننده‌ها را آشکار می‌کند و Authority منابعی را بر جسته می‌سازد که توسط هاب‌های قوی لینک می‌شوند.

۶.۲ نگاهی به محتوای مطالب (TF-IDF روی فایل‌های XLSX)

در داده‌های ضمیمه برای هر پلتفرم، 5 فایل XLSX وجود دارد که هر کدام شامل 5000 مطلب پرانتشار در روزهای 15 تا 19 دی هستند. برای تفسیر دقیق‌تر اختلاف میان فهرست‌ها، یک تحلیل متنی ساده انجام شد:

- برای چند اکانت شاخص (از میان نودهای Top-10)، همه متن‌های موجود در فایل‌های XLSX گردآوری شد.
- سپس با TF-IDF (پس از حذف برخی واژه‌های توقف) واژه‌های شاخص هر اکانت استخراج شد.

۱.۶.۲ نمونه واژه‌های شاخص در Twitter

- irannewsagency0: خیابان، مردمی، خیش، ملی، دی، مردم، قیام، حضور.
- iranintl: ایران، نشان، چهارشنبه، دی، ۱۷، مردم، خیابان، ویدیوهای.
- tanhayektiger: جاویدشاه، ایران، انقلاب، ملی، این_آخرین_نبرده_پهلوی_برمیگرده، فراخوان، بازار، فردا.
- ario1401: جاویدشاه، مردم، گاز، خیابان، دست، ایران، کردین، الان.

برداشت محتوایی Twitter بر اساس واژه‌های شاخص می‌توان مشاهده کرد که برخی اکانت‌ها بیشتر نقش «خبر/گزارش صحنه» دارند (برای نمونه، واژه‌هایی مانند «خیابان»، «خیش»، «قیام» و «ویدیوها») و برخی دیگر نقش «فراخوان/هشتگ محور» (مانند «جاویدشاه» یا هشتگ‌های خاص). این تمایز می‌تواند توضیح دهد که چرا بعضی اکانت‌ها با وجود بازنشر زیاد، در PageRank (که به کیفیت لینک‌دهندگان حساس است) دقیقاً همان رتبه را کسب نمی‌کنند.

۲.۶.۲ نمونه واژه‌های شاخص در Telegram

- tasnimnews: مردم، امریکا، تهران، ایران، رییس، ونزویلا، شود، امنیت.
- akhbarefori: ایران، مجازی، فضای، امریکا، ترامپ، مردم، ونزویلا، اتش.
- iribnews: مردم، امریکا، ایران، ونزویلا، رییس، امنیت، تهران، جمهور.
- farsna: بازار، ایران، پلیس، مردم، امریکا، ترامپ، کند، شود.
- mehrnews: امنیت، ایران، مسلح، ترویریست، اتش، مردم، استان، کشور.
- iran_uk1: ایران، شاه، حکومت_کثیف_اخوندی، رضاشاھ_روحت_شاد، تلگرامی، بزرگ، شاهزاده_رضاپهلوی، شاهنشاھ_اریامهر.

برداشت محتوایی Telegram در کانال‌های خبری رسمی‌تر، واژه‌های شاخص می‌تواند حول رویدادهای خبری خارجی/سیاسی نیز شکل گیرد (مانند «امریکا»، «ونزویلا»، «ترامپ» و «امنیت»). در مقابل، کانال‌هایی که در Hub/Authority به شدت برجسته شده‌اند، ممکن است از نظر گفتمنی (برای مثال، هشتگ‌ها/کلیدواژه‌های خاص) یک خوشبی‌سیار همگن بسازند و همین خوشبی HITS غالب شود.

۳.۶.۲ نمونه واژه‌های شاخص در Instagram

- avatoday1 ●: خیزش، مردم، ازادی، سراسری، رژیم، مزدوران، روز، خواهان.
- iranintltv ●: ایران، مردم، ملی، نشان، دی، خیزش، ایرانیان، خیابان.
- indypersian ●: ایران، مردم، پهلوی، رضا، شاهزاده، معترضان، جمهوری، ملی.
- courosh_iranzamin ●: ایران، جاویدشاه، همبستگی ملی، همه به ایران فکر کنیم ، صدای نوین ایران، نیست، خامنه، پاینده ایران‌ایران را پس میگیریم اتحاد برای ازادیرضا شاه دومگارد جاویدان.

برداشت محتوایی Instagram واژه‌های شاخص در برخی صفحات به «اعتصاب/بازاریان/سراسری» اشاره دارد که با ماهیت بسیج و انتشار فراخوان‌ها سازگار است. با این وجود، به دلیل مصنوعی بودن لینک‌های بازنشر در Instagram، در تفسیر دقیق نقش‌های Authority/Hub لازم است احتیاط بیشتری رعایت شود.

تحلیل k-core و k-shell پس از حذف جهت یال‌ها

۱.۳ هدف مسئله و چرایی حذف جهت

در سؤال ۳، لازم است جهت‌دار بودن گراف بازنثر نادیده گرفته شود و گراف به صورت «بدون جهت» تحلیل گردد؛ سپس با اجرای تجزیه k-core (یا core decomposition)، برای هر نو德 مقدار k-shell محاسبه شود و در نهایت بررسی گردد که نودهای دارای بیشترین k-shell از نظر محتوایی چه ویژگی‌هایی دارند.

نکته مفهومی مهم: در سؤال ۲، معیارهایی مانند PageRank، In-Degree و HITS به جهت یال‌ها وابسته بودند؛ اما در این سؤال، در k-core عمدتاً جهت حذف می‌شود تا «چسبندگی/انسجام ساختاری» سنجیده شود؛ یعنی نودهایی که در یک زیرشبکه متراکم، با تعداد زیادی همسایه مشترک و پیوندهای فراوان قرار گرفته‌اند.

۲.۳ تعاریف دقیق و روابط ریاضی

۱.۲.۳ تعریف k-core

برای یک گراف بدون جهت $G = (V, E)$ ، زیرگراف بیشینه‌ای است که در آن تمام نودها (در همان زیرگراف) درجه‌ای حداقل k دارند. اگر زیرگراف $H \subseteq G$ باشد، شرط k-core به صورت زیر است:

$$\forall v \in V(H) : \deg_H(v) \geq k.$$

بر اساس این تعریف، ممکن است نودي در کل گراف درجه بزرگی داشته باشد، اما اگر در زیرگراف‌های متراکم حضور نداشته باشد، در پوسته‌های بزرگ‌تر باقی نمی‌ماند.

۲.۲.۳ تعریف core number و ارتباط آن با k-shell

یک نوD v (که در این گزارش از آن با عنوان index k-shell نیز یاد می‌شود) برابر است با بزرگ‌ترین k که نوD v عضو k-core مرتبه k باشد:

$$c(v) = \max\{k \mid v \in \text{k-core}(k)\}.$$

به صورت شهودی، اگر $k(v) = c$ ، نود v تا پوسته k می‌تواند در فرایند «پوست‌اندازی» باقی بماند و در مقادیر بزرگ‌تر حذف می‌شود. به همین دلیل، در بسیاری از منابع گفته می‌شود «نود در k-shell شماره k قرار دارد» اگر $k(v) = c$.

۳.۲.۳ الگوریتم محاسبه (پوست‌اندازی/هرس تکراری)

روش استاندارد محاسبه k-core یک فرایند هرس تکراری است:

۱. برای یک k ثابت، تمام نودهایی که درجه آن‌ها کمتر از k است حذف می‌شوند.
۲. حذف یک نود، درجه همسایه‌های آن را کاهش می‌دهد؛ بنابراین ممکن است نودهای دیگری نیز اکنون درجه کمتر از k پیدا کنند و لازم باشد حذف شوند.
۳. این روند تا زمانی ادامه می‌یابد که هیچ نودی با درجه کمتر از k باقی نماند.

خروجی پایدار، همان k-core است. سپس با افزایش k ، هسته‌های عمیق‌تر به دست می‌آیند و در نهایت، بزرگ‌ترین k که هنوز هسته خالی نشده است با k_{\max} یا «هسته اصلی» معروفی می‌شود.

۴.۲.۳ چرایی تفاوت k-shell با «محبوبیت»

k-shell بیش از آن‌که «محبوبیت» را بسنجد، «جایگاه در ساختار متراکم» را اندازه‌گیری می‌کند و صرفاً با «بیشترین بازنشر/بیشترین درجه» یکسان نیست. در گراف‌های بازنشر ممکن است:

- یک اکانت بسیار بازنشر شود، اما عمدتاً توسط حساب‌های پراکنده و کم ارتباط؛ در این حالت درجه آن افزایش می‌یابد، ولی الزاماً در هسته متراکم قرار نمی‌گیرد.
- یک گروه از اکانت‌ها که به طور گسترده از یکدیگر بازنشر می‌کنند/به یکدیگر لینک می‌دهند، یک هسته متراکم بسازند؛ حتی اگر محبوبیت عمومی آن‌ها از چند اکانت بزرگ خبری کمتر باشد.

۳.۳ روش اجرای مسئله روی داده تمرین

۱.۳.۳ گام ۱: تبدیل گراف جهت‌دار به گراف بدون جهت

یال‌های داده اصلی جهت‌دار هستند (از بازنشر کننده به مرجع). در سؤال ۳، جهت حذف شد و یک گراف undirected ساخته شد. برای جلوگیری از حذف اطلاعات مربوط به تکرار بازنشر،

وزن‌ها نگه داشته شد و اگر بین دو نود u و v چند بار ارتباط گزارش شده بود (یا هر دو جهت وجود داشت)، وزن یال بدون جهت به صورت تجمعی محاسبه شد:

$$w_{uv} \leftarrow \sum w_{uv}^{(i)}.$$

توجه: الگوریتم استاندارد k-core وزن را وارد محاسبه نمی‌کند و تنها «درجه بدون وزن» را ملاک قرار می‌دهد. بنابراین، وزن برای تحلیل‌های تکمیلی (برای نمونه، weighted degree) حفظ شد، اما k-shell بر اساس درجه ساده محاسبه گردید.

۲.۳.۳ گام 2: محاسبه k-shell برای هر نод

برای هر پلتفرم:

۱. فایل gdf خوانده شد و فهرست نودها و یال‌ها استخراج شد.
۲. گراف بدون جهت ساخته شد.
۳. با تابع core_number مقدار $c(v)$ برای همه نودها محاسبه شد.
۴. مقدار $k_{\max} = \max_v c(v)$ به عنوان پوسته بیشینه گزارش شد.
۵. نودهای با $c(v) = k_{\max}$ استخراج شدند (این‌ها «نودهای پوسته بیشینه» هستند).

۳.۳.۳ گام 3: تحلیل محتوایی نودهای پوسته بیشینه با TF-IDF

برای پاسخ به بخش «ویژگی‌های موضوعی/محتوایی»، از فایل‌های xlsx روزهای مختلف استفاده شد:

۱. تمام پست‌ها در چند روز (برای هر پلتفرم) خوانده شد.
۲. پست‌ها بر اساس username گروه‌بندی شدند.
۳. برای هر username (در میان نودهای پوسته بیشینه)، همه متن‌ها به یک سند واحد تبدیل شد.
۴. با مدل TF-IDF (کیسه‌واژه‌ها) واژه‌ها/عبارت‌های شاخص هر اکانت استخراج شد.

فرمول TF-IDF برای یک واژه t در سند d و مجموعه اسناد D به صورت زیر است:

$$\text{tfidf}(t, d, D) = \text{tf}(t, d) \times \log \left(\frac{|D|}{1 + \text{df}(t)} \right),$$

که در آن $\text{tf}(t, d)$ فراوانی واژه در سند و $\text{df}(t)$ تعداد اسنادی است که واژه در آنها ظاهر شده است.

۴.۳ نتایج عددی و مقایسه سه پلتفرم

پلتفرم	تعداد نودهای پوسته بیشینه	k_{\max}	تعداد نود
Twitter	290	29	85021
Telegram	125	38	48632
Instagram	81	13	35825

جدول ۲۶: خلاصه نتایج k -core (گراف بدون جهت) برای سه پلتفرم

بر اساس جدول ۲۶، بیشترین پوسته در Telegram برابر 38 است که از Twitter (29) بزرگ‌تر و از Instagram (13) به مرتب بزرگ‌تر است. این اختلاف معمولاً نشان‌دهنده تفاوت در «چگالی ساختاری» شبکه بدون جهت است: هرچه k_{\max} بزرگ‌تر باشد، یعنی زیرگرافی بسیار متراکم‌تر وجود دارد که در آن همه نودها (در همان زیرگراف) دست‌کم k_{\max} همسایه دارند.

۵.۳ توزیع k -shell (کوانتاپل‌ها و تفسیر)

مقدار	کوانتاپل	k -shell
min	0	
25%	0	
50%	1	
75%	1	
90%	5	
95%	9	
99%	22	
max	29	

جدول ۲۷: کوانتاپل‌های k -shell برای Twitter

مقدار	کوانتایل k-shell
min	0
25%	0
50%	0
75%	1
90%	3
95%	7
99%	25
max	38

جدول ۲۸: کوانتایل‌های k-shell برای Telegram

مقدار	کوانتایل k-shell
min	0
25%	0
50%	0
75%	0
90%	1
95%	1
99%	7
max	13

جدول ۲۹: کوانتایل‌های k-shell برای Instagram

برداشت از جداول کوانتایل در هر سه پلتفرم، بخش بزرگی از نودها در پوسته‌های بسیار پایین قرار دارند (کوانتایل‌های پایین برابر ۰ یا ۱ هستند). این الگو در شبکه‌های بازنشر رایج است: شمار زیادی نod «حاشیه‌ای» وجود دارد که تنها یکبار یا چندبار در بازنشر ظاهر شده‌اند و در مقابل، تعداد محدودی نod در هسته‌های عمیق‌تر قرار می‌گیرند.

تفاوت مهم آن است که در Telegram، کوانتایل ۹۹% تا ۲۵ افزایش می‌یابد و بیشینه ۳۸ است؛ یعنی دُمی سنگین از نودهای با coreness بالا مشاهده می‌شود. در Instagram، مقدار ۹۹% برابر ۷ و بیشینه ۱۳ است که نشان‌دهنده هسته‌های کم عمق‌تر است؛ وضعیتی که علاوه بر ویژگی‌های پلتفرم، می‌تواند تحت تأثیر ساخت مصنوعی لینک‌ها نیز قرار گرفته باشد.

۶.۳ نودهای پوسته بیشینه: 10 نود با بیشترین درجه در هر پلتفرم

ردیف	شناسه	نام کاربری	درجه	درجه وزندار
1	1613903983030026240	ario1401	2949	3480
2	797159363811147776	tanhayektiger	2692	3063
3	1638372580464685056	aryammehr_2	2592	4205
4	728252911675944961	iranintl	2078	3978
5	38154763	pahlavireza	1848	2103
6	1550545991778340865	irannewsagency0	1672	4864
7	1925636410557521920	gohardaddy	1667	2002
8	1065658304884547585	indypersian	1549	2952
9	1408558266	aamirhossein	1343	2149
10	410718052	manotonews	1325	2139

جدول ۳۰: ده نود برتیر k-shell (پوسته بیشینه) در Twitter بر اساس درجه کل

ردیف	شناسه	نام کاربری	درجه	درجه وزندار
1	2037713460	choonkehghashange	1129	7850
2	1562786584	kianmeli1	474	1555
3	1798869200	hadafazadi2022	402	1157
4	1007936468	sepehrzadi	395	1099
5	1212387582	eterazebar	360	1290
6	1107431118	javanmardi77	307	1384
7	1503443333	mardom_khiaban	281	529
8	1648796579	iliaen	269	488
9	1254295693	protest_iran	268	839
10	1662852594	ghyam_meli	264	786

جدول ۳۱: ده نود برتیر k-shell (پوسته بیشینه) در Telegram بر اساس درجه کل

ردیف	شناسه	نام کاربری	درجه وزندار	درجه
1	10322006202	avatoday1	265	644
2	1949382920	officialrezapahalvi	245	483
3	5496840166	iranintltv	148	309
4	59807915743	gord_afarid_tv	124	142
5	13521455499	indypersian	119	304
6	364081256	manotoofficial	111	191
7	49743276108	azaddokht_	98	178
8	57725855438	babak_b2006	83	107
9	56463117828	iran.janam.2022	79	224
10	7219800846	shemiranisaied	77	113

جدول ۳۲: ده نویسنده بیشینه در Instagram بر اساس درجه کل پستهای k-shell (پستهای پوششی)

۷.۳ تحلیل محتوایی/ موضوعی نویسندگان پوششی

۱.۷.۳ روش خواندن متن و محدودیت‌ها

تحلیل محتوایی این سؤال بر اساس فایل‌های *xlsx* انجام شد که «پست‌های پرانشمار» را در چند روز مشخص پوشش می‌دهند. بنابراین:

- اگر یک اکانت در هسته ساختاری قرار داشته باشد اما در آن چند روز، پست پرانشمار اندکی منتشر کرده باشد، تعداد پست‌های متناظر آن در تحلیل متنی کاهش می‌یابد و در نتیجه کلیدواژه‌های استخراج شده ممکن است ناپایدار باشند.
- روش TF-IDF قادر به معناشناختی عمیق است و بیشتر برای شناسایی «کلیدواژه‌های شاخص» کاربرد دارد؛ برای تحلیل پیشرفته‌تر می‌توان از topic modeling embedding استفاده کرد که در بخش ایده‌های ابتکاری به آن اشاره می‌شود.

۲.۷.۳ کلیدواژه‌های شاخص نویسندگان پوششی در Twitter

- ario1401: تعداد پست‌های یافتشده در *xlsx* برابر 16. کلیدواژه‌های TF-IDF: رو، کردن، پهلوی برمیگردد، برمیگردد، همین الان، الان، خیابان، میکن، خیابان رو، نبرده پهلوی، آخرین نبرده، نبرده.

- tanhayektiger: تعداد پست‌های یافتشده در *xlsx* برابر 24. کلیدواژه‌های TF-IDF: جاویدشاه، ایران، برمیگردد جاویدشاه، انقلاب ملی، انقلاب، ملی، جاویدشاه بازار، رژیم، بازار، پهلوی برمیگردد، برمیگردد، پادشاه.

● aryammehr_2: تعداد پست‌های یافت شده در *xlsx* برابر 99. کلیدواژه‌های TF-IDF: جاویدشاه، مردم، رو، longlivetheshah، longlivetheshah ۱۷ دی، دریافتی، جاویدشاه دریافتی، میکنن، دارن، رضاشاه، جاویدشاه جاویدشاه.

● iranintl: تعداد پست‌های یافت شده در *xlsx* برابر 197. کلیدواژه‌های TF-IDF: ایران اینترنشنال، اینترنشنال، رسیده ایران، نشان، ایران، رسیده، رسیده، معترضان، اعتصاب، سردر.

● pahlavireza: تعداد پست‌های یافت شده در *xlsx* برابر 3. کلیدواژه‌های TF-IDF: ایران، است، خواهید، ثبت نام، مردم، ملت، اید، ثانیه، آینده، آدرس، بهمن، نتیجه.

● irannewsagency0: تعداد پست‌های یافت شده در *xlsx* برابر 336. کلیدواژه‌های TF-IDF: تروریستی جمهوری، رژیم تروریستی، مردمی، قیام مردمی، تروریستی، رژیم، قیام، تهران، مردم، سراسری، جمهوری، بیانیه.

● gohardaddy: تعداد پست‌های یافت شده در *xlsx* برابر 24. کلیدواژه‌های TF-IDF: جاویدشاه، جمهوری، دیگه، دارن، باید، رو، حتی، حتما، خیابان، مردم، رژیم، سراسری.

● indypersian: تعداد پست‌های یافت شده در *xlsx* برابر 207. کلیدواژه‌های TF-IDF: اعتراضات ملی، فارسی، ملی، است، معترضان، اعتراضات، ایران، گزارش، ویدیو، اعلام، شهر، دی.

● aamirhoossein: تعداد پست‌های یافت شده در *xlsx* برابر 93. کلیدواژه‌های TF-IDF: اعتراضات سراسری، اعتراضات، سراسری، نشر، نشرحداکثري، حدакثري، خیابان، فراخوان، دی، مردم، رژیم، اعتصاب.

● manotonews: تعداد پست‌های یافت شده در *xlsx* برابر 99. کلیدواژه‌های TF-IDF: ۱۴۰۴، ۱۴۰۴، دی، اعتراضات، دی، اعتراضات شبانه، شبانه، فراخوان، سراسری، بازار، ۱۷، ۱۷، دی، تهران.

تفسیر محتوایی Twitter در Twitter، بخش قابل توجهی از نودهای پوسته بیشینه از نظر کلیدواژه‌ها حول چند محور مشترک متمرکز هستند:

● خبر و پوشش ویدئویی رویدادها: واژگانی مانند «دریافتی»، «ویدیو»، «رسیده»، «اعتراضات» و «خیابان».

● بسیج و فراخوان: حضور پررنگ واژه‌هایی همچون «فراخوان»، «اعتصاب»، «سراسری» و «بازار».

● شعارها و هشتگ‌های سیاسی: تکرار واژه «جاویدشاه» و ترکیبات وابسته به آن در چند اکانت پرتکرار.

از منظر ساختاری، این الگو با انتظار نظری از k-core سازگار است؛ هسته متراکم شبکه معمولاً شامل اکانت‌هایی است که در زنجیره‌های بازنشر متقابل، هم نقش منع خبری و هم نقش بازنشرکننده فعال را ایفا می‌کنند و به صورت گسترده به یکدیگر متصل‌اند.

۳.۷.۳ کلیدواژه‌های شاخص نودهای پوسته بیشینه در Telegram

● choonkehghashange: تعداد پست‌های یافت شده در xlsx برابر 1. کلیدواژه‌های TF-IDF: خیابان سعدی، انداز، سعدی، جاویدشاه، طنین، سه شنبه، دی، مشهد، جمعیت، شعار، مردم، ایران.

● kianmeli1: تعداد پست‌های یافت شده در xlsx برابر 66. کلیدواژه‌های TF-IDF: کیان، ملی، قیام، اعتصاب، خبرفوری، مردم، اتش، به اختیار، بازار، دی، تهران، سراسری.

● hadafazadi2022: تعداد پست‌های یافت شده در xlsx برابر 85. کلیدواژه‌های TF-IDF: اعتصاب، همه، دی، سلام، ۱۷ دی، مغازه، بازار، ۱۷، تعطیل، کسبه، سراسری، تهران.

● sepehrzadi: تعداد پست‌های یافت شده در xlsx برابر 484. کلیدواژه‌های TF-IDF: sepehrzadi 1404، دی، مردم، ۱۶ دی، ۱۷ دی، ۱۸ دی، تهران، دریافتی، ویدیوی دریافتی، گزارش، خیابان.

● eterazebazar: تعداد پست‌های یافت شده در xlsx برابر 38. کلیدواژه‌های TF-IDF: بازار، است، اعتصاب، ۱۷ دی، ۱۷، دی، مغازه، تعطیلی، کسبه، تهران، سراسری، بازاریان.

● javanmardi77: تعداد پست‌های یافت شده در xlsx برابر 87. کلیدواژه‌های TF-IDF: خیزش، خیزش سراسری، مردم، رژیم، است، سراسری، مزدوران، قیام، تهران، اعتصاب، بازار، جمهوری.

● mardom_khiaban: تعداد پست‌های یافت شده در xlsx برابر 51. کلیدواژه‌های TF-IDF: مردم، مردم خیابان، رو، کanal، خیابان، است، ایران، تهران، همه، قیام، دی، سراسری.

● تعداد پست‌های یافت شده در `xlsx` برابر 12. کلیدواژه‌های TF-IDF: دی مشهد، گاهی، ۱۷ دی، بعد ظهر، ۱۷، چهارشنبه، اعتراضات، شعار، خیابان، مردم، تجمع، ایران.

● تعداد پست‌های یافت شده در `xlsx` برابر 23. کلیدواژه‌های TF-IDF: protest_iran، اگه، and، زخم، رو، وحش سرکوبگر، حتما، رژیم، مردم، خیابان، سراسری، اعتصاب.

● تعداد پست‌های یافت شده در `xlsx` برابر 14. کلیدواژه‌های TF-IDF: ماه ۲۵۸۴ شاهنشاهی، هجدهم دی، شه، دی، فراخوان، سراسری، مردم، ایران، بازار، قیام، اعتصاب.

تفسیر محتوایی **Telegram** در **Telegram**، پوسته بیشینه عمق بیشتری دارد (38) و از نظر کلیدواژه‌ها چند الگوی غالب مشاهده می‌شود:

● گزارش لحظه‌ای و دریافتی از شهرها: کلیدواژه‌هایی مانند «دریافتی»، «ویدیوی دریافتی»، تاریخ‌های روزانه (مانند «۱۶ دی» و «۱۷ دی») و نام مکان‌ها.

● فراخوان اعتصاب و بازار: تکرار واژگانی نظیر «اعتصاب»، «بازار»، «معازه»، «تعطیلی» و «کسبه».

● چارچوب‌های هویتی و شعاری: واژه‌هایی مانند «قیام» و ترکیبات نمادین در برخی کانال‌ها.

عمق بیشتر هسته در **Telegram** می‌تواند بازتاب این واقعیت باشد که کانال‌ها الگوهای فوروارد «یکدیگر-محور» (و نه صرفاً منبع-محور) ایجاد کرده‌اند. در چنین شرایطی، شبکه بدون جهت متراکم‌تر شده و مقدار k_{\max} افزایش می‌یابد.

٤.٧.٣ کلیدواژه‌های شاخص نودهای پوسته بیشینه در **Instagram**

● تعداد پست‌های یافت شده در `xlsx` برابر 91. کلیدواژه‌های TF-IDF: خیش، رژیم، است، مزدوران، خیش سراسری، مردم، اعتراضات، تهران، سراسری، جمهوری، دی، بازار.

● تعداد پست‌های یافت شده در `xlsx` برابر 3. کلیدواژه‌های TF-IDF: officialrezapahlavi، the، iran، of، people، regime، nation، freedom، and، خواهید، ملت، the

- iranintltv: تعداد پست‌های یافتشده در xlsx برابر 123. کلیدواژه‌های TF-IDF: اینترنشنال، ایران اینترنشنال، ایران، رسیده ایران، رسیده، ویدیو، گزارش، معارضان، اعتصاب، تهران، دی، سراسری.
 - gord_afarid_tv: در فایل‌های xlsx یافت نشد؛ کلیدواژه‌ای استخراج نشد.
 - indypersian: تعداد پست‌های یافتشده در xlsx برابر 137. کلیدواژه‌های TF-IDF: اعتراضات ملی، فارسی، است، ملی، شاهزاده رضا، ایران، گزارش، ویدیو، معارضان، دی، تهران، سراسری.
 - manotoofficial: تعداد پست‌های یافتشده در xlsx برابر 49. کلیدواژه‌های TF-IDF: ۱۴۰۴، دی، اعتراضات، ۱۷، ۱۷ دی، شاهزاده، فراخوان، سراسری، بازار، تهران، اعتصاب.
 - azaddokht_: تعداد پست‌های یافتشده در xlsx برابر 39. کلیدواژه‌های TF-IDF: ایران، جاوید، جاوید شاه، شاه، پهلوی، پاینده ایران، مردم، سراسری، دی، تهران، اعتصاب، قیام.
 - babak_b2006: تعداد پست‌های یافتشده در xlsx برابر 31. کلیدواژه‌های TF-IDF: دیماه، ۱۴۰۴، اعدام، زمانی، شورشگران، شهید، اعتراضات، تهران، سراسری، دی، بازار، اعتصاب.
 - iran.janam.2022: تعداد پست‌های یافتشده در xlsx برابر 33. کلیدواژه‌های TF-IDF: خیزش، مردم، رژیم، خیزش سراسری، است، ای، دی، سراسری، تهران، اعتراضات، بازار، اعتصاب.
 - shemiranisaied: تعداد پست‌های یافتشده در xlsx برابر 12. کلیدواژه‌های TF-IDF: بازار، حکومت جنایتکار، جنایتکار اسلامی، جنایتکار، تهران، مردم، قیام، بازار بزرگ، حکومت، بزرگ، شهید، کسبه بازاریان.
- تفسیر محتوایی Instagram** در Instagram، دو نکته باید هم‌زمان مدنظر قرار گیرد:
- از منظر ساختاری، پوسته بیشینه کوچک‌تر و کم‌عمق‌تر است (13) که با کم‌چگال بودن شبکه سازگار است.
 - از منظر داده، لینک‌های بازنشر به صورت مصنوعی ساخته شده‌اند؛ بنابراین تفسیر هسته ساختاری باید با احتیاط انجام شود، زیرا بخشی از ساختار به روش ساخت لینک وابسته است.

با وجود این محدودیت، کلیدوازه‌ها نشان می‌دهد نودهای پوسته بیشینه عمدتاً حول محورهایی مانند «اعتراضات»، «خیش سراسری»، «بازار/اعتصاب»، «جمهوری اسلامی» و برخی شعارها یا هشتگ‌ها متمرکز هستند. این امر نشان می‌دهد که حتی در مدل‌سازی ساده نیز هسته شبکه بیشتر با محتوای خبری-سیاسی و بسیج اجتماعی گره خورده است.

۸.۳ جمع‌بندی تفسیری: k-shell بالا چه معنایی دارد؟

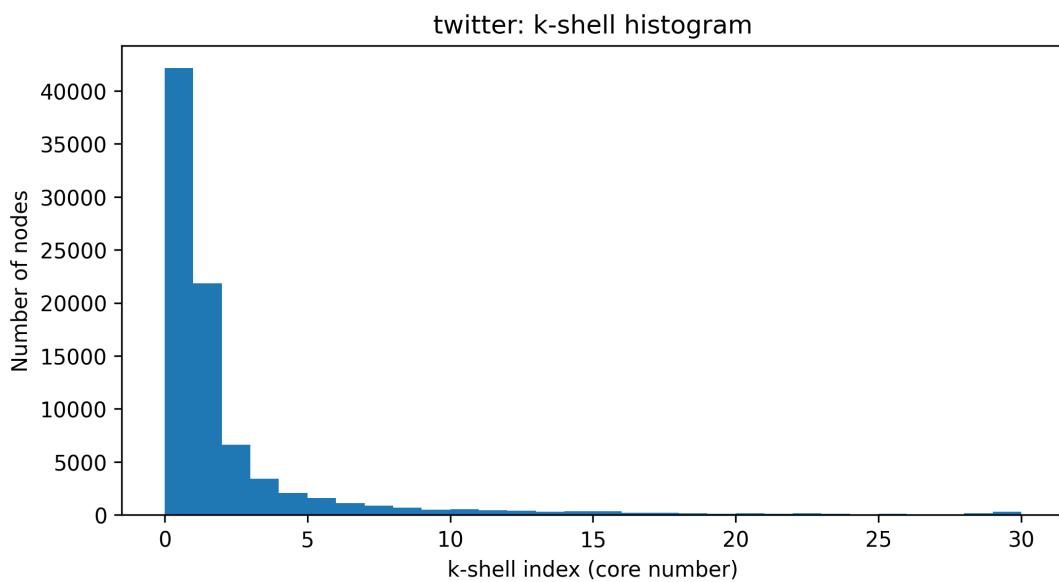
از منظر شبکه‌ای، نودهای دارای بیشترین مقدار k-shell لزوماً بیشترین بازنشر را ندارند، بلکه عموماً در یک حلقه متراکم از بازنشرهای متقابل قرار گرفته‌اند. در داده این تمرین، هسته متراکم در هر سه پلتفرم عمدتاً شامل اکانت‌هایی است که یا:

- نقش «منبع خبر، ویدیو یا گزارش» را ایفا می‌کنند،
- یا نقش «پخش‌کننده و هماهنگ‌کننده فراخوان‌ها» (مانند اعتصاب، تجمع و بازار)،
- یا در یک خوش‌گفتمانی/شعاری با بازنشر متقابل جای گرفته‌اند.

بر این اساس، k-core در چنین دیتاست‌هایی بیش از آنکه صرفاً شاخصی از محبوبیت باشد، ابزاری برای شناسایی «هسته اطلاعاتی و هماهنگی» در شبکه محسوب می‌شود.

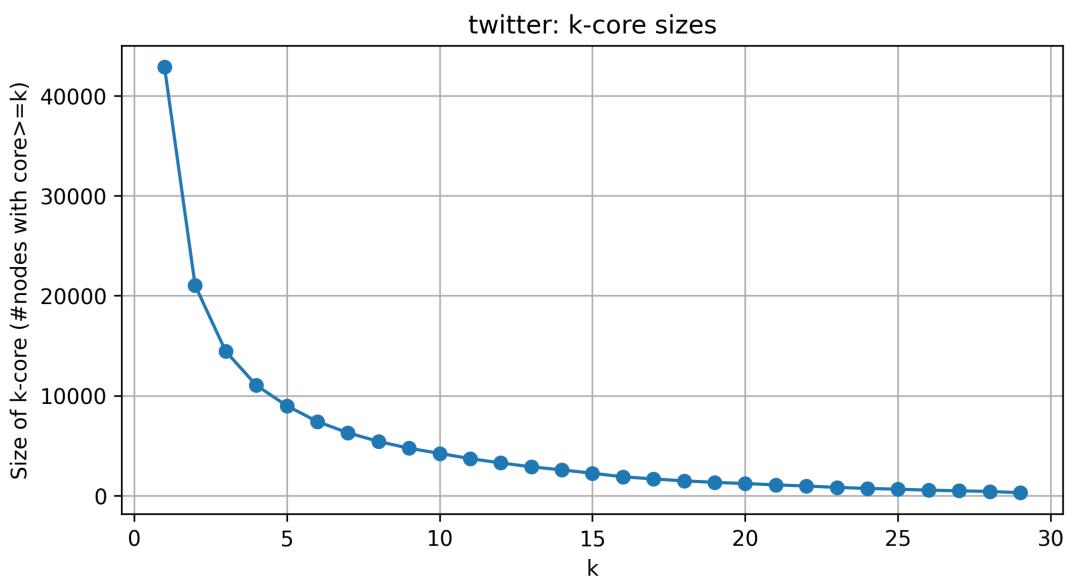
۹.۳ شکل‌ها

- توزیع فراوانی k-shell در Twitter.



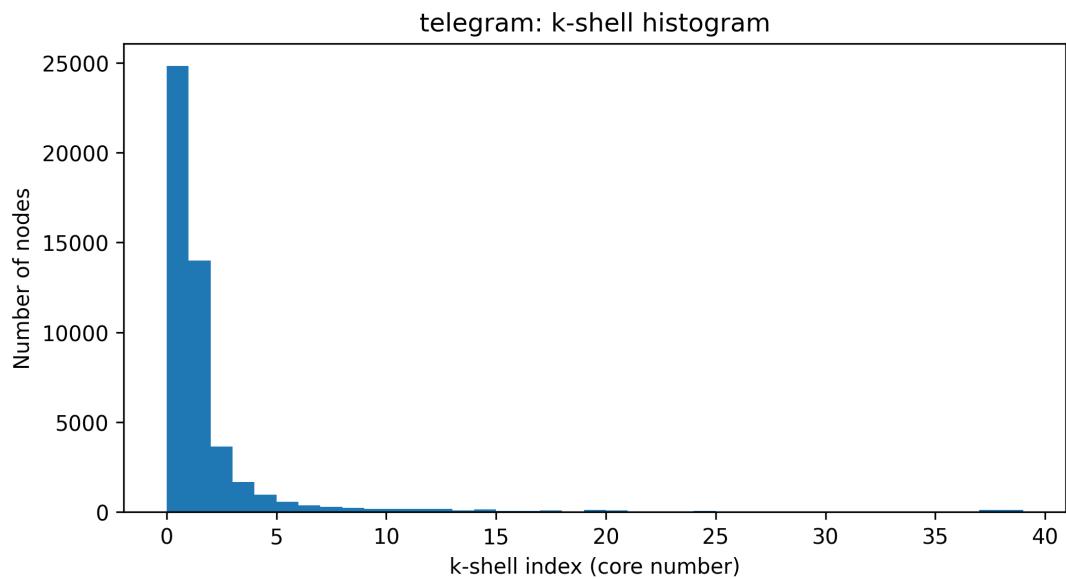
شکل ۱: توزیع k-shell در Twitter

- منحنی اندازه k-core (اندازه زیرگراف با شرط درجه $\leq k$) در Twitter.



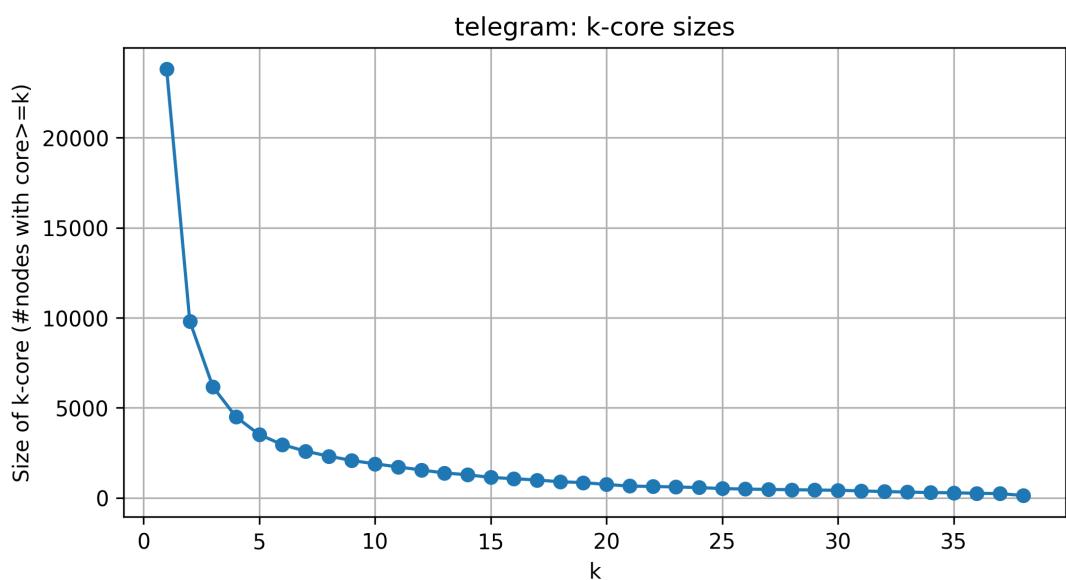
شکل ۲: اندازه k-core‌ها در Twitter

- توزیع فراوانی k-shell در Telegram.



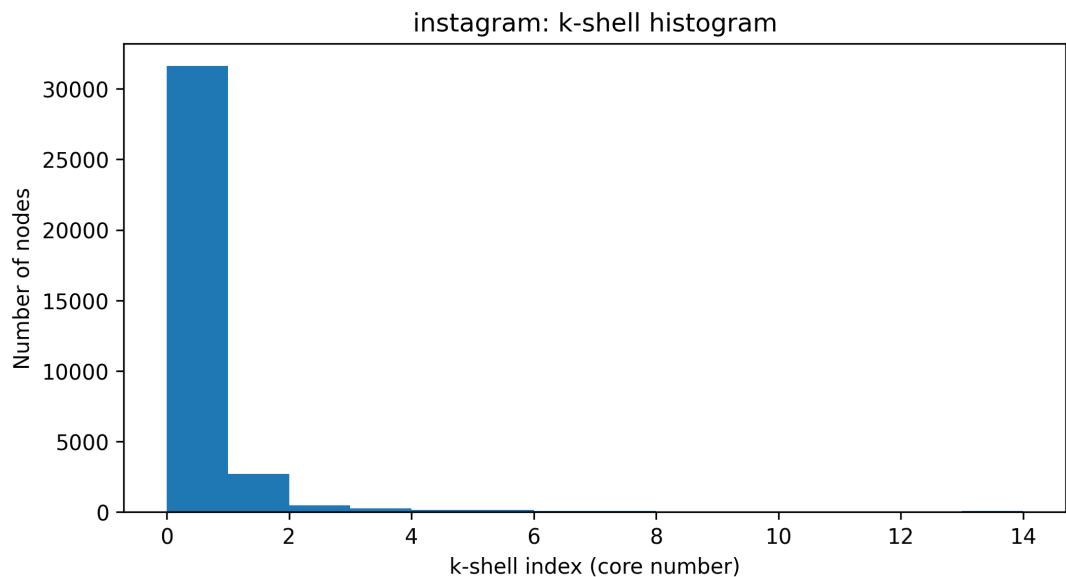
شکل ۳: توزیع k-shell در Telegram

● منحنی اندازه k-core (اندازه زیرگراف با شرط درجه $\leq k$) در Telegram.



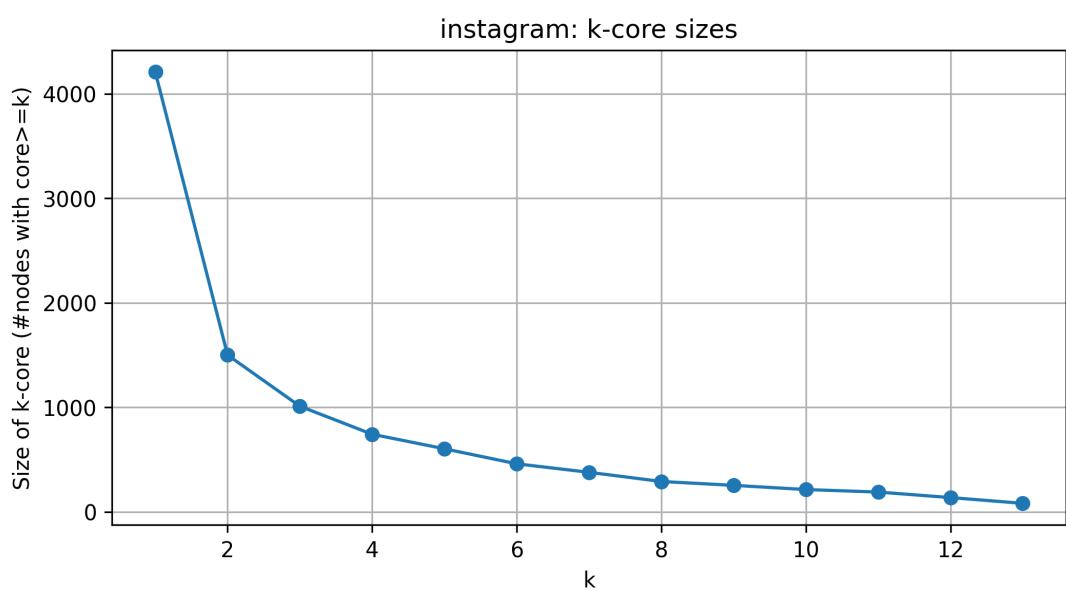
شکل ۴: اندازه k-core ها در Telegram

● توزیع فراوانی k-shell در Instagram



شکل ۵: توزیع k-shell در Instagram

● منحنی اندازه k-core (اندازه زیرگراف با شرط درجه $\leq k$) در Telegram.



شکل ۶: اندازه k-core‌ها در Instagram

تشخیص انجمن با Louvain بر روی شبکه بازنشر

۱.۴ هدف سؤال ۴ و تعریف دقیق گراف بازنشر

هدف سؤال 4 تقسیم شبکه بازنشر هر پلتفرم به انجمن‌هایی است که اعضای آن‌ها بیش از سایر گره‌ها با یکدیگر تعامل بازنشر دارند. سپس:

- انجمن‌های بزرگ (حداقل 5% از کل گراف) شناسایی می‌شوند،
- در هر انجمن بزرگ، اکانت‌های اثرگذار با استفاده از معیارهای شبکه‌ای معرفی می‌شوند،
- و در صورت امکان، تحلیل محتوایی یا رویکردهای یادگیری ماشین بر روی متن‌های مرتبط با همان انجمن انجام می‌شود.

در گراف بازنشر:

- هر گره نمایانگر یک account است.
- هر یال بیانگر وقوع بازنشر یا ارجاع بین دو گره است.
- وزن یال با ویژگی weight برابر با تعداد دفعات بازنشر بین دو گره در نظر گرفته شده است.

خروجی‌های GEXF شامل گراف‌های نهایی از نوع networkx.Graph و بدون جهت هستند؛ بدین معنا که جهت بازنشر در مرحله انجمن‌سازی حذف شده و تعامل‌ها به صورت رابطه‌ای وزن‌دار و دو طرفه مدل‌سازی شده‌اند. در فایل‌های louvain.gexf *_with_louvain.cid برای هر گره صفت louvain.cid ذخیره شده است که شناسه انجمن آن گره را مشخص می‌کند.

۲.۴ الگوریتم Louvain و معیار Modularity (فرمول‌ها)

الگوریتم Louvain روشی ابتکاری برای بیشینه‌سازی معیار modularity است. ایده اصلی modularity بررسی این موضوع است که آیا مجموع وزن یال‌های درون‌انجمنی بیش از مقدار مورد انتظار در یک مدل تصادفی با حفظ درجه گره‌ها است یا خیر.

برای یک گراف بدون جهت وزندار با پارامتر تفکیک‌پذیری γ (معادل resolution)، یکی از فرم‌های رایج modularity به صورت زیر تعریف می‌شود:

$$Q = \frac{1}{2m} \sum_{i,j} \left(A_{ij} - \gamma \frac{k_i k_j}{2m} \right) \delta(c_i, c_j),$$

که در آن:

- A_{ij} وزن یال بین i و j است،
- $k_i = \sum_j A_{ij}$ درجه وزن‌دار گره i است،
- $m = \frac{1}{2} \sum_{i,j} A_{ij}$ مجموع کل وزن یال‌ها است،
- c_i شناسه انجمن گره i است،
- $\delta(c_i, c_j)$ در صورت همان‌جامن بودن i و j برابر 1 و در غیر این صورت برابر 0 است.

اثر پارامتر resolution پارامتر γ تعیین می‌کند که انجمن‌ها در چه مقیاسی استخراج شوند:

- γ بزرگ‌تر \Rightarrow انجمن‌های کوچک‌تر و تعداد انجمن‌های بیشتر،
- γ کوچک‌تر \Rightarrow انجمن‌های بزرگ‌تر و تعداد انجمن‌های کمتر.

۳.۴ انتخاب resolution و فیلتر انجمن‌های بزرگ

- برای Twitter و Telegram مقدار $\text{resolution}=1.0$ استفاده شده است.
- برای Instagram مقدار $\text{resolution}=0.2$ در نظر گرفته شده است.

حداکثر اندازه یک انجمن بزرگ برای هر پلتفرم مطابق رابطه زیر تعیین شده است:

$$\min_{\text{size}} = \lceil 0.05 N \rceil,$$

که در آن N تعداد کل گره‌های گراف همان پلتفرم است.

۴.۴ آمار کلی شبکه‌های خروجی

در جدول ۳۳ آمار کلی هر گراف بر اساس فایل‌های `*_with_louvain.gexf` گزارش شده است. در ستون modularity مقدار Q بر اساس برجسب‌های `louvain_cid` محاسبه شده و مقدار `resolution` مطابق تنظیمات به کاررفته در هر پلتفرم درج شده است.

پلتفرم	# گره	# یال	چگالی	میانگین درجه	ایزوله #	مؤلفه #	resolution	modularity
Twitter	85,021	150,896	4.1750e-05	3.5496	42,142	42,693	1.0	0.4955
Telegram	48,632	73,798	6.2408e-05	3.0350	24,824	26,332	1.0	0.6595
Instagram	35,825	9,149	1.4257e-05	0.5108	31,617	32,042	0.2	0.8501

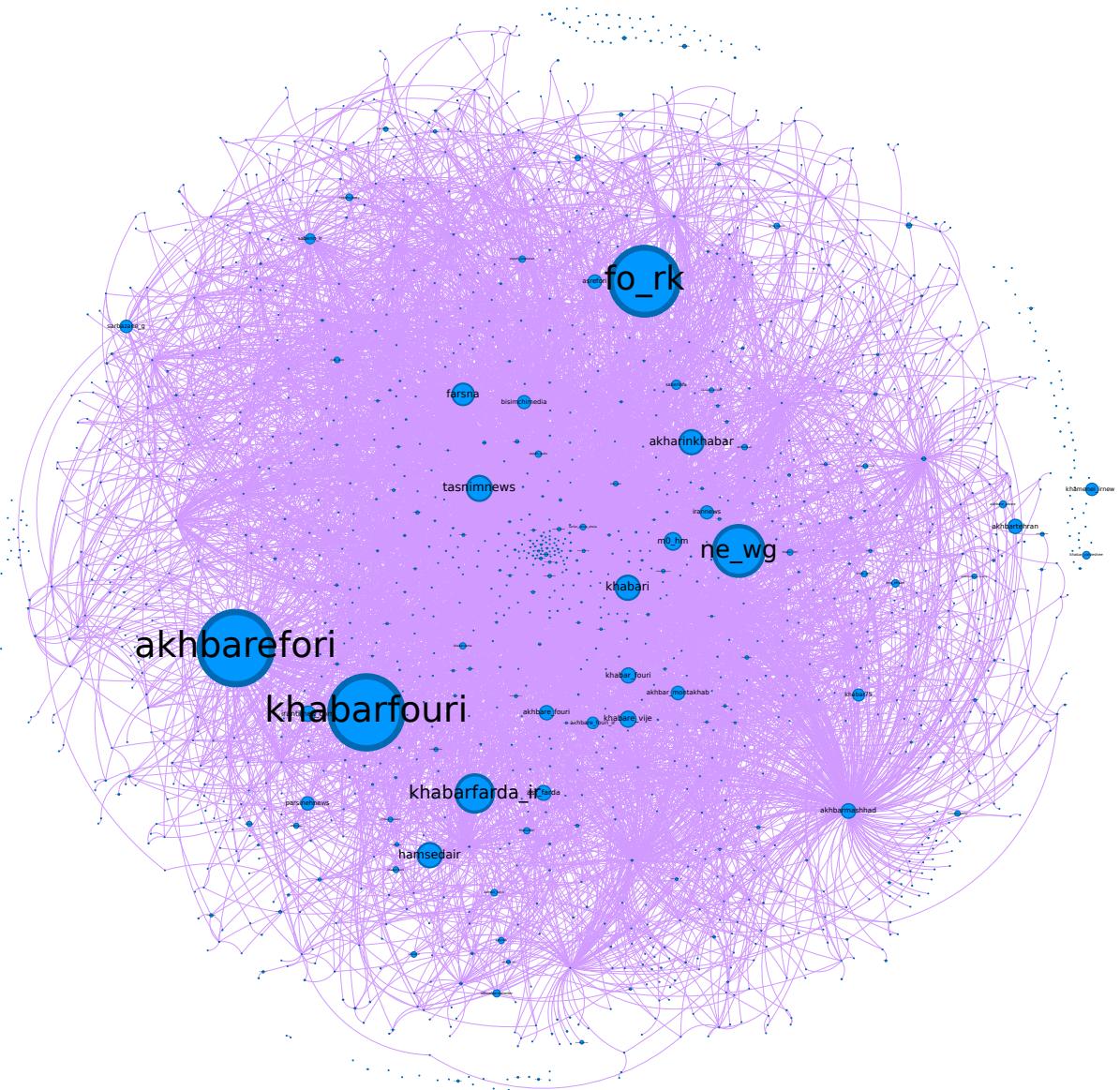
جدول ۳۳: آمار کلی گراف‌های خروجی.

نکته تفسیری در هر سه پلتفرم، تعداد گره‌های ایزوله (با درجه صفر) قابل توجه است. این موضوع سبب افزایش تعداد انجمان‌های استخراج شده با Louvain می‌شود، زیرا هر گره ایزوله معمولاً یک انجمان تک‌گرهی تشکیل می‌دهد. از این‌رو، اعمال شرط 5% برای تمرکز بر خوش‌های بزرگ و معنادار ضروری است.

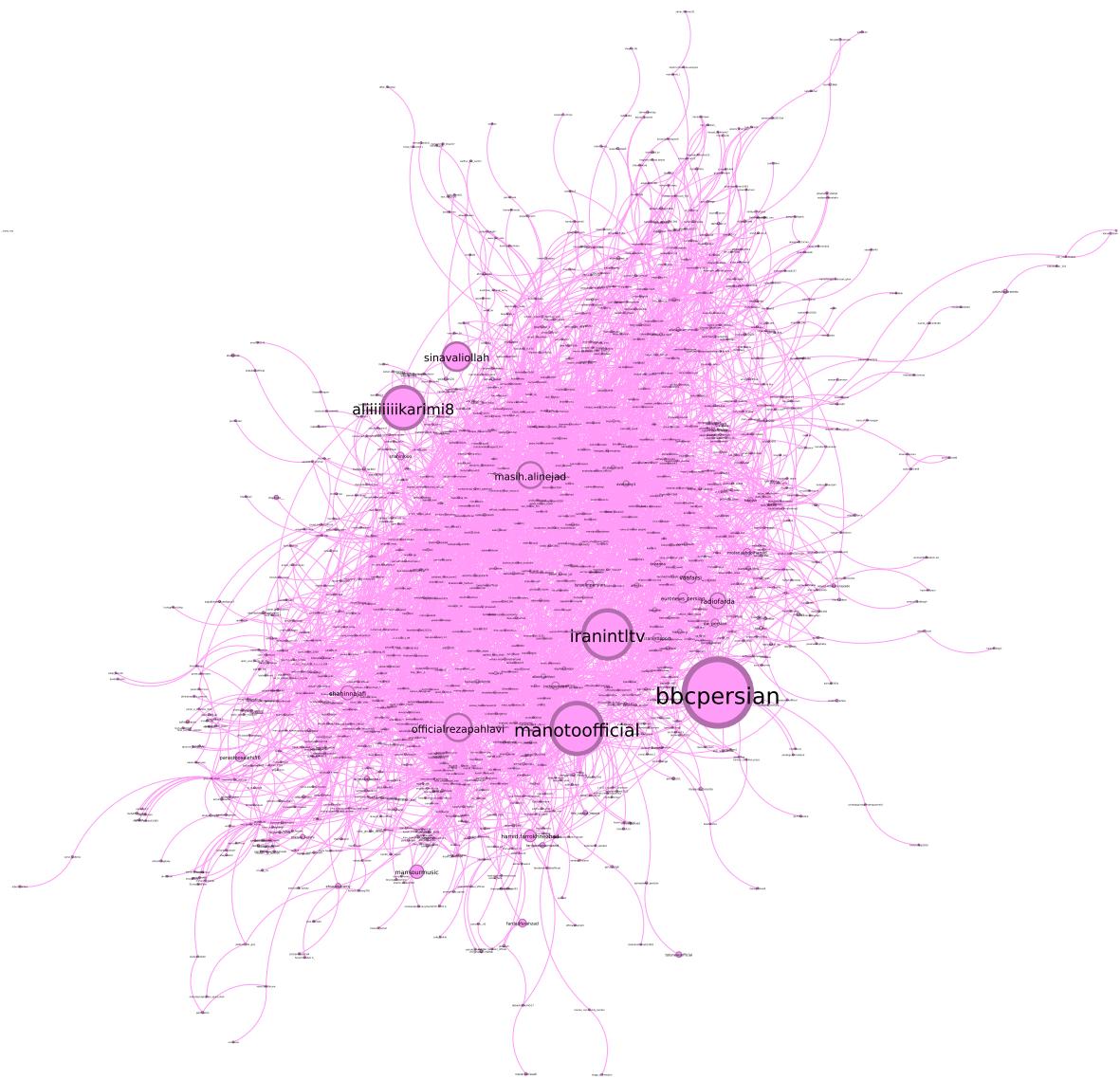
۵.۴ نمای کلی گراف‌ها با شرط 5%



شکل ۷: نمای کلی گراف Twitter با رنگ‌بندی انجمان‌های Louvain



شکل ۸: نمای کلی گراف Telegram با رنگبندی انجمن‌های Louvain



شکل ۹: نمای کلی گراف Instagram با رنگبندی انجمن‌های Louvain

۶.۴ پلتفرم Twitter: انجمن‌های بزرگ

حداقل اندازه یک انجمن بزرگ در Twitter با $N = 85021$ به صورت زیر محاسبه می‌شود:

$$\min_size = \lceil 0.05 \times 85021 \rceil = \lceil 4251.05 \rceil = 4252.$$

بنابراین هر انجمنی که حداقل 4252 عضو داشته باشد، «بزرگ» تلقی می‌شود.
براساس فایل q4_outputs/twitter_big_communities_summary.csv، دقیقاً 4 انجمن بزرگ شناسایی شده است:

community_id	اندازه	سهم از کل گراف
0	6,570	7.7275%
1	6,485	7.6275%
2	6,187	7.2770%
3	4,749	5.5857%

جدول ۳۴: انجمن‌های بزرگ در Twitter.

۷.۴ تحلیل انجمن ۰ در Twitter

در این بخش، انجمن $community_id=0$ به عنوان بزرگ‌ترین انجمن Twitter مورد تحلیل قرار می‌گیرد.

۱.۷.۴ زیرگراف انجمن و شاخص‌های ساختاری (بر اساس GEXF)

فایل زیرگراف این انجمن در خروجی موجود است:
`.q4_outputs/twitter_louvain_id_0.gexf`
 این زیرگراف دقیقاً 6570 گره (مطابق جدول ۳۴) و 15060 یال درون‌انجمن دارد.

چگالی درون‌انجمن (بدون وزن) اگر n تعداد گره‌ها و m_{in} تعداد یال‌های درون‌انجمن باشد:

$$\text{Density} = \frac{2m_{\text{in}}}{n(n - 1)}.$$

برای انجمن ۰:

$$n = 6570, \quad m_{\text{in}} = 15060 \quad \Rightarrow \quad \text{Density} = 0.0006978954.$$

مجموع وزن یال‌های درون‌انجمن اگر w_e وزن یال e باشد:

$$W_{\text{in}} = \sum_{e \in E(S)} w_e.$$

برای این انجمن:

$$W_{\text{in}} = 20894.$$

برای سنجش میزان اتصال این انجمن به بیرون، از معیار **conductance** استفاده می‌شود:

$$\phi(S) = \frac{\text{cut}(S, \bar{S})}{\text{vol}(S)}.$$

در این رابطه:

- $\text{cut}(S, \bar{S})$ مجموع وزن یال‌هایی است که یک سر آن‌ها درون انجمن و سر دیگر آن‌ها بیرون انجمن قرار دارد.

- $\text{vol}(S)$ مجموع درجه وزندار همه گره‌های درون انجمن است (یعنی مجموع وزن یال‌های incident به اعضای انجمن).

برای انجمن 0:

$$\text{vol}(S) = 70188, \quad \text{cut}(S, \bar{S}) = 28400, \quad \phi(S) = 0.4046.$$

تفسیر: مقدار 0.4046 نشان می‌دهد این انجمن، با وجود برخورداری از ساختار درونی، اتصال قابل توجهی به بیرون نیز دارد و کاملاً ایزوله نیست.

۲.۷.۴ آمار محتوایی و احساسات در انجمن 0 (بر اساس صفات گره در GEXF)

در فایل GEXF برای هر گره، صفات `posts`, `views`, `positive`, `negative` و `neutral` ذخیره شده است. با تجمعی این صفات بر روی اعضای انجمن 0، نتایج زیر به دست می‌آید:

- مجموع `posts`: 70,570
- مجموع `views`: 30,058,617
- مجموع پست‌های `positive` (سهم: 18.21%) 12,853
- مجموع پست‌های `negative` (سهم: 80.37%) 56,718
- مجموع پست‌های `neutral` (سهم: 1.42%) 999

۳.۷.۴ توزیع followers در انجمن ۰

برای توصیف ناهمگنی اندازه مخاطب اعضای انجمن ۰، شاخص‌های توصیفی زیر گزارش می‌شود:

- میانگین 3635.48 :followers

- میانه 118 :followers

- صدک 2245.2 :followers 90

- بیشینه 2,164,668 :followers

تفسیر: اختلاف چشمگیر میانگین و میانه نشان می‌دهد توزیع followers راست‌چوله است؛ به این معنا که تعداد اندکی اکانت بزرگ مقدار میانگین را افزایش می‌دهند، در حالی که بخش عمده‌ای از اکانت‌ها دنبال‌کنندگان کمتری دارند.

۴.۷.۴ اکانت‌های اثرگذار

فایل q4_outputs/twitter_community_top_accounts.csv برای انجمن‌های بزرگ، دو رتبه‌بندی زیر را ذخیره کرده است:

- top10_weighted_degree

- top10_pagerank

این شاخص‌های مرکزیت بر روی زیرگراف همان انجمن محاسبه شده‌اند (نه بر روی کل گراف).

#	اکانت	followers	weighted_degree
1	iranint1 (728252911675944961)	2,164,668	2,657
2	indypersian (1065658304884547585)	336,492	1,934
3	manotonews (410718052)	1,311,254	1,255
4	kayhanlondon (2377612094)	257,885	1,022
5	iranintlsport (1041674345926389760)	314,439	983
6	alighazizade (2177326935)	634,595	766
7	haghiae (88389777)	54,701	763
8	noorpahlavi20 (1639733491611959296)	15,729	664
9	vahid__bahman (2572902906)	80,076	595
10	rppress0 (1249443053988769792)	28,332	488

جدول ۳۵: اکانت‌های برتر انجمن ۰ در Twitter بر اساس Weighted Degree

#	اکانت	followers	pagerank
1	iranintl (728252911675944961)	2,164,668	0.051518424
2	indypersian (1065658304884547585)	336,492	0.034844253
3	noorpahlavi20 (1639733491611959296)	15,729	0.024956533
4	manotonews (410718052)	1,311,254	0.023626526
5	alighazizade (2177326935)	634,595	0.019890882
6	iranintlsport (1041674345926389760)	314,439	0.018828555
7	kayhanlondon (2377612094)	257,885	0.018316745
8	haghaie (88389777)	54,701	0.016388977
9	vahid__bahman (2572902906)	80,076	0.013142799
10	rppress0 (1249443053988769792)	28,332	0.013062132

جدول ۳۶: اکانت‌های برتر انجمن ۰ در Twitter بر اساس PageRank وزن دار.

تفسیر اختلاف Weighted Degree و PageRank در یک انجمن معيار Weighted Degree بيش از همه بيانگر «تعامل مستقيم» است (تعداد و وزن ارتباط‌های مستقيم با ساير اعضا). در مقابل، PageRank علاوه بر تعداد و وزن ارتباط‌ها، به «اعتبار همسایه‌ها» نيز حساس است؛ بدین معنا که دريافت بازنشر از گره‌های اثرگذارتر می‌تواند مقدار PageRank را افزایش دهد. از اين‌رو ممکن است اکانتی در رتبه‌بندی PageRank بالاتر از اکانتی قرار گيرد که درجه وزن دار بيشتری دارد، زيرا از مسیرهای مهم‌تر انتشار تغذيه می‌کند.

۵.۷.۴ تحليل محتوايی با TF-IDF و NMF

بر اساس فایل q4_outputs/ml_topics.json، برای Twitter و انجمن ۰ تعداد اسناد به کاررفته برابر 2306 بوده و ۳ موضوع استخراج شده است. ۱۵ عبارت برتر هر موضوع به شرح زير است:

- **موضوع ۱:** در، به، از، ايران، https، co، co، https، اين، می، دی، مردم، است، را

- **موضوع ۲:** اي سرکردگان، سرکردگان فاسد، تقاص، جنایت را، سرکردگان، که خامنه، اسلامی تقاص، قاتل جمهوري، هم درد، مردم رنج، درد همراه، دوست ملکشاهی، کشیده میهن، همراه تان، تان هستم

- **موضوع ۳:** تسخیر، قدر فرصت، کلان شهرها، انصورت، انصورت Trump، شليک ميشود، امده تسخیر، شهرها تهران، انگاه رژيم، يا مجبور، شيراز تبريز، کنيم با، زمين را، شاهزاده انگاه، کند قدر

نکته تفسیری درباره موضوعها در موضوع 1 چند واژه بسیار عمومی (مانند «در»، «به»، «از» و «که») مشاهده می شود که بهدلیل نقش توقف واژه، بعنهایی محتوای متمایزی را مشخص نمی کند. در مقابل، موضوعهای 2 و 3 شامل واژه ها و عبارت های اختصاصی تری هستند و برای برچسب گذاری محتوایی انجمن 0 ظرفیت تبیینی بالاتری دارند.

۶.۷.۴ جمع بندی تحلیلی انجمن ۰

با تجمعی سه دسته خروجی ساختار شبکه ای، شاخص های مرکزیت و موضوع های استخراج شده، نتایج زیر درباره انجمن ۰ قابل بیان است:

۱. از نظر ساختار انتشار: با وجود چگالی پایین (که در شبکه های اجتماعی پدیده ای رایج است)، تعداد یال های درون انجمن (15060) و مجموع وزن درون انجمن (20894) نشان می دهد این انجمن دارای یک «هسته بازنیش» معنادار است.
۲. از نظر جدا افتادگی: مقدار conductance برابر ۰.۴۰۴۶ است؛ به این معنا که انجمن، ضمن برخورداری از ساختار درونی، اتصال بیرونی قابل توجهی نیز دارد و بنابراین می توان آن را خوش های بزرگ محتوایی / رسانه ای دانست که با سایر خوش ها نیز تعامل دارد.
۳. از نظر نقش اکانت ها: رتبه بندی ها نشان می دهد چند اکانت هم زمان از نظر درجه وزن دار و نیز PageRank نقش محوری دارند (برای مثال iranint و indypersian). چنین هم پوشانی ای معمولاً نشانه وجود «هاب های انتشار» درون خوش است.
۴. از نظر محتوایی: خروجی NMF نشان می دهد در کنار واژه های عمومی، دو موضوع دیگر دارای عبارت های مشخص تر هستند؛ از این رو الگوی بازنیش در این انجمن حول چند محور محتوایی نسبتاً متمایز متمرکز شده است.

۸.۴ انجمن ۱ در Twitter

۱.۸.۴ مشخصات زیر گراف و سازگاری ساختاری

زیر گراف این انجمن در فایل زیر موجود است:

q4_outputs/twitter_louvain_id_1.gexf.

این زیر گراف دارای ویژگی های زیر است:

● تعداد گره: 6485

- تعداد یال درون انجمن: 16619
- تعداد مؤلفه همبند: 1 (یعنی زیرگراف کاملاً همبند است)
- تعداد گره ایزوله داخلی: 0 (یعنی هیچ عضوی وجود ندارد که صرفاً به بیرون انجمن یال داشته باشد)

چگالی و میانگین درجه $m_{in} = 16619$ و $n = 6485$ با:

$$\text{Density} = 0.0007904636, \quad \overline{\deg} = \frac{2m_{in}}{n} = 5.1253662.$$

این مقدار میانگین درجه نسبت به بخش بزرگی از کل گراف (با وجود گره‌های ایزوله فراوان) بزرگ‌تر است؛ بنابراین انجمن‌های بزرگ عموماً «هسته ارتباطی» پرنگ‌تری دارند.

مجموع وزن‌های درون انجمن

$$W_{in} = 21052.$$

در نتیجه، میانگین درجه وزن‌دار داخلی برابر است با:

$$\overline{\deg}_w = \frac{2W_{in}}{n} = 6.4925212.$$

آمار توزیع وزن یال‌ها درون انجمن با استخراج همه مقدارهای weight برای یال‌های درون انجمن:

- میانگین وزن یال: 1.2667429
- میانه وزن یال: 1
- صدک 90 وزن یال: 2
- بیشینه وزن یال: 105

تفسیر: میانه 1 بیانگر آن است که بخش بزرگی از روابط از نوع «یک‌بار بازنشر» است؛ با این حال، وجود بیشینه 105 نشان می‌دهد چند جفت اکانت دارای رابطه بازنشر بسیار تکرارشونده هستند و می‌توانند مسیرهای اصلی انتشار در این انجمن را شکل دهند.

اتصال به بیرون: Conductance و Cut ، Volume
بر اساس گراف کامل :twitter_with_louvain.gexf

$$\text{vol}(S) = 75286, \quad \text{cut}(S, \bar{S}) = 33182, \quad \phi(S) = 0.44074596.$$

همچنین سهم داخلی و خارجی از حجم به صورت زیر است:

$$\frac{2W_{\text{in}}}{\text{vol}} = 0.5592540, \quad \frac{\text{cut}}{\text{vol}} = 0.4407460.$$

تفسیر: حدود 55.9% از وزن incident اعضا مربوط به یالهای داخلی و حدود 44.1% مربوط به ارتباط با بیرون است؛ بنابراین انجمن 1 نسبتاً «نفوذپذیر» است و مرز آن کاملاً بسته نیست.

خوشبندی (سه تایی ها)

$$0.0852993 = \text{Average Clustering} \bullet$$

$$0.00440085 = \text{Transitivity} \bullet$$

تفسیر: مقدار میانگین خوشبندی نشان می‌دهد برخی گره‌ها در همسایگی خود سه تایی‌هایی تشکیل می‌دهند، اما مقدار کوچک transitivity بیانگر آن است که نسبت کلی مثلث‌ها به کل سه تایی‌ها در مقیاس انجمن بالا نیست؛ رفتاری که در شبکه‌های بازنشر (با ساختارهای ستاره‌ای/هاب محور) رایج است.

۲.۸.۴ آمار محتوایی و احساسات در انجمن 1

این آمار با تجمعی صفات گره‌ها در فایل GEXF به دست آمده است:

$$77065 : \text{posts} \bullet$$

$$16123605 : \text{views} \bullet$$

$$(16.0151\% : \text{سهم}) \quad 12342 : \text{positive} \bullet$$

$$(82.5757\% : \text{سهم}) \quad 63637 : \text{negative} \bullet$$

$$(1.4092\% : \text{سهم}) \quad 1086 : \text{neutral} \bullet$$

۳.۸.۴ توزیع followers در انجمن ۱

برای توصیف ناهمگنی اندازه مخاطب اعضای انجمن، شاخص‌های توصیفی زیر گزارش می‌شود:

● میانگین 1383.2435 :followers

● میانه 155 :followers

● صدک 1810.8 :followers 90

● بیشینه 1370401 :followers

تفسیر: اختلاف قابل توجه میانگین و میانه نشان می‌دهد توزیع followers راست‌چوله است؛ بدین معنا که اغلب اکانت‌ها تعداد دنبال‌کننده اندکی دارند، اما تعداد کمی اکانت بزرگ مقدار میانگین را افزایش می‌دهند.

۴.۸.۴ اکانت‌های اثرگذار در انجمن ۱

برای این انجمن، دو رتبه‌بندی بر اساس PageRank و Weighted Degree گزارش می‌شود.

#	اکانت	followers	weighted_degree
1	irannewsagency0 (1550545991778340865)	117,861	2,988
2	ario1401 (1613903983030026240)	41,479	1,728
3	tanhayektiger (797159363811147776)	24,960	1,663
4	scharllinn (1525051315801542659)	12,755	923
5	2002kingfox (1810765787847872513)	50,185	897
6	gohardaddy (1925636410557521920)	23,600	880
7	ayatolahjaghian (1525913373958975491)	52,341	743
8	agha_rashti2 (1676213344062021632)	29,679	626
9	king_fishpoet (1573142641046421504)	15,690	615
10	akhoondemorde (25033880)	43,106	588

. جدول ۳۷: اکانت‌های برتر انجمن ۱ در Twitter بر اساس Weighted Degree

اکانت‌های برتر بر اساس Weighted Degree

#	اکانت	followers	pagerank
1	irannewsagency0 (1550545991778340865)	117,861	0.0477693232
2	ario1401 (1613903983030026240)	41,479	0.0418763981
3	tanhayektiger (797159363811147776)	24,960	0.0403601907
4	gohardaddy (1925636410557521920)	23,600	0.0203439819
5	2002kingfox (1810765787847872513)	50,185	0.0173546519
6	scharllinn (1525051315801542659)	12,755	0.0167289688
7	ayatolahjaghian (1525913373958975491)	52,341	0.0158022557
8	agha_rashti2 (1676213344062021632)	29,679	0.0141020923
9	akhoondemorde (25033880)	43,106	0.0136230600
10	arixolin (49127938)	51,939	0.0119639592

جدول ۳۸: اکانت‌های برتر انجمان ۱ در Twitter بر اساس PageRank.

اکانت‌های برتر بر اساس PageRank

مقایسه دو رتبه‌بندی (همپوشانی و تفاوت) در انجمان ۱، میزان همپوشانی دو فهرست ۱۰ تایی برابر با ۹ است:

● فقط در king_fishpoet :Weighted Degree

● فقط در arixolin :PageRank

تفسیر: Weighted Degree عمدتاً «عامل مستقیم» را می‌سنجد (تعداد و وزن ارتباط‌های درون‌انجمان)، در حالی که PageRank علاوه بر آن به «اعتبار همسایه‌ها» نیز حساس است. بنابراین ممکن است اکانتی که اتصال مستقیم کمتری دارد، اما از مسیرهای اثرگذارتر بازنیش تغذیه می‌کند، در رتبه‌بندی PageRank وارد Top10 شود.

۵.۸.۴ تحلیل آماری رابطه followers با مرکزیت‌ها (محاسبه بر اساس زیرگراف انجمان)

با توجه به چولگی شدید توزیع followers، دو نوع همبستگی گزارش می‌شود:

● (همبستگی رتبه‌ای) برای سنجش رابطهٔ یکنواخت، بدون فرض خطی بودن، Spearman

● Pearson بین $\log(1 + \text{followers})$ و مرکزیت‌ها، به‌منظور کاهش اثر چولگی و نزدیک شدن به الگوی خطی.

1 در انجمان Spearman

$$\rho(\text{followers}, \text{weighted_degree}) = 0.1686,$$

$$\rho(\text{followers}, \text{pagerank}) = 0.1582,$$

$$\rho(\text{weighted_degree}, \text{pagerank}) = 0.9136.$$

تفسیر: ارتباط followers با مرکزیت‌ها در بازهٔ ضعیف تا متوسط قرار دارد؛ به عبارت دیگر، «بزرگی مخاطب» الزاماً به معنای «محوریت در بازنشر درون‌انجمان» نیست. در مقابل، رابطهٔ pagerank و weighted_degree پسیار قوی است که با توجه به اینکه هر دو شاخص نوعی «اهمیت شبکه‌ای» را بر اساس یال‌ها استخراج می‌کنند، نتیجه‌ای طبیعی محسوب می‌شود.

با تبدیل لگاریتمی در انجمان 1 Pearson

$$r(\log(1 + \text{followers}), \text{weighted_degree}) = 0.1452,$$

$$r(\log(1 + \text{followers}), \text{pagerank}) = 0.1515.$$

تفسیر: حتی پس از تبدیل لگاریتمی، رابطهٔ خطی بین followers و شاخص‌های مرکزیت همچنان ضعیف باقی می‌ماند؛ بنابراین ساختار بازنشر درون‌انجمان بیش از آنکه صرفاً تابع اندازهٔ مخاطب باشد، توسط الگوهای ارتباطی شکل می‌گیرد.

۶.۸.۴ موضوع‌یابی NMF برای انجمان 1

برای Twitter و $\text{community_id} = 1$

● تعداد اسناد استفاده شده: 2128

● تعداد موضوع‌ها: 3

15 عبارت/واژهٔ برتر هر موضوع

● موضوع 1: در، به، co، https، https_co، از، با، است، جاویدشاه، را، مردم، این، ایران، که، می

● موضوع 2: چه در، چه، فراخوان، دی، اولین فراخوان، ۱۹ دی، جمعه ۱۹، صادر خواهد، فراخوان عمومی، دی راس، دادن شعاردر، ان فراخوان، خانه چه، بعدی صادر، دی جمعه

- موضوع ۳: رضا پهلوی، پهلوی، رضا، شاهزاده رضا، شاهزاده، هستم، می دهم، به شما، که خامنه، اطمنان می، پهلوی مردم، ای سرکردگان، شما اطمینان، قاتل جمهوری، مردم رنج

تفسیر روش شناختی خروجی NMF

- موضوع ۱ شامل توقف واژه‌ها و توکن‌های عمومی (مانند «در»، «به»، «از»، «که» و «می») و نیز توکن‌های فنی مرتبط با پیوندها (مانند https و co) است. این موضوع معمولاً «پس زمینه زبانی/فنی» را جذب می‌کند و به تنایی دلالت محتوایی روشنی ندارد.
- موضوع‌های ۲ و ۳ ترکیب‌های اختصاصی‌تری دارند و برای برچسب‌گذاری محتوایی انجمن ظرفیت تبیینی بیشتری فراهم می‌کنند.

۹.۴ انجمن ۲ در Twitter

۱.۹.۴ مشخصات زیرگراف و سازگاری ساختاری

در این زیرگراف:

- تعداد گره: 6187
- تعداد یال درون‌انجمن: 14083
- تعداد مؤلفه همبند: 1 (زیرگراف همبند است)
- تعداد گره ایزوله داخلی: 0

چگالی و میانگین درجه با $m_{\text{in}} = 14083$ و $n = 6187$

$$\text{Density} = 0.0007359277, \quad \overline{\deg} = \frac{2m_{\text{in}}}{n} = 4.5524487.$$

مجموع وزن‌های درون‌انجمن

$$W_{\text{in}} = 17573 \Rightarrow \overline{\deg}_w = \frac{2W_{\text{in}}}{n} = 5.6806207.$$

آمار توزیع وزن یال‌ها درون انجمن

- میانگین وزن یال: 1.2478165
- میانه وزن یال: 1
- صدک 90 وزن یال: 2
- بیشینه وزن یال: 97

اتصال به بیرون: Conductance، Cut و Volume بر اساس گراف کامل:

$$\text{vol}(S) = 49867, \quad \text{cut}(S, \bar{S}) = 14721, \quad \phi(S) = 0.29520525.$$

سهم داخلی و خارجی از حجم:

$$\frac{2W_{\text{in}}}{\text{vol}} = 0.7047948, \quad \frac{\text{cut}}{\text{vol}} = 0.2952053.$$

تفسیر: این انجمن نسبت به انجمن 1 «بسته‌تر» است (حدود 70.5% از وزن incident داخلی و حدود 29.5% مربوط به ارتباط با بیرون است); بنابراین مرزهای انجمن 2 از منظر شبکه‌ای روشن‌تر و جدآفتداده‌تر به نظر می‌رسد.

خوشه‌بندی

$$0.0212104 = \text{Average Clustering} \bullet$$

$$0.00781761 = \text{Transitivity} \bullet$$

تفسیر: میانگین خوشه‌بندی این انجمن از انجمن 1 کوچک‌تر است، اما transitivity اندکی بزرگ‌تر گزارش شده است. این وضعیت می‌تواند زمانی رخ دهد که مثلث‌ها بیشتر پیرامون بخشی از شبکه متمرکز باشند و به صورت یکنواخت در کل گره‌ها پخش نشده باشند.

۲.۹.۴ آمار محتوایی و احساسات در انجمن 2

$$62262 : \text{posts} \bullet$$

$$31850564 : \text{views} \bullet$$

(سهم: 8447 :positive ● 13.5669%)

(سهم: 52015 :negative ● 83.5421%)

(سهم: 1800 :neutral ● 2.8910%)

۳.۹.۴ توزیع followers در انجمن ۲

● میانگین 3538.2500 :followers

● میانه 171 :followers

● صدک 2510.8 :followers

● بیشینه 2383067 :followers

تفسیر: در مقایسه با انجمن ۱، هم میانگین و هم بیشینه followers بزرگ‌تر است؛ بنابراین حضور تعداد بیشتری اکانت بزرگ (یا چند اکانت بسیار بزرگ) محتمل است. با این حال، بزرگ بودن مخاطب‌الزاماً به معنای مرکزیت در بازنثر درون‌انجمن نیست.

۴.۹.۴ اکانت‌های اثرگذار در انجمن ۲

#	اکانت	followers	weighted_degree
1	aamirhoossein (1408558266)	12,307	1,096
2	arash_cosmo (52033979)	2,753	754
3	iranwire (1378970222)	142,744	550
4	mamadporii (334161563)	423,110	520
5	corrrdeliiiaaa (1316122268020420609)	1,473	445
6	soranmansournia (1455566016013012998)	18,550	440
7	tavaana (142772512)	154,355	421
8	nastaran_mn (1472237454551887880)	27	319
9	elhamomidvari (1955124895)	13,027	252
10	maroofian_n (1612794343911297025)	26,332	235

جدول ۳۹: اکانت‌های برتر انجمن ۲ در Twitter بر اساس Weighted Degree

اکانت‌های برتر بر اساس Weighted Degree

#	اکانت	followers	pagerank
1	aamirhoossein (1408558266)	12,307	0.0262593829
2	arash_cosmo (52033979)	2,753	0.0254010909
3	iranwire (1378970222)	142,744	0.0165159856
4	mamadporii (334161563)	423,110	0.0110053960
5	corrddeliiiaaa (1316122268020420609)	1,473	0.0102057938
6	tavaana (142772512)	154,355	0.0099021861
7	hamid_kp_ (1177620173731323906)	1,933	0.0098826619
8	soranmansournia (1455566016013012998)	18,550	0.0091180264
9	maroofian_n (1612794343911297025)	26,332	0.0067394186
10	mahsa_piraei (1591396206797881344)	8,054	0.0065876491

جدول ۴۰: اکانت‌های برتر انجمن ۲ در Twitter بر اساس PageRank.

اکانت‌های برتر بر اساس PageRank

مقایسه دو رتبه‌بندی (همپوشانی و تفاوت) در انجمن ۲، میزان همپوشانی دو فهرست برابر با ۸ است:

• فقط در nastaran_mn و elhamomidvari :Weighted Degree

• فقط در mahsa_piraei و hamid_kp_ :PageRank

تفسیر: اختلاف بین دو فهرست Top10 در این انجمن کمی بیشتر از انجمن ۱ است (همپوشانی ۸ بهجای ۹). این وضعیت می‌تواند نشان‌دهنده آن باشد که ساختار نفوذ در این انجمن به مسیرهای چندمرحله‌ای انتشار (که PageRank بهتر بازتاب می‌دهد) حساس‌تر است.

۵.۹.۴ تحلیل آماری رابطه followers با مرکزیت‌ها (محاسبه بر اساس زیرگراف انجمن)

در انجمن ۲ Spearman

$$\rho(\text{followers}, \text{weighted_degree}) = 0.2108,$$

$$\rho(\text{followers}, \text{pagerank}) = 0.2025,$$

$$\rho(\text{weighted_degree}, \text{pagerank}) = 0.8907.$$

تفسیر: در این انجمن، رابطه رتبه‌ای followers با مرکزیت‌ها اندکی قوی‌تر از انجمن ۱ است، اما

همچنان در سطح «قوی» قرار نمی‌گیرد. همچنین همبستگی بین pagerank و weighted_degree بین ۰.۱۹۴۰ و ۰.۱۸۸۵ است. همچنان بالا گزارش شده است.

Pearson با تبدیل لگاریتمی در انجمان ۲

$$r(\log(1 + \text{followers}), \text{weighted_degree}) = 0.1940,$$

$$r(\log(1 + \text{followers}), \text{pagerank}) = 0.1885.$$

تفسیر: پس از تبدیل لگاریتمی، شدت رابطه خطی افزایش می‌یابد، اما همچنان در حد متوسط رو به پایین باقی می‌ماند؛ بنابراین تحلیل «اثرگذاری شبکه‌ای» لازم است مستقل از «اندازه مخاطب» نیز انجام شود.

۶.۹.۴ خروجی موضوع‌یابی NMF برای انجمان ۲

برای Twitter و $\text{community_id} = 2$

● تعداد اسناد: ۳۴۹۷

● تعداد موضوع‌ها: ۳

۱۵ عبارت/واژه برتر هر موضوع

● موضوع ۱: در، از، به، که، می، است، این، را، ایران، با، های، مردم، اعتراضات، کوردستان،
ها

● موضوع ۲: https://co، اعتراضات_سراسری، اعتراضات_سراسری، https://co، دی،
دی، ۱۴۰۴، زن_زندگی_ازادی، ۱۴۰۴، اعتراضات_سراسری، در، زن_زندگی_ازادی
اعتصابات_سراسری، تهران، ۱۶، https://

● موضوع ۳: زندگی، زن، زندگی_ازادی، ازادی، زن_زندگی، ازادی_اعتصابات_سراسری،
زن، رو، اعتراضات_سراسری_اعتصابات_و_اعتصابات_سراسری، iranrevolution2026،
اعتصابات_و_اعتصابات_سراسری_iranrevolution2026، اعتراضات_iranrevolution2026،
اعتصابات_سراسری_iranrevolution2026، اعتصابات_سراسری، تا، زن_زندگی_ازادی
iranrevolution2026 https://

تفسیر روش شناختی

- موضوع 1 (همانند بسیاری از انجمن‌ها) ترکیبی از واژه‌های عمومی و چند واژه پر تکرار است و لزوماً به تهایی «برچسب معنایی دقیق» فراهم نمی‌کند.
- موضوع‌های 2 و 3 به دلیل حضور توکن‌های خاص‌تر (هشتگ‌ها و عبارت‌های متمایز) برای توصیف محتوایی این انجمن مناسب‌تر هستند.

۱۰.۴ انجمن 3 در Twitter

۱.۱۰.۴ مشخصات زیرگراف و سازگاری ساختاری

در این زیرگراف:

- تعداد گره: 4749
- تعداد یال‌داخلی (بدون جهت): 18448
- تعداد مؤلفه همبند: 1 (زیرگراف به طور کامل همبند است)
- تعداد گره ایزوله داخلی: 0

۲.۱۰.۴ چگالی، میانگین درجه و شدت تعامل داخلی

چگالی (بدون وزن) برای این انجمن:

$$n = 4749, \quad m_{\text{in}} = 18448 \quad \Rightarrow \quad \text{Density} = 0.0016363131.$$

تفسیر: چگالی مطلق همچنان کوچک است (که در شبکه‌های بزرگ پدیده‌ای طبیعی است)، اما در مقایسه با سه انجمن بزرگ دیگر Twitter، این انجمن متراکم‌تر است (مطابق جدول مقایسه انتهای همین بخش).

میانگین درجه داخلی (بدون وزن)

$$\overline{\deg} = \frac{2m_{\text{in}}}{n} = \frac{2 \times 18448}{4749} = 7.7692146.$$

تفسیر: هر عضو به طور متوسط حدود 7.77 ارتباط بازنیش داخلی دارد که نسبت به انجمن‌های 0، 1 و 2 مقدار بزرگ‌تری است.

مجموع وزن داخلی و میانگین درجه وزندار داخلی

$$W_{\text{in}} = \sum_{e \in E(S)} w_e = 23986.$$

بنابراین:

$$\overline{\deg}_w = \frac{2W_{\text{in}}}{n} = \frac{2 \times 23986}{4749} = 10.1014951.$$

۳.۱۰.۴ توزیع وزن یال‌ها درون انجمن و تفسیر آن

در شبکه بازنشر، وزن یال (weight) بیانگر تعداد تکرار بازنشر بین دو اکانت است؛ از این‌رو، توزیع وزن‌ها نشان می‌دهد تعامل‌ها تا چه حد «یکباره» یا «تکرارشونده و پایدار» هستند. برای یال‌های داخلی انجمن ۳:

- میانگین وزن یال: ۱.۳۰۰۱۹۵۱
- میانه وزن یال: ۱
- صدک ۹۰ وزن یال: ۲
- بیشینه وزن یال: ۶۳

تفسیر میانه ۱ نشان می‌دهد بخش بزرگی از روابط بازنشر تنها یکبار ثبت شده‌اند؛ با این حال، وجود بیشینه ۶۳ بیانگر آن است که برخی زوج اکانت‌ها به صورت تکرارشونده از یکدیگر بازنشر می‌کنند. این «پیوندهای قوی» می‌توانند ستون فقرات انتشار محتوا در انجمن را شکل دهند.

۴.۱۰.۴ اتصال به بیرون و جداول‌گذگی انجمن با Conductance

بر اساس گراف کامل :twitter_with_louvain.gexf

$$\text{vol}(S) = 80965, \quad \text{cut}(S, \bar{S}) = 32993.$$

از آنجا که $\text{vol}(S) < \text{vol}(\bar{S})$ است و داریم:

$$\phi(S) = \frac{32993}{80965} = 0.4074971.$$

بررسی سازگاری حجم

$$2W_{\text{in}} + \text{cut} = 2 \times 23986 + 32993 = 80965 = \text{vol}(S).$$

این برابری نشان می‌دهد محاسبه وزن‌های داخلی و برش بیرونی به صورت سازگار انجام شده است.

سهم تعامل داخلی نسبت به تعامل کل

$$\frac{2W_{\text{in}}}{\text{vol}(S)} = \frac{47972}{80965} = 0.5925029.$$

یعنی حدود 59.25% از وزن تعاملات incident اعضای انجمن در داخل خود انجمن رخ می‌دهد و حدود 40.75% مربوط به تعامل با بیرون است.

تفسیر این نتایج نشان می‌دهد انجمن 3 با وجود تراکم داخلی بالا، همچنان ارتباط بیرونی قابل توجهی دارد؛ بنابراین می‌توان آن را «هسته‌ای پرتعامل» دانست که کاملاً منزوی نیست و با سایر خوشه‌ها نیز تعامل دارد.

۵.۱۰.۴ مثلفها و خوشه‌بندی: Transitivity و Average Clustering

برای بررسی اینکه ساختار شبکه بیشتر «هاب‌محور» است یا «گروه‌های سه‌تایی/بسته» را بازتاب می‌دهد، دو شاخص مکمل زیر گزارش می‌شود:

$$0.0412579 = \text{Average Clustering} \bullet$$

$$0.0118184 = \text{Transitivity} \bullet$$

تفسیر Average Clustering نشان می‌دهد برای بخشی از گره‌ها، همسایه‌ها تا حدی با یکدیگر در ارتباط هستند و مثلفهایی تشکیل می‌شود؛ با این حال، مقدار کوچک transitivity بسیاری از شبکه‌های بازنیش (یا بستانگر آن است که شبکه در مقیاس کلان بیشتر با الگوهای ستاره‌ای/هاب‌محور شکل گرفته است تا گروه‌های کوچک کاملاً بسته).

۶.۱۰.۴ آمار posts و احساسات (neutral/negative/positive views)

در فایل‌های GEXF برای هر گره، صفات neutral، negative، positive، views، posts ذخیره شده است. با تجمعی این صفات بر روی اعضای انجمن 3:

- مجموع posts: 75751

- مجموع views: 36848356

تقسیم احساسات:

- (سهم: 19.60%) 14850 :positive ●

- (سهم: 78.79%) 59687 :negative ●

- (سهم: 1.60%) 1214 :neutral ●

۷.۱۰.۴ توزیع followers و ناهمگنی اندازه مخاطب در انجمن

برای سنجش ناهمگنی اندازه مخاطب اعضای انجمن، آمارهای زیر از صفت followers استخراج شده است:

- میانگین followers: 2730.90

- میانه followers: 149

- صدک followers: 90

- بیشینه followers: 1933855

تفسیر کوچکتر بودن چشمگیر میانه نسبت به میانگین نشان می‌دهد توزیع followers به شدت راست‌چوله است: اغلب اکانت‌ها دنبال‌کنندگان کمی دارند، اما چند اکانت بسیار بزرگ مقدار میانگین را افزایش می‌دهند. این نکته در تحلیل شبکه‌های اجتماعی اهمیت دارد، زیرا اتكای صرف به میانگین می‌تواند تصویری گمراه‌کننده از ساختار مخاطب ارائه کند.

۸.۱۰.۴ اکانت‌های برتر انجمن ۳ بر اساس Weighted Degree

#	اکانت	followers	weighted_degree
1	aryammehr_2 (1638372580464685056)	54,538	2,134
2	sonofthebeach22 (1638665833240035330)	15,009	1,427
3	pahlavireza (38154763)	1,933,855	1,283
4	hichkasofficial (181198159)	903,055	1,180
5	alirezajafarzade (1490047253319333890)	140,223	1,047
6	hahamzadeh (1697244779413559296)	128,630	945
7	gh0lch0magh (1473474686977773568)	17,556	928
8	khodash_ (1242133768239155200)	33,692	886
9	yaarghadimi (1009441107364499456)	14,301	881
10	javidshah_ (827937322046083074)	37,790	826

جدول ۴۱: اکانت‌های برتر انجمن ۳ در Twitter بر اساس Weighted Degree

۹.۱۰.۴ اکانت‌های برتر انجمن ۳ بر اساس PageRank

#	اکانت	followers	pagerank
1	pahlavireza (38154763)	1,933,855	0.0417529339
2	aryammehr_2 (1638372580464685056)	54,538	0.0386080189
3	hichkasofficial (181198159)	903,055	0.0217266690
4	alirezajafarzade (1490047253319333890)	140,223	0.0173302302
5	hahamzadeh (1697244779413559296)	128,630	0.0163203210
6	khodash_ (1242133768239155200)	33,692	0.0132014149
7	purple9494 (1837507695215902720)	50,435	0.0114888443
8	javidshah_ (827937322046083074)	37,790	0.0109552078
9	mayekhal (151079169)	102,862	0.0105147051
10	sonofthebeach22 (1638665833240035330)	15,009	0.0097446662

جدول ۴۲: اکانت‌های برتر انجمن ۳ در Twitter بر اساس PageRank وزن‌دار.

۱۰.۱۰.۴ مقایسه Top10‌های Weighted Degree و PageRank

در این انجمن، میزان همپوشانی دو فهرست Top10 برابر با ۸ اکانت است. دو اکانت فقط در فهرست Weighted Degree قرار دارند:

yaarghadimi gh0lch0magh

و دو اکانت فقط در فهرست PageRank قرار گرفته‌اند:

purple9494 mayekhal

تفسیر Weighted Degree عمدتاً «تعامل مستقیم» را می‌سنجد (جمع وزن یال‌های مجاور هر گره درون انجمن). در مقابل، PageRank مفهومی از «نفوذ» را بازنمایی می‌کند که افزون بر وزن ارتباط‌ها، به اهمیت همسایه‌ها نیز وابسته است. از این‌رو ممکن است اکانتی که از نظر درجه وزن‌دار اندکی پایین‌تر است، اما توسط گره‌های مرکزی‌تر بازنشر می‌شود، در PageRank رتبه بالاتری کسب کند.

۱۱.۱۰.۴ رابطه followers با مرکزیت‌ها: آیا بزرگی مخاطب به معنای اثرباری در بازنشر است؟

برای بررسی کمی این موضوع، همبستگی بین followers و دو شاخص مرکزیت (درون همین زیرگراف) محاسبه شده است. با توجه به چولگی توزیع followers، دو سنجه گزارش می‌شود:

• (همبستگی رتبه‌ای) برای سنجش رابطه یکنواخت، بدون فرض خطی بودن، Spearman •

• بین $\log(1 + \text{followers})$ و مرکزیت‌ها، به منظور کاهش اثر چولگی Pearson.

نتایج برای انجمن ۳:

$$\rho(\text{followers}, \text{weighted_degree}) = 0.1945,$$

$$\rho(\text{followers}, \text{pagerank}) = 0.2068,$$

$$\rho(\text{weighted_degree}, \text{pagerank}) = 0.9361.$$

و برای Pearson :

$$r(\log(1 + \text{followers}), \text{weighted_degree}) = 0.1944,$$

$$r(\log(1 + \text{followers}), \text{pagerank}) = 0.1908.$$

تفسیر

• همبستگی‌های در بازه ۰.۱۹ تا ۰.۲۱ نشان می‌دهد «اندازه مخاطب» بی‌تأثیر نیست، اما عامل تعیین‌کننده اصلی در ساختار بازنشر محسوب نمی‌شود.

- همبستگی بسیار بالای weighted_degree طبیعی است، زیرا هر دو شاخص از ساختار یال‌ها استخراج می‌شوند؛ با این تفاوت که pagerank به کیفیت همسایه‌ها نیز حساس است.

۱۲.۱۰.۴ خروجی موضوع‌یابی NMF برای انجمن ۳

۱۵ عبارت/токن برتر هر موضوع

- موضوع ۱: چه در، چه، فراخوان، دی، بعدی صادر، پهلوی‌یو عده، فراخوان_عمومی اولین، خانه چه، فراخوان عمومی، پهلوی‌یو عده ما، عمومی شاهزاده، دی جمعه، دی راس، ان فراخوان، شعاردر

- موضوع ۲: co، https، co، https، https، co، https، co، جاویدشاه، دی، ۱۷ دی، ۱۷، ۲۵۸۴، ۲۵۸۴، دی ۲۵۸۴، ۲۵۸۴ جاویدشاه، مردم، دی ماه، ماه، به

- موضوع ۳: به، از، که، در، این، ایران، رو، را، می، است، پهلوی، مردم، با، برای، رضا پهلوی

نکتهٔ تفسیری روش‌شناختی در بسیاری از خروجی‌های NMF، یکی از موضوع‌ها معمولاً نقش «پس‌زمینه» را ایفا می‌کند (واژه‌های عمومی یا تکن‌های فنی مانند پیوندها). در این انجمن نیز موضوع ۲ شامل تکن‌های مرتبط با لینک (مانند https و co) است و می‌تواند بخش «فنی/لینک‌محور» متن‌ها را جذب کند. برای نامگذاری معنایی انجمن، معمولاً موضوع‌هایی که تکن‌های خاص‌تر و کم‌عمومی‌تر دارند (در اینجا موضوع‌های ۱ و ۳) اطلاعات تبیینی بیشتری فراهم می‌کنند.

۱۱.۴ مقایسهٔ چهار انجمن بزرگ Twitter با تمرکز بر ساختار و احساسات

برای نشان دادن این‌که تفاوت انجمن‌ها صرفاً ناشی از «اندازه» نیست، چند شاخص کلیدی به صورت همزمان مقایسه می‌شود.

cid	# گره	# یال داخلی	چگالی	conductance	سهم داخلی	posts	views	neg%
0	6,570	15,060	0.0006978954	0.4046276	0.5953724	70,570	30,058,617	80.37
1	6,485	16,619	0.0007904636	0.4407460	0.5592540	77,065	16,123,605	82.58
2	6,187	14,083	0.0007359277	0.2952052	0.7047948	62,262	31,850,564	83.54
3	4,749	18,448	0.0016363131	0.4074971	0.5925029	75,751	36,848,356	78.79

جدول ۴۳: مقایسه ۴ انجمن بزرگ Twitter.

۱۱.۴ کدام انجمن بسته‌تر است و چرا؟

کمترین مقدار conductance مربوط به انجمن 2 است (0.2952) و در عین حال، «سهم داخلی» آن بیشترین مقدار را دارد (0.7048). همزمانی این دو شاخص بیانگر آن است که:

- اعضای انجمن 2 بخش بزرگ‌تری از تعاملات خود را درون همان انجمن انجام می‌دهند،
- و نسبت وزن یال‌های خروجی آن کمتر است.

در نتیجه، مرزهای انجمن 2 از منظر شبکه‌ای شفاف‌تر بوده و گردش محظوظ به احتمال زیاد درون همان خوش‌نمایی‌تر است.

۱۱.۴ کدام انجمن متراکم‌تر است؟

بیشترین چگالی به انجمن 3 تعلق دارد (0.0016363) و میانگین درجه داخلی آن نیز بیشترین مقدار گزارش شده است (7.77). این نتیجه نشان می‌دهد با وجود آنکه انجمن 3 از نظر اندازه از سه انجمن دیگر کوچک‌تر است، پیوندهای داخلی آن فشرده‌تر بوده و احتمالاً «هسته بازنیز» فعال‌تری دارد.

۱۱.۴ مقایسه احساسات: آیا یک انجمن از سایرین مثبت‌تر است؟

در هر 4 انجمن، سهم negative بسیار بالا است (حدود 78% تا 84%). با این حال، انجمن 3 کمترین سهم negative را دارد (78.79%) و بیشترین سهم positive نیز برای آن گزارش شده است (19.60%). این اختلاف هرچند کوچک است، اما می‌تواند نشان دهد فضای گفت‌وگو/بازنشر در انجمن 3 نسبت به سایر انجمن‌های بزرگ Twitter اندکی کمتر منفی است.

۱۲.۴ گراف کامل Telegram: آمار ساختاری در مقیاس کل شبکه

این بخش برای درک زمینه کلی اهمیت دارد؛ زیرا رفتار انجمن‌ها به ساختار کلی گراف وابسته است. بر اساس فایل `:telegram_with_louvain.gexf`

- تعداد گره‌ها: 48632
- تعداد یال‌ها: 73798
- چگالی کل: 0.00006241
- میانگین درجه (بدون وزن): 3.03496

● تعداد گرهای ایزوله: 24824

● تعداد مؤلفه‌های همبند: 26332

تفسیر این مقادیر.

● وجود 24824 گره ایزوله نشان می‌دهد بخش بزرگی از «کانال‌ها/حساب‌ها» در بازه زمانی گردآوری شده، در شبکه بازنشر ثبت شده نقشی نداشته‌اند (یا تعامل آن‌ها در فایل یال‌ها ثبت نشده است).

● تعداد بسیار زیاد مؤلفه‌ها (26332) نشان می‌دهد شبکه Telegram در این داده به شدت قطعه‌قطعه است و تنها چند «هسته بزرگ» واقعاً شبکه پیوسته تشکیل می‌دهند.

● در چنین وضعیتی، وجود حتی یک انجمن بزرگ و منسجم از نظر تحلیلی بسیار مهم است؛ زیرا آن انجمن به احتمال زیاد کانون اصلی انتشار/بازنشر در داده بوده است.

کیفیت کلی انجمن‌بندی با **Modularity**. برای Telegram، مقدار modularity کل (با resolution=1.0) برابر است با:

$$Q_{\text{Telegram}} = 0.6594635.$$

این مقدار نسبتاً بالا بوده و نشان می‌دهد الگوریتم Louvain در این گراف (با این داده و این وزن‌ها) ساختار انجمنی قابل توجهی استخراج کرده است.

۱۳.۴ انجمن‌های بزرگ در Telegram

در این داده، تنها یک انجمن بزرگ از آستانه تعریف شده عبور کرده است:

● با اندازه 5482 community_id=0 (سهم 11.2724% از کل گرهای

بنابراین، برخلاف Twitter (که 4 انجمن بزرگ داشت)، در Telegram یک «هسته بزرگ غالب» مشاهده می‌شود و سایر انجمن‌ها کوچک‌تر بوده یا در مؤلفه‌های جداافتاده پراکنده شده‌اند. برای زیرگراف این انجمن:

● تعداد گره: 5482

● تعداد یال داخلی: 20392

● تعداد مؤلفه همبند: 1 (کاملاً همبند)

● تعداد گره ایزوله درون انجمن: 0

تفسیر «همبند بودن». همبند بودن کامل یک انجمن بزرگ به این معنا است که هر عضو، دست‌کم از طریق یک مسیر بازنشر (احتمالاً چند مرحله‌ای) به سایر اعضاء متصل است. این ویژگی در شبکه‌های بازنشر معمولاً نشان‌دهنده‌یک «اکوسیستم انتشار» واقعی است، نه صرفاً مجموعه‌ای از گره‌ها که صرفاً از نظر برچسب در یک گروه قرار گرفته‌اند.

۱.۱۳.۴ چگالی و میانگین درجه (بدون وزن) و معنای آن‌ها

برای $n = 5482$ و $m_{\text{in}} = 20392$

$$\text{Density} = \frac{2m_{\text{in}}}{n(n-1)} = 0.0013573473,$$

$$\overline{\text{deg}} = \frac{2m_{\text{in}}}{n} = 7.4396206.$$

تفسیر.

● میانگین درجه 7.44 نشان می‌دهد هر عضو این انجمن، به‌طور متوسط با حدود 7 عضو دیگر در داخل انجمن رابطه بازنشر ثبت شده دارد (این مقدار میانگین است و توزیع می‌تواند بسیار ناهمگن باشد).

● در مقایسه با میانگین درجه کل گراف Telegram (3.03)، این انجمن تقریباً «دو برابر» پرارتباطر است؛ بنابراین می‌توان آن را هسته اصلی فعالیت شبکه‌ای در این داده دانست.

۲.۱۳.۴ تحلیل وزن یال‌ها (شدت بازنشرهای تکراری)

مجموع وزن داخلی و درجه وزن‌دار متوسط.

$$W_{\text{in}} = 26141,$$

$$\overline{\text{deg}}_w = \frac{2W_{\text{in}}}{n} = 9.5370303.$$

آمار توزیع وزن یال‌های داخلی.

● میانگین وزن یال: 1.2819243

● میانه وزن یال: 1

● صدک 90: 2

● بیشینه: 69

برداشت تحلیلی از توزیع وزن‌ها. هم‌زمانی میانه 1 و صدک 90 برابر 2 نشان می‌دهد:

- بیشتر روابط بازنشر «کم‌تکرار» هستند (یک یا دو بار)،
● با این وجود، بیشینه 69 بیانگر وجود تعداد اندکی رابطه بسیار تکراری است.

چنین یال‌های سنگینی معمولاً میان کانال‌هایی شکل می‌گیرد که به صورت منظم از یکدیگر محتوا کپی/بازنشر می‌کنند و می‌توانند ستون‌های اصلی پخش محتوا باشند.

۳.۱۳.۴ جداول‌دادگی انجمن از کل شبکه: Conductance و Cut، Volume

این محاسبات بر روی گراف کامل telegram_with_louvain.gexf انجام شده است،
اما مجموعه اعضاء همان گرهای داخل telegram_louvain_id_0.gexf است.

مقادیر کلیدی.

$$\begin{aligned}\text{vol}(S) &= 57290, \\ \text{cut}(S, \bar{S}) &= 5008, \\ \phi(S) &= \frac{\text{cut}}{\text{vol}} = 0.0874149.\end{aligned}$$

سهم تعامل داخلی/خارجی.

$$\begin{aligned}\frac{2W_{\text{in}}}{\text{vol}(S)} &= 0.9125851, \\ \frac{\text{cut}}{\text{vol}(S)} &= 0.0874149.\end{aligned}$$

.conductance تفسیر

● مقدار 0.087 برای conductance کوچک است (در مقایسه با بسیاری از انجمن‌های Twitter که حدود 0.3 تا 0.44 گزارش شده‌اند).

- این نتیجه بیانگر درون‌گرایی بالای اعضای انجمن است: حدود 91.3% از وزن تعاملات incident به یال‌های داخلی اختصاص دارد.
- در شبکه‌های بازنشر، چنین الگویی معمولاً نشان‌دهنده خوش‌های نسبتاً بسته از کanal‌ها/اکانت‌ها است که عمدتاً از یکدیگر بازنشر می‌کنند و تبادل اندکی با بیرون دارند.

٤.١٣.٤ خوش‌بندی و مثلث‌ها: Transitivity و Average Clustering

$$\text{Clustering Average} = 0.1613486,$$

$$\text{Transitivity} = 0.0382658.$$

معنای این مقادیر در شبکه بازنشر.

- مقدار Average Clustering نسبتاً بزرگ است؛ یعنی در همسایگی بسیاری از گره‌ها، همسایه‌ها نیز تا حدی به یکدیگر متصل‌اند و ساختارهای سه‌تایی/گروهی شکل گرفته است.
- Transitivity نیز (در قیاس با مقادیر بسیار کوچک Twitter) بزرگ‌تر است و می‌تواند بیانگر وقوع بیشتر «حلقه‌های بازنشر» یا «گروه‌های سه‌تایی فعال» در Telegram باشد.
- توجه: در شبکه‌های بازنشر، خوش‌بندی بالا می‌تواند ناشی از بازنشر زنجیره‌ای چند کanal از یکدیگر و شکل‌گیری یک «گروه هماهنگ» باشد؛ این تفسیر صرفاً در حد فرضیه است و برای نتیجه‌گیری قطعی، بررسی‌های محتوایی/زمانی نیز لازم است.

٥.١٣.٤ تحلیل محتوایی/رفتاری انجمن Telegram

در داده‌ها، برای هر گره صفاتی مانند posts، views و نیز سه ستون neutral/negative/positive ذخیره شده است. با تجمعی مقادیر بر روی اعضای انجمن 0:

- مجموع posts: 132921
- مجموع views: 36641156

احساسات (به صورت تعداد و درصد).

(23.116% : سهم positive ●

(72.380% : سهم negative ●

(4.504% : سهم neutral ●

بنابراین، این سه ستون «تعداد پست‌ها به تفکیک احساس» را نشان می‌دهند (نه نمره احساس).

برداشت تحلیلی از توزیع احساسات.

● سهم negative همچنان غالب است (بیش از 72%)؛ با این حال، نسبت به انجمنهای بزرگ Twitter که معمولاً 78% تا 84% منفی گزارش شده‌اند، این انجمن Telegram اندکی کمتر منفی است.

● سهم neutral در این انجمن حدود 4.5% است که از Twitter (که در برخی انجمنهای حدود 1% تا 3% بوده است) بیشتر است. این امر می‌تواند نشان دهد بخشی از محتوا ماهیت «اطلاع‌رسانی/خنثی» داشته یا مدل احساس‌سنجی بر روی داده‌های Telegram خروجی خنثی بیشتری تولید کرده است.

۶.۱۳.۴ اکانت‌های برتر بر اساس Weighted Degree

#	اکانت	followers	weighted_degree
1	akhbarmashhad (1002941104)	1,174,590	1,060
2	iribnews (1004576149)	45,008	1,023
3	farsna (1006939659)	1,793,542	1,014
4	isna94 (1005665139)	131,765	855
5	mehrnews (1002765601)	331,280	838
6	chandsanieh_news (1069291200)	105,628	758
7	tasnimnews (1001976122)	2,077,397	571
8	akhbarefori (1004069147)	6,142,175	547
9	bisimchimedia (1006165084)	1,052,842	431
10	fararunews (1004135749)	89,651	393

جدول ۱۰: اکانت برتر انجمن ۰ در Telegram بر اساس Weighted Degree

۷.۱۳.۴ اکانت‌های برتر بر اساس PageRank

#	اکانت	followers	pagerank
1	akhbarmashhad (1002941104)	1,174,590	0.02787683
2	isna94 (1005665139)	131,765	0.01466079
3	farsna (1006939659)	1,793,542	0.01463345
4	mehrnews (1002765601)	331,280	0.01356863
5	iribnews (1004576149)	45,008	0.01267241
6	chandsanieh_news (1069291200)	105,628	0.01092414
7	akhbarefori (1004069147)	6,142,175	0.00965236
8	tabnak (1374693748)	252,575	0.00854018
9	tasnimnews (1001976122)	2,077,397	0.00753567
10	fararunews (1004135749)	89,651	0.00684734

جدول ۴۵: ۱۰ اکانت برتر انجمن ۰ در Telegram بر اساس PageRank.

۸.۱۳.۴ مقایسه PageRank و Weighted Degree: همپوشانی و برداشت مفهومی

دو فهرست Top10 در این انجمن ۹ عضو مشترک دارند و تنها یک تفاوت اصلی مشاهده می‌شود:

• در Weighted Degree حضور دارد ولی در PageRank نیست: bisimchimedia

• در Weighted Degree حضور دارد ولی در PageRank نیست: tabnak

تفسیر. معیار Weighted Degree بیشتر «حجم تعامل مستقیم» را می‌سنجد (این‌که یک کanal با چه تعداد کanal و با چه شدتی تبادل بازنشیر داشته است). در مقابل، PageRank علاوه بر حجم، به ساختار اتصالات نیز حساس است؛ بدین معنا که اگر یک کanal از سوی کanal‌های مرکزی‌تر تغذیه شود، ممکن است با درجه کمتر نیز مقدار PageRank بالاتری دریافت کند.

۹.۱۳.۴ رابطه followers با مرکزیت‌ها: آیا بزرگی به معنای اثرگذاری است؟

از آن‌جا که توزیع followers در شبکه‌های اجتماعی بهشدت چوله است، دو سنجه همبستگی گزارش می‌شود:

• Spearman (رتبه‌ای): بدون فرض خطی بودن.

• Pearson بر روی $\log(1 + \text{followers})$: برای کاهش اثر تعداد اندکی اکانت بسیار بزرگ.

.cid=0 ، Spearman برای Telegram

$$\rho(\text{followers}, \text{weighted_degree}) = 0.2900320,$$

$$\rho(\text{followers}, \text{pagerank}) = 0.2690197,$$

$$\rho(\text{weighted_degree}, \text{pagerank}) = 0.9362927.$$

با تبدیل لگاریتمی Pearson.

$$r(\log(1 + \text{followers}), \text{weighted_degree}) = 0.1876301,$$

$$r(\log(1 + \text{followers}), \text{pagerank}) = 0.1765950.$$

تفسیر.

- همبستگی رتبه‌ای در بازه 0.27 تا 0.29 نشان می‌دهد در این انجمن، «اندازه مخاطب» تا حدی با «مرکزیت شبکه‌ای» همراه است، اما تعیین‌کننده مطلق نیست.
- پس از تبدیل لگاریتمی، همبستگی خطی کاهش می‌یابد (حدود 0.18)؛ یعنی بخش قابل توجهی از رابطه ظاهری، ناشی از چند کanal بسیار بزرگ است و در مقیاس اکثریت کanal‌ها، بزرگی مخاطب الزاماً به معنای نقش مرکزی در بازنشر نیست.

۱۰.۱۳.۴ Telegram followers در انجمن

برای درک ناهمگنی اندازه مخاطب:

- میانگین 31202.946 :followers
- میانه 2175.5 :followers
- صدک 90 :followers
- بیشینه 6142175

تفسیر. فاصله بسیار زیاد میانگین و میانه نشان می‌دهد توزیع followers به شدت راست‌چوله است؛ بدین معنا که:

- بیشتر کanal‌ها/اکانت‌ها تعداد دنبال‌کننده نسبتاً کمی دارند،
- اما تعداد اندکی کanal بسیار بزرگ هستند و میانگین را به سمت بالا می‌کشند.

۱۱.۱۳.۴ موضوعیابی NMF برای Telegram/انجمن ۰

برای Telegram و ۰ community_id=

- تعداد اسناد استفاده شده: 8186

- تعداد موضوع ها: ۳

۱۲.۱۳.۴ واژه ها / توکن های برتر هر موضوع

- موضوع ۱: در، به، از، که، می، را، این، است، با، های، ایران، ها، برای، مردم، امریکا

- موضوع ۲: d9 86، d9 85، af، 86، a7 d8، d8 b1، db 8c، db a7، d9 88

- موضوع ۳: یارانه، ثبت، کالا برج، کنند، می توانند، تواند، می، کد، سامانه، کاربر، از، قبلا یارانه، خانوار، که قبلا، درخواست

۱۳.۱۳.۴ تفسیر محتوایی و تشخیص یک مسئله مهم در پیش پردازش

موضوع ۱ (پس زمینه زبانی). موضوع ۱ عمدهاً شامل واژه های پرتکرار و عمومی فارسی است (توقف واژه هایی مانند «در/به/از/که/می»). این نوع موضوع معمولاً «پس زمینه» را جذب می کند و برای نامگذاری معنایی انجمان به تنها یی چندان مفید نیست، مگر آن که در کنار موضوع های دیگر تحلیل شود.

موضوع ۳ (موضوع نسبتاً معنادار). موضوع ۳ واژه های نسبتاً مشخص و هم خانواده ای دارد (مانند «یارانه»، «ثبت»، «کالا برج»، «سامانه»، «خانوار»). این الگو نشان می دهد بخش قابل توجهی از متن های این انجمان حول محور اطلاع رسانی / بحث درباره یارانه، کالا برج، ثبت نام و مفاهیم مرتبط شکل گرفته است.

موضوع ۲ (علامت هشدار: توکن های خراب شده). موضوع ۲ تقریباً فقط از توکن های نامعمول مانند d8 و d9 تشکیل شده است. این الگو معمولاً زمانی مشاهده می شود که:

- متن فارسی به درستی decode نشده باشد، یا

- در مرحله پاکسازی، بخشی از کدهای UTF-8 به صورت جداگانه به عنوان «توکن» وارد مدل شده باشند، یا

- یک tokenizer/regex نامناسب، متن را به قطعات غیرمعنادار شکسته باشد.

۱۴.۴ گراف کامل Instagram: تحلیل ساختار کل شبکه

- تعداد گره‌ها: 35825
- تعداد یال‌ها: 9149
- چگالی کل: 0.0000142575
- میانگین درجه (بدون وزن): 0.5107606
- تعداد گره‌های ایزوله: 31617
- تعداد مؤلفه‌های همبند: 32042
- بیش از 88% گره‌ها (31617 گره) ایزوله هستند؛ یعنی در داده یال‌ها هیچ ارتباط بازنگشی/تعامل ثبت شده‌ای ندارند.
- وجود تعداد بسیار زیاد مؤلفه (32042) نشان می‌دهد شبکه عمدتاً از «جزیره‌های کوچک» و «گره‌های منفرد» ساخته شده و فقط چند هسته ارتباطی واقعی وجود دارد.
- در چنین ساختاری، روش Louvain معمولاً تعداد بسیار زیادی انجمن تک‌گرهی/چندگرهی تولید می‌کند و تنها یک یا دو انجمن واقعاً بزرگ و معنادار باقی می‌مانند.

۱۴.۴ کیفیت کلی انجمن‌بندی با Modularity

با همان برچسب‌های Louvain ذخیره شده در فایل و با $:resolution=0.2$

$$Q_{\text{Instagram}} = 0.8500877.$$

این مقدار بالا نشان می‌دهد (با توجه به وزن‌ها و ساختار داده) جداسازی انجمنی در شبکه قوی است.

۲.۱۴.۴ تعداد انجمن‌های استخراج شده و معنای آن

در این اجرا، تعداد کل انجمن‌ها برابر 32061 است. این عدد بزرگ با وجود حجم زیاد گره‌های ایزوله سازگار است؛ زیرا هر گره ایزوله عملاً می‌تواند یک انجمن جداگانه تشکیل دهد و شمار انجمن‌ها را افزایش دهد.

۱۵.۴ انجمن‌های بزرگ در Instagram

بر اساس فایل `instagram_big_communities_summary.csv` تنها یک انجمن از آستانه «بزرگ بودن» عبور کرده است:

- `community_id=0` با اندازه 2205 (سهم 6.1549% از کل گره‌ها)

بنابراین تمرکز تحلیل انجمن‌های بزرگ در Instagram بر انجمن 0 قرار می‌گیرد.

۱۶.۴ تحلیل ساختاری انجمن بزرگ Instagram با شناسه ۰

۱.۱۶.۴ مشخصات پایه زیرگراف انجمن

زیرگراف القایی انجمن 0 در فایل

`instagram_louvain_id_0.gexf`

قرار دارد و دارای ویژگی‌های زیر است:

- تعداد گره: 2205

- تعداد یال داخلی: 7265

- تعداد مؤلفه همبند: 1 (کاملاً همبند)

- تعداد گره ایزوله درون انجمن: 0

۲.۱۶.۴ چگالی و درجه متوسط (بدون وزن)

$$\text{Density} = \frac{2m_{\text{in}}}{n(n-1)} = \frac{2 \times 7265}{2205 \times 2204} = 0.0029898227,$$

$$\overline{\text{deg}} = \frac{2m_{\text{in}}}{n} = \frac{2 \times 7265}{2205} = 6.5895692.$$

- این چگالی نسبت به چگالی کل شبکه (0.0000143) چندین مرتبه بزرگ‌تر است؛ بنابراین شبکه Instagram دارای یک هسته بسیار متراکم بوده و بیرون از آن عمدتاً کم تعامل/ایزوله است.
- میانگین درجه 6.59 بیانگر آن است که هر عضو، به طور متوسط با حدود 6 تا 7 عضو دیگر درون همین تعامل بازنشر ثبت شده دارد.

۳.۱۶.۴ سهم انجمن ۰ از کل شبکه

با وجود آنکه انجمن ۰ تنها 6.15% از گره‌ها را در بر می‌گیرد، از نظر تعداد یال و وزن، سهم بسیار بزرگ‌گی دارد:

- سهم از کل یال‌ها (فقط یال‌های داخلی انجمن):

$$\frac{7265}{9149} = 0.7940759 \quad (79.41\%)$$

- سهم از کل وزن یال‌ها (فقط وزن داخلی انجمن):

$$\frac{W_{\text{in}}}{W_{\text{total}}} = \frac{11559}{14208} = 0.8135557 \quad (81.36\%)$$

این نتیجه نشان می‌دهد بخش عمده «فعالیت بازنشر/تعامل» در داده Instagram بر روی تعداد نسبتاً کمی اکانت متمرکز شده و سایر گره‌ها یا فاقد تعامل‌اند یا در جزیره‌های کوچک پراکنده شده‌اند.

۴.۱۶.۴ تحلیل وزن یال‌ها (شدت تعامل‌های تکرارشونده)

مجموع وزن داخلی و درجه وزن‌دار متوسط.

$$W_{\text{in}} = \sum_{e \in E(S)} w_e = 11559, \quad \overline{\deg}_w = \frac{2W_{\text{in}}}{n} = \frac{2 \times 11559}{2205} = 10.4843537.$$

آمار توزیع وزن یال‌های داخلی.

- میانگین وزن یال: 1.5910530

● میانه وزن یال: 1

● صدک 90: 2

● بیشینه وزن یال: 110

هم زمانی میانه 1 و صدک 90 برابر 2 نشان می دهد اکثریت روابط کم تکرار هستند (یک یا دو بار). با این حال، بیشینه 110 بیانگر آن است که برخی ارتباطها بسیار تکراری بوده و احتمالاً میان اکانت هایی شکل گرفته اند که به صورت سیستماتیک از یکدیگر بازنشر می کنند یا رابطه پایداری دارند.

۱۶.۴ اتصال به بیرون: Conductance و Cut ، Volume

با مجموعه اعضای انجمن 0 (گره های زیرگراف) بر روی گراف کامل `instagram_with_louvain.gexf` به صورت زیر انجام شده است.

مقادیر کلیدی.

$$\text{vol}(S) = 23326, \quad \text{cut}(S, \bar{S}) = 208.$$

حجم متمم:

$$\text{vol}(\bar{S}) = 5090.$$

از آنجا که $\min(\text{vol}(S), \text{vol}(\bar{S})) = \text{vol}(\bar{S})$ ، تعريف استاندارد conductance به صورت زیر به دست می آید:

$$\phi(S) = \frac{208}{5090} = 0.04086444.$$

بررسی سازگاری حجم.

$$2W_{\text{in}} + \text{cut} = 2 \times 11559 + 208 = 23326 = \text{vol}(S).$$

این برابری نشان می دهد وزن داخلی و وزن یال های خروجی به صورت دقیق و سازگار محاسبه شده اند.

نکته روش شناختی درباره conductance در این داده. در بسیاری از مثال ها، انجمن S کوچک تر از متمم آن است و در نتیجه مخرج conductance برابر $\text{vol}(S)$ قرار می گیرد. اما در Instagram،

انجمن 0 از نظر «حجم وزن دار» آنقدر بزرگ است که متمم آن حجم کوچکتری دارد؛ از این رو علاوه‌گلوبال اتصال «بخش کوچک باقی‌مانده» به «هسته بزرگ» را می‌سنجد. برای مشاهده درون‌گرایی از زاویه خود انجمن 0 نیز «سهم مرزی از دید انجمن» گزارش می‌شود:

$$\frac{\text{cut}}{\text{vol}(S)} = \frac{208}{23326} = 0.00891709 \quad (0.89\%).$$

همچنان، سهم تعامل داخلی از دید اعضای انجمن:

$$\frac{2W_{\text{in}}}{\text{vol}(S)} = \frac{23118}{23326} = 0.99108291 \quad (99.11\%).$$

بنابراین، از دید اعضای انجمن، تقریباً تمام تعامل وزن دار درون انجمن رخ می‌دهد و انجمن 0 یک خوش‌بیانگر است؛ حتی اگر تعریف استاندارد conductance (به علت کوچکتر بودن حجم بیرون) مخرج را $\text{vol}(\bar{S})$ در نظر بگیرد.

۶.۱۶.۴ خوشبندی و مثلث‌ها: Transitivity و Average Clustering

در زیرگراف انجمن 0:

$$\text{Clustering Average} = 0.10802324, \quad \text{Transitivity} = 0.08178773.$$

- مقدار Average Clustering حدود 0.108 نشان می‌دهد در همسایگی بسیاری از گره‌ها، میزانی از «بسته شدن سه‌تایی» وجود دارد (همسایه‌ها تا حدی به هم متصل‌اند).
- مقدار Transitivity نیز بیانگر آن است که در مقیاس کل انجمن، مثلث‌ها نسبتاً قابل توجه‌اند؛ در شبکه‌های بازنشر، افزایش این مقدار معمولاً زمانی رخ می‌دهد که گروه‌هایی از اکانت‌ها به صورت چندجانبه از یکدیگر بازنشر کنند.

۷.۱۶.۴ تحلیل محتوایی/رفتاری انجمن Instagram

برای هر گره، صفات neutral/negative/posts و نیز views ذخیره شده است. برای انجمن 0 (با تجمعی بر روی همه اعضا):

• مجموع 56945 posts

• مجموع 852867558 views

میانگین‌های رفتاری.

$$\frac{\text{views}}{\text{posts}} = \frac{852867558}{56945} = 14977.0403$$

(میانگین بازدید به ازای هر پست)،

$$\frac{\text{posts}}{\#\text{accounts}} = \frac{56945}{2205} = 25.8254$$

(میانگین پست به ازای هر اکانت).

این دو مقدار امکان کمی‌سازی شدت فعالیت و سطح توجه در این انجمن را فراهم می‌کند؛ هرچند توزیع‌ها می‌توانند ناهمگن باشند و میانگین تحت تأثیر تعداد اندکی اکانت بسیار بزرگ قرار گیرد.

۸.۱۶.۴ تقسیم احساسات (تعداد و درصد)

(سهم: 23.8212%) 13565 :positive •

(سهم: 71.2319%) 40563 :negative •

(سهم: 4.9469%) 2817 :neutral •

در این انجمن، احساسات منفی غالب است (حدود 71%)، اما سهم مثبت نیز قابل توجه است (23.8%). سهم neutral نزدیک به 5% می‌تواند بیانگر آن باشد که بخشی از محتوا خنثی/اطلاع‌رسانی بوده یا مدل احساس‌سنجی در این پلتفرم خروجی خنثی بیشتری تولید کرده است.

۹.۱۶.۴ توزیع followers در انجمن Instagram و ناهمگنی مخاطب

برای صفت followers در انجمن 0:

• میانگین 161230.134 :followers

• میانه 17815 :followers

• صدک 90:90

• بیشینه 23601288 :followers

فاصله زیاد میانگین و میانه نشان می‌دهد توزیع followers به شدت راست‌چوله است: اکثر اکانت‌ها مخاطب متوسط/کمتری دارند، اما چند اکانت بسیار بزرگ (با میلیون‌ها دنبال‌کننده) میانگین را افزایش می‌دهند. از این رو، در گزارش علمی، میانه و صدک‌ها برای ارائه «تصویر دقیق از عضو معمولی» ضروری‌اند.

۱۰.۱۶.۴ اکانت‌های اثربار انجمن Instagram بر اساس مرکزیت‌های شبکه‌ای

در خروجی‌ها دو معیار اصلی گزارش شده است:

Weighted Degree: مجموع وزن یال‌های incident به هر گره در زیرگراف انجمن.

PageRank: نفوذ شبکه‌ای بر مبنای random walk وزن‌دار (در NetworkX).

۱۱.۱۶.۴ اکانت‌های برتر بر اساس Weighted Degree

#	اکانت	followers	weighted_degree
1	avatoday1 (10322006202)	1,549,983	639
2	officialrezapahlavi (1949382920)	9,278,948	483
3	iranintltv (5496840166)	16,906,381	306
4	indypersian (13521455499)	2,111,917	303
5	unknown (52063150785)	N/A	242
6	reza.davoudi.pahlavism (54539779400)	53,021	238
7	iran.janam.2022 (56463117828)	4,710	222
8	yasamankhodajo (73222501641)	31,711	207
9	manotoofficial (364081256)	17,857,820	190
10	azaddokht_ (49743276108)	23,948	172

جدول ۴۶: اکانت برتر انجمن ۰ در Instagram بر اساس Weighted Degree.

۱۲.۱۶.۴ اکانت‌های برتر بر اساس PageRank

#	اکانت	followers	pagerank
1	avatoday1 (10322006202)	1,549,983	0.02030569
2	officialrezapahlavi (1949382920)	9,278,948	0.01833885
3	unknown (52063150785)	N/A	0.01338219
4	iranintltv (5496840166)	16,906,381	0.00980049
5	indypersian (13521455499)	2,111,917	0.00945002
6	reza.davoudi.pahlavism (54539779400)	53,021	0.00867674
7	saalhaay_e_door (1576482205)	50,710	0.00741708
8	azaddokht_ (49743276108)	23,948	0.00673272
9	manotoofficial (364081256)	17,857,820	0.00653352
10	iran.janam.2022 (56463117828)	4,710	0.00651746

جدول ۴۷: اکانت برتر انجمن ۰ در Instagram بر اساس PageRank وزن‌دار.

۱۳.۱۶.۴ همپوشانی دو فهرست و برداشت مفهومی

دو فهرست Top10 در این انجمن ۹ عضو مشترک دارند:

● فقط در (73222501641) yasamankhodajo :Weighted Degree

● فقط در (1576482205) saalhaay_e_door :PageRank

PageRank بیشتر «حجم تعامل مستقیم» را بازنمایی می‌کند؛ در حالی که Weighted Degree علاوه بر حجم، به کیفیت اتصالات نیز حساس است. از این رو، ممکن است یک اکانت با درجه اندکی پایین‌تر، به دلیل اتصال به hub‌های قوی‌تر، مقدار PageRank بالاتری به دست آورد.

۱۴.۱۶.۴ آیا followers با مرکزیت شبکه‌ای همبسته است؟

با توجه به چوله‌بودن شدید followers، دو سنجه همبستگی ارائه می‌شود:

● Spearman: همبستگی رتبه‌ای (بدون فرض خطی بودن).

● Pearson: بر روی $\log(1 + \text{followers})$ برای کاهش اثر تعداد اندکی اکانت بسیار بزرگ.

برای :cid=0، Instagram

$$\rho(\text{followers}, \text{weighted_degree}) = 0.09438068,$$

$$\rho(\text{followers}, \text{pagerank}) = 0.10031866,$$

$$\rho(\text{weighted_degree}, \text{pagerank}) = 0.92152698.$$

و برای Pearson :

$$r(\log(1 + \text{followers}), \text{weighted_degree}) = 0.12193478,$$

$$r(\log(1 + \text{followers}), \text{pagerank}) = 0.10642470.$$

● همبستگی followers با مرکزیت‌ها در این انجمن ضعیف است (0.09 تا 0.12)؛ بنابراین بزرگی مخاطب الزاماً به معنای «نقش مرکزی در بازنشر/تعامل داخل این خوشه» نیست.

● همبستگی بسیار بالای weighted_degree با pagerank با قابل انتظار است؛ زیرا هر دو معیار از ساختار یال‌ها استخراج می‌شوند، با این تفاوت که PageRank کیفیت همسایه‌ها را نیز در نظر می‌گیرد.

۱۵.۱۶.۴ موضوعیابی NMF برای Instagram/انجمن ۰

● تعداد اسناد استفاده شده: 5745

● تعداد موضوعات: 3

۱۶.۱۶.۴ توکن‌ها/عبارت‌های برتر هر موضوع

● موضوع ۱: در، به، از، که، است، را، این، می، با، ایران، های، ها، مردم، شده، برای

● موضوع ۲: دی، ۱۷، ۱۷ دی، ۱۴۰۴، در، دی ماه، ماه، چهارشنبه، دی ۱۴۰۴، چهارشنبه ۱۷، اعتصاب، ۱۸، ۱۸ دی، اعتراضات سراسری، اعتراضات

● موضوع ۳: شاه، جاوید، جاوید شاه، جاوید شاه، پهلوی، رضا، پاینده، شعار، فراخوان، این آخرین نبرده پهلوی برمیگردد، شاهزاده رضا پهلوی، رضا پهلوی

۱۷.۱۶.۴ تفسیر موضوعات و نکته روش‌شناختی

موضوع ۱ (پس زمینه/توقف و ازهای). این موضوع عمدتاً از واژه‌های پر تکرار عمومی فارسی تشکیل شده است و معمولاً «پس زمینه زبانی» را جذب می‌کند. برای بهبود کیفیت موضوعیابی، در حالت ایده‌آل توقف و ازهای فارسی باید حذف شوند تا مدل بر واژه‌های محتوایی تر تمرکز کند.

موضوع ۲ (زمان/رخداد). تکرار الگوهایی مانند «۱۷ دی»، «۱۸ دی»، «۱۴۰۴»، «اعتصاب» و «اعتراضات» نشان می‌دهد بخشی از محتوای این انجمن حول زمان‌بندی یک رخداد/فراخوان/اعتراضات شکل گرفته است.

موضوع ۳ (هویت/شعار/هشتگ). وجود ترکیب‌هایی مانند «جاوید شاه»، «پهلوی»، «رضا پهلوی»، «فراخوان» و هشتگ‌های طولانی نشان می‌دهد بخش دیگری از محتوای این انجمن حول شعارها و هشتگ‌های هویتی-سیاسی جریان دارد.

مقایسه انجمن‌یابی و Louvain و Infomap

۱.۵ تعریف شبکه و وزن یال‌ها

در فایل‌های GDF، یال‌های شبکه جهت‌دار هستند و جهت هر یال از «اکانت بازنشرکننده/ارجاع‌دهنده» به «اکانت مرجع بازنشر» ثبت شده است. وزن یال نیز تعداد دفعات بازنشر از اکانت ارجاع‌دهنده به اکانت مرجع را نشان می‌دهد. در پلتفرم Instagram، به‌دلیل مشخص نبودن مسیر دقیق بازنشر، اکانتی که نخستین بار یک محتوا را منتشر کرده به عنوان مرجع در نظر گرفته شده و سایر اکانت‌هایی که همان محتوا را منتشر کرده‌اند به آن لینک شده‌اند؛ بنابراین، تفسیر یال‌ها در این پلتفرم به صورت «ارجاع به نخستین منتشرکننده» است.

۲.۵ روش اجرا و معیارهای مقایسه

برای هر پلتفرم، گراف جهت‌دار و وزن‌دار ساخته شد و دو روش Louvain و Infomap بر روی همین گراف اجرا شدند. سپس برای مقایسه دو خروجی، انجمن‌ها/ماژول‌ها بر اساس اندازه مرتب شدند و تنها 10 خوش‌بزرگ‌تر از هر روش در نظر گرفته شد. از آنجا که شناسه خوش‌ها در دو الگوریتم هم‌تراز نیست، برای هر خوشة C از روش نخست، خوشة D از روش دوم که بیشترین هم‌پوشانی را دارد با معیار Jaccard انتخاب شد:

$$J(C, D) = \frac{|C \cap D|}{|C \cup D|}.$$

علاوه بر Jaccard، دو معیار پوشش نیز برای تفسیر «جهت هم‌پوشانی» گزارش شد:

$$\frac{|C \cap D|}{c}, \quad \frac{|C \cap D|}{|D|}.$$

بر این اساس، اگر مقدار $\frac{|C \cap D|}{c}$ بزرگ و مقدار $\frac{|C \cap D|}{|D|}$ کوچک باشد، می‌توان نتیجه گرفت که خوشة C تقریباً به صورت زیرمجموعه‌ای از خوشة D قرار گرفته است و برعکس.

۳.۵ خلاصه ابعاد شبکه و خروجی کلی

پلتفرم	تعداد نود	تعداد یال	تعداد انجمن در Louvain	تعداد مژول در Infomap	زمان اجرای Infomap (ثانیه/دقیقه)	codeLength
Twitter	85021	151318	42726	3261	(13.10) 785.94	6.922
Telegram	48632	75395	26445	2618	(1.07) 64.32	6.850
Instagram	35825	9429	32077	652	(0.02) 1.39	6.045

جدول ۴۸: خلاصه ابعاد شبکه و خروجی کلی دو روش روی گراف جهت‌دار وزن‌دار.

۴.۵ نتایج پلتفرم Twitter

در این پلتفرم، گراف جهت‌دار وزن‌دار شامل 85021 نود و 151318 یال است. اجرای Louvain روی این گراف به 42726 انجمن منجر شد که میانگین اندازه هر انجمن تقریباً 1.99 نود است؛ در مقابل، Infomap با تنظیمات سریع به 3261 مژول رسید و میانگین اندازه مژول‌ها حدود 26.07 نود شد. بزرگ‌ترین انجمن Louvain در میان 10 انجمن بزرگ‌تر اندازه 6532 داشت و بزرگ‌ترین مژول Infomap در میان 10 مژول بزرگ‌تر اندازه 1552 بود. زمان اجرای Infomap برابر 785.94 ثانیه (حدود 13.10 دقیقه) و مقدار codeLength برابر 6.922 به دست آمد.

به عنوان نمونه، در میان 10 مژول بزرگ‌تر Infomap، بهترین همپوشانی مربوط به مژول 1 با اندازه 1470 و انجمن 29868 از Louvain با اندازه 1633 است؛ در این حالت، $|C \cap D| = 1434$ ، مقدار $J(C, D) = 0.859$ و پوشش‌ها به ترتیب $\frac{|C \cap D|}{c} = 0.976$ و $\frac{|C \cap D|}{|D|} = 0.878$ گزارش شد. جزئیات تطبیق کامل 10 خوش‌بزرگ‌تر در جدول‌های ۴۹ و ۵۰ ارائه شده است.

ردیف	شناسه	Infomap	اندازه Infomap	بهترین تطبیق در Louvain	اندازه Louvain	Louvain	Infomap	$\frac{ C \cap D }{ D }$	$\frac{ C \cap D }{c}$	$J(C, D)$	$ C \cap D $
1	1	2	1552	27262	5636	1214	0.203	0.215	0.782	0.203	1214
2	2	1	1470	29868	1633	1434	0.859	0.878	0.976	0.859	1434
3	3	32	1170	25333	4972	1111	0.221	0.223	0.950	0.221	1111
4	4	6	1139	35038	3434	808	0.215	0.235	0.709	0.215	808
5	5	3	858	20085	5193	623	0.115	0.120	0.726	0.115	623
6	6	45	537	37056	6532	519	0.079	0.079	0.966	0.079	519
7	7	88	526	8140	1067	513	0.507	0.481	0.975	0.507	513
8	8	4	465	23880	1130	412	0.348	0.365	0.886	0.348	412
9	9	87	415	8140	1067	397	0.367	0.372	0.957	0.367	397
10	10	49	381	8140	1067	360	0.332	0.337	0.945	0.332	360

جدول ۴۹: تطبیق 10 مژول بزرگ‌تر روش Infomap با بهترین انجمن متناظر در Louvain برای پلتفرم Twitter.

$\frac{ C \cap D }{ D }$	$\frac{ C \cap D }{c}$	$J(C, D)$	$ C \cap D $	اندازه Infomap	بهترین نتیجه در Infomap	اندازه Louvain	اندازه Louvain	شناسه	رتبه
0.966	0.079	0.079	519	537	45	6532	37056	1	
0.782	0.215	0.203	1214	1552	2	5636	27262	2	
0.149	0.044	0.035	231	1552	2	5207	37948	3	
0.726	0.120	0.115	623	858	3	5193	20085	4	
0.950	0.223	0.221	1111	1170	32	4972	25333	5	
0.716	0.121	0.115	535	747	11	4422	24066	6	
0.709	0.235	0.215	808	1139	6	3434	35038	7	
0.326	0.110	0.089	371	1139	6	3386	37734	8	
0.565	0.123	0.112	414	733	47	3359	7118	9	
0.041	0.024	0.024	60	1470	1	2486	38418	10	

جدول ۵۰: تطبیق ۱۰ انجمن بزرگ‌تر روش Louvain با بهترین ماذول متناظر در Infomap برای پلتفرم Twitter.

۵.۵ نتایج پلتفرم Telegram

در این پلتفرم، گراف جهت‌دار وزن دار شامل 48632 نود و 75395 یال است. اجرای Louvain روی این گراف به 26445 انجمن منجر شد که میانگین اندازه هر انجمن تقریباً ۱.۸۴ نود است؛ در مقابل، Infomap با تنظیمات سریع به 2618 ماذول رسید و میانگین اندازه ماذول‌ها حدود ۱۸.۵۸ نود شد. بزرگ‌ترین انجمن Louvain در میان ۱۰ انجمن بزرگ‌تر اندازه ۵۵۸۵ داشت و بزرگ‌ترین ماذول Infomap در میان ۱۰ ماذول بزرگ‌تر اندازه ۲۸۴۴ بود. زمان اجرای Infomap برابر ۶۴.۳۲ ثانیه (حدود ۱.۰۷ دقیقه) و مقدار codelength برابر ۶.۸۵۰ به دست آمد.

به عنوان نمونه، در میان ۱۰ ماذول بزرگ‌تر Infomap، بهترین همپوشانی مربوط به ماذول ۴ با اندازه ۴۶۳ و انجمن ۱۶ از Louvain با اندازه ۶۶۱ است؛ در این حالت، $|C \cap D| = 451$ و $J(C, D) = 0.670$ مقدار $\frac{|C \cap D|}{|D|} = 0.682$ و $\frac{|C \cap D|}{c} = 0.974$ گزارش شد. جزئیات تطبیق کامل ۱۰ خوشه بزرگ‌تر در جدول‌های ۵۱ و ۵۲ ارائه شده است.

$\frac{ C \cap D }{ D }$	$\frac{ C \cap D }{c}$	$J(C, D)$	$ C \cap D $	Louvain	اندازه Louvain	بهترین تطبیق در Infomap	اندازه Infomap	شناسه Infomap	رتبه
0.462	0.907	0.439	2581	5585	4394	2844	0	1	
0.951	0.957	0.915	2146	2257	12237	2242	1	2	
0.200	0.520	0.170	1115	5585	4394	2145	2	3	
0.109	0.494	0.097	606	5585	4394	1227	3	4	
0.682	0.974	0.670	451	661	16	463	4	5	
0.648	0.928	0.619	428	661	16	461	5	6	
0.564	0.836	0.509	373	661	16	446	6	7	
0.812	0.960	0.790	407	501	25	424	7	8	
0.048	0.633	0.046	266	5585	4394	420	8	9	
0.578	0.965	0.569	382	661	16	396	9	10	

جدول ۵۱: تطبیق ۱۰ مژوول بزرگ‌تر روش Infomap با بهترین انجمن متناظر در Louvain برای Telegram پلتفرم.

$\frac{ C \cap D }{ D }$	$\frac{ C \cap D }{c}$	$J(C, D)$	$ C \cap D $	Infomap	اندازه Infomap	بهترین تطبیق در Louvain	اندازه Louvain	شناسه Louvain	رتبه
0.907	0.462	0.439	2581	2844	0	5585	4394	1	
0.957	0.951	0.915	2146	2242	1	2257	12237	2	
0.138	0.177	0.066	393	2844	0	2223	5234	3	
0.180	0.237	0.075	511	2844	0	2158	21826	4	
0.102	0.150	0.040	289	2844	0	1933	25344	5	
0.025	0.033	0.024	56	2242	1	1681	7301	6	
0.983	0.970	0.960	1206	1227	3	1243	16543	7	
0.974	0.682	0.670	451	463	4	661	16	8	
0.960	0.812	0.790	407	424	7	501	25	9	
0.083	0.072	0.041	35	424	7	489	838	10	

جدول ۵۲: تطبیق ۱۰ انجمن بزرگ‌تر روش Louvain با بهترین مژوول متناظر در Infomap برای Telegram پلتفرم.

۶.۵ نتایج پلتفرم Instagram

در این پلتفرم، گراف جهت دار وزن دار شامل 35825 نود و 9429 یال است. اجرای Louvain روی این گراف به 32077 انجمن منجر شد که میانگین اندازه هر انجمن تقریباً 1.12 نود است؛ در مقابل، Infomap با تنظیمات سریع به 652 مژوول رسید و میانگین اندازه مژوولها حدود 54.95 نود شد. بزرگ‌ترین انجمن Louvain در میان 10 انجمن بزرگ‌تر اندازه 446 داشت و بزرگ‌ترین مژوول Infomap در میان 10 مژوول بزرگ‌تر اندازه 214 بود. زمان اجرای Infomap برابر 1.39 ثانیه (حدود 0.02 دقیقه) و مقدار codelength برابر 6.045 به دست آمد.

به عنوان نمونه، در میان 10 مژوول بزرگ‌تر Infomap، بهترین همپوشانی مربوط به مژوول 21 با اندازه 140 و انجمن 12731 از Louvain با اندازه 148 است؛ در این حالت، $|C \cap D| = 132$

مقدار $J(C, D) = 0.846$ و پوشش‌ها به ترتیب $\frac{|C \cap D|}{|D|} = 0.943$ و $\frac{|C \cap D|}{c} = 0.892$ گزارش شد. جزئیات تطبیق کامل ۱۰ خوشه بزرگ‌تر در جدول‌های ۵۳ و ۵۴ ارائه شده است.

$\frac{ C \cap D }{ D }$	$\frac{ C \cap D }{c}$	$J(C, D)$	$ C \cap D $	Louvain	اندازه	بهترین تطبیق در	Infomap	اندازه	Infomap	شناسه	رتبه
0.928	0.967	0.896	207	223	7902		214		0		1
0.858	0.960	0.831	194	226	5908		202		1		2
0.956	0.975	0.923	196	205	16924		201		2		3
0.010	0.010	0.005	2	205	16924		196		3		4
0.897	0.961	0.867	174	194	7253		181		4		5
0.879	0.950	0.841	153	174	12345		161		5		6
0.009	0.013	0.005	2	226	5908		155		6		7
0.013	0.020	0.008	3	223	7902		149		7		8
0.878	0.951	0.841	137	156	14989		144		8		9
0.892	0.943	0.846	132	148	12731		140		21		10

جدول ۵۳: تطبیق ۱۰ ماثول بزرگ‌تر روش Infomap با بهترین انجمن متناظر در Louvain برای پلتفرم Instagram.

$\frac{ C \cap D }{ D }$	$\frac{ C \cap D }{c}$	$J(C, D)$	$ C \cap D $	Infomap	اندازه	بهترین تطبیق در	Louvain	اندازه	Louvain	شناسه	رتبه
0.967	0.928	0.896	207	214	0		223		7902		1
0.960	0.858	0.831	194	202	1		226		5908		2
0.975	0.956	0.923	196	201	2		205		16924		3
0.961	0.897	0.867	174	181	4		194		7253		4
0.950	0.879	0.841	153	161	5		174		12345		5
0.951	0.878	0.841	137	144	8		156		14989		6
0.943	0.892	0.846	132	140	21		148		12731		7
0.943	0.975	0.916	116	123	12		119		4136		8
0.006	0.009	0.003	1	161	5		108		25190		9
0.910	0.981	0.847	101	111	18		103		5423		10

جدول ۵۴: تطبیق ۱۰ انجمن بزرگ‌تر روش Louvain با بهترین ماثول متناظر در Infomap برای پلتفرم Instagram.

تخمین داخل/خارج بودن اکانت‌های اثرگذار با اتكا به قطع اینترنت

۱.۶ داده‌ها و ساخت گراف

برای هر پلتفرم (Twitter، Telegram و Instagram) یک فایل GDF شامل شبکه‌ی بازنشر در بازه‌ی 10 تا 24 دی در دسترس بود. مطابق تعریف مسئله، یال‌ها جهت‌دار هستند و جهت هر یال از «اکانت بازنشرکننده/ارجاع‌دهنده» به «اکانت مرجع بازنشر» در نظر گرفته می‌شود؛ وزن هر یال نیز برابر با تعداد دفعات بازنشر از مبدأ به مقصد است. برای این اساس، برای هر پلتفرم گراف جهت‌دار و وزن‌دار $G = (V, E)$ به صورت یک DiGraph ساخته شد و در صورت وجود یال‌های تکراری بین یک زوج مرتب (u, v) ، وزن‌ها به صورت زیر تجمعی گردید:

$$w_{uv} = \sum_{e: u \rightarrow v} w(e).$$

پس از ساخت گراف‌ها، اندازه‌ی آن‌ها به ترتیب چنین بود: Twitter با $|V| = 85021$ و $|E| = 151318$ ، Instagram با $|V| = 75395$ و $|E| = 48632$ ، Telegram با $|V| = 35825$ و $|E| = 9429$ (همگی جهت‌دار).

۲.۶ محاسبه‌ی PageRank وزن‌دار و استخراج 1000 اکانت برتر

برای هر گراف، مقدار PageRank وزن‌دار با ضریب میرایی $\alpha = 0.85$ محاسبه شد. در نسخه‌ی وزن‌دار، سهم هر یال ورودی متناسب با وزن آن و پس از نرمال‌سازی بر مجموع وزن خروجی‌های مبدأ تعریف می‌شود. برای هر گره v داریم:

$$\text{PR}(v) = \frac{1 - \alpha}{|V|} + \alpha \sum_{(u,v) \in E} \frac{w_{uv}}{\sum_{(u,x) \in E} w_{ux}} \text{PR}(u).$$

پس از همگرایی، گره‌ها بر اساس مقدار PR مرتب شدند و 1000 اکانت نخست به عنوان «اکانت‌های اثرگذار» انتخاب گردید. با توجه به امکان تکرار شناسه‌ی گره در برخی فایل‌ها (به دلیل تکرار در بخش nodedef و الحاق ویژگی‌ها)، پس از مرتب‌سازی بر حسب PageRank، شناسه‌ها

یکتا شدند و در صورت تکرار، رکورد با PageRank بزرگ‌تر نگه داشته شد تا در نهایت دقیقاً 1000 اکانت یکتا برای هر پلتفرم باقی بماند.

۳.۶ ویژگی‌های فعالیت زمانی و برچسب‌گذاری داخل/خارج ایران

برای مکان‌یابی تقریبی اکانت‌ها، از سرنخ مسئله درباره‌ی «قطع تدریجی اینترنت در روز 17 دی از حوالی عصر تا پایان شب» استفاده شد و این رخداد به مثابه‌ی یک آزمایش طبیعی در رفتار زمانی کاربران مدل‌سازی گردید. برای هر پلتفرم، 5 فایل XLSX مربوط به روزهای 15 تا 19 دی در دسترس بود که هر فایل شامل 5000 پست پربازنشر همان روز است. ستون زمان (date) ابتدا در قالب UTC تفسیر و سپس به ناحیه‌ی زمانی ایران Asia/Tehran تبدیل شد. برای سازگاری با برخی محیط‌های Windows که داده‌ی ناحیه‌ی زمانی به صورت پیش‌فرض در دسترس نیست، استفاده از IANA TZ Database از طریق zoneinfo و در صورت نیاز بسته‌ی tzdata لحاظ شد.
از آنجا که زمان دقیق شروع قطع در داده‌ها مشخص نبود، یک آستانه‌ی زمانی ثابت به عنوان تقریب در نظر گرفته شد:

$$t_{\text{cut}} = 18:00 \quad (\text{به وقت تهران}).$$

برای هر اکانت (بر اساس username نرم‌الشده به حروف کوچک)، ویژگی‌های زیر از داده‌های پست‌ها استخراج شد:

- تعداد کل پست‌های مشاهده شده در بازه‌ی 15 تا 19 دی: .posts_total
- تعداد پست‌ها در هر روز: posts_dey_17، posts_dey_16، posts_dey_15، posts_dey_19 و posts_dey_18
- برای روز 17 دی، تعداد پست‌های قبل و بعد از آستانه‌ی قطع:

$$n_{17}^{\text{pre}}(u) = \text{posts_dey17_before_cut}, \quad n_{17}^{\text{post}}(u) = \text{posts_dey17_after_cut},$$

و سهم پساقطع:

$$s_{17}^{\text{post}}(u) = \frac{n_{17}^{\text{post}}(u)}{n_{17}^{\text{pre}}(u) + n_{17}^{\text{post}}(u)} = \text{share_dey17_after_cut}.$$

- مقایسه‌ی «قبل/بعد از روز قطع» با تجمعی روزهای 15 و 16 در برابر 18 و 19:

$$r(u) = \frac{n_{18}(u) + n_{19}(u)}{n_{15}(u) + n_{16}(u) + \varepsilon} = \text{post_pre_ratio}, \quad \varepsilon = 10^{-6}.$$

- میانگین و انحراف معیار ساعت انتشار به وقت تهران: hour_std و hour_mean .

قانون اولیه (برچسب‌های شبه‌واقعی). در گام نخست، یک قاعده‌ی ساده برای تولید برچسب‌های با اطمینان بالا تعریف شد. شرط حداقل شواهد در روز 17 دی به صورت

$$n_{17}^{\text{pre}}(u) + n_{17}^{\text{post}}(u) \geq 2$$

در نظر گرفته شد. سپس برچسب قاعده‌محور به شکل زیر تعیین گردید:

$$\hat{y}_{\text{rule}}(u) = \begin{cases} \text{داخل ایران} & n_{17}^{\text{post}}(u) = 0 \wedge n_{17}^{\text{pre}}(u) > 0, \\ \text{خارج ایران} & n_{17}^{\text{post}}(u) > 0, \\ \text{نامعلوم} & \text{در غیر این صورت.} \end{cases}$$

همچنین برای این قانون، یک امتیاز اطمینان قاعده‌محور تعریف شد:

$$c_{\text{in}}(u) = \min\left(1, \frac{n_{17}^{\text{pre}}(u)}{5}\right), \quad c_{\text{out}}(u) = \min\left(1, \frac{n_{17}^{\text{post}}(u)}{3}\right),$$

و در حالت نامعلوم مقدار ثابت 0.2 در نظر گرفته شد. با اعمال این قاعده، اندازه‌ی داده‌ی آموزشی شبه‌برچسب‌خورده (صرفًا شامل داخل/خارج) در Twitter برابر با $437 + 197 = 634$ ، در Telegram برابر با 371 و در Instagram برابر با $252 + 152 = 404$ نمونه به دست آمد.

مدل نهایی (یادگیری نیمه‌نظری). در ادامه، برای تعمیم برچسب‌گذاری به کاربرانی که شواهد مستقیم در روز 17 دی ندارند یا در ناحیه‌ی مرزی قرار می‌گیرند، یک مدل Logistic Regression بر روی همین برچسب‌های شبه‌واقعی آموزش داده شد. پیش از آموزش، ویژگی‌ها استانداردسازی شدند و مقدار random_state برابر 42 قرار گرفت. بردار ویژگی هر کاربر به صورت زیر تعریف

شد:

$$\mathbf{x}(u) = [\text{posts_total}, \text{posts_pre_15_16}, \text{posts_post_18_19}, \\ r(u), \text{posts_dey17_total}, n_{17}^{\text{post}}(u), s_{17}^{\text{post}}(u), \\ \text{hour_mean}, \text{hour_std}].$$

خروجی مدل دو احتمال $(\mathbf{x} | \text{داخل}(\mathbb{P}) \text{ و } \mathbf{x} | \text{خارج}(\mathbb{P}))$ را ارائه می‌دهد. با در نظر گرفتن آستانه‌ی تصمیم 0.60، برچسب نهایی به صورت زیر تعیین شد:

$$\hat{y}(u) = \begin{cases} \text{داخل ایران} & \max(p_{\text{in}}, p_{\text{out}}) \geq 0.60 \wedge p_{\text{in}} \geq p_{\text{out}}, \\ \text{خارج ایران} & \max(p_{\text{in}}, p_{\text{out}}) \geq 0.60 \wedge p_{\text{out}} > p_{\text{in}}, \\ \text{نامعلوم} & \max(p_{\text{in}}, p_{\text{out}}) < 0.60. \end{cases}$$

در نهایت، این برچسب‌ها برای 1000 اکانت برتر هر پلتفرم (بر اساس PageRank) گزارش شد.

۴.۶ نتایج برچسب‌گذاری برای 1000 اکانت برتر هر پلتفرم

جدول ۵۵ توزیع برچسب‌ها را برای 1000 اکانت برتر (با شناسه‌ی یکتا) در هر پلتفرم نشان می‌دهد. منظور از «نامعلوم» این است که یا اکانت در داده‌های پست‌ها شواهد کافی نداشته است (به‌ویژه به‌دلیل محدود بودن داده‌ها به 5000 پست پربازنشر در هر روز)، یا خروجی مدل پایین‌تر از آستانه‌ی اطمینان 0.60 قرار گرفته است.

جدول ۵۵: تعداد و درصد برچسب‌های نهایی برای 1000 اکانت برتر هر پلتفرم (پس از یکتا‌سازی شناسه‌ها).

پلتفرم	داخل ایران	خارج ایران	نامعلوم
Twitter	513 (51.3%)	366 (36.6%)	121 (12.1%)
Telegram	509 (50.9%)	287 (28.7%)	204 (20.4%)
Instagram	618 (61.8%)	293 (29.3%)	89 (8.9%)

نمونه‌ی اکانت‌های بالای PageRank. به منظور ارزیابی کیفی نتایج، جدول‌های ۵۶ تا ۵۸ ده اکانت نخست هر پلتفرم را (بر حسب PageRank) به همراه برچسب نهایی و میزان اطمینان مدل نمایش می‌دهند.

جدول ۵۶: ده اکانت نخست Twitter بر حسب PageRank به همراه برچسب نهایی.

اکانت	PageRank	برچسب	اطمینان
iranint1	0.011518	خارج ایران	1.000
tanhayektiger	0.009246	خارج ایران	0.999
spyssee	0.007945	داخل ایران	0.996
sharifizarchi	0.007821	خارج ایران	1.000
pahlavireza	0.007568	خارج ایران	1.000
rayzareei	0.007122	داخل ایران	0.956
yasamankhodajo	0.006873	خارج ایران	1.000
iranintlbrk	0.006869	خارج ایران	1.000
gohardaddy	0.006357	خارج ایران	1.000
ario1401	0.006318	داخل ایران	0.856

جدول ۵۷: ده اکانت نخست Telegram بر حسب PageRank به همراه برچسب نهایی.

اکانت	PageRank	برچسب	اطمینان
choonkehghashange	0.019180	داخل ایران	0.998
iribnews	0.011974	خارج ایران	1.000
akhbarmashhad	0.007917	خارج ایران	0.999
farsna	0.007510	خارج ایران	1.000
isna94	0.006714	خارج ایران	0.981
mehrnews	0.006483	داخل ایران	0.874
tasnimnews	0.005603	خارج ایران	1.000
irdamavand	0.005065	خارج ایران	0.988
irna_1313	0.004566	داخل ایران	0.961
iranintltv	0.004148	خارج ایران	1.000

جدول ۵۸: ده اکانت نخست Instagram بر حسب PageRank به همراه برچسب نهایی.

اکانت	PageRank	برچسب	اطمینان
avatoday1	0.008906	خارج ایران	1.000
iranintltv	0.007452	خارج ایران	1.000
officialrezapahalvi	0.005641	داخل ایران	0.999
gord_afarid_tv	0.003633	نامعلوم	0.000
avatodaypersian.official	0.003364	خارج ایران	1.000
indypersian	0.002849	خارج ایران	1.000
farzan43	0.002489	داخل ایران	0.992
farzam1213	0.002320	داخل ایران	0.999
par.e.simorgh	0.002195	داخل ایران	0.972
ali.javanmardi	0.001912	داخل ایران	0.989

نکته‌ی روش‌شناختی. برچسب‌گذاری ارائه شده یک تخمین رفتاری است و به دو عامل اصلی حساسیت دارد: (الف) زمان آستانه‌ی قطع t_{cut} که در اینجا به صورت ثابت 18:00 در نظر گرفته شد و امکان انجام تحلیل حساسیت نسبت به آن وجود دارد؛ (ب) این واقعیت که داده‌های پست‌ها صرفاً نمونه‌ای محدود از «پست‌های پربازنشر» هستند و لزوماً تمام فعالیت کاربران را پوشش نمی‌دهند. با وجود این محدودیت‌ها، ترکیب شواهد روز 17 دی با ویژگی‌های فعالیت در روزهای مجاور، معیاری عملی برای تفکیک «داخل/خارج/نامعلوم» در میان اکانت‌های اثرگذار فراهم می‌کند.

تحلیل رخدادهای بازتاب یافته در باره شهر یزد در پست‌های برتر شبکه‌های اجتماعی

۱.۷ داده و روش

در این بخش، خروجی‌های استخراج شده از سه پلتفرم Instagram، Twitter و Telegram (محدود به پست‌های یزد-محور در بازه ۱۵ تا ۱۹ دی ۱۴۰۴) بررسی شد. بر اساس نگاشت تقویمی، این بازه متناظر با ۵ تا ۹ ژانویه ۲۰۲۶ است. به منظور قابل مقایسه کردن میزان واکنش به پست‌ها، یک شاخص برهمنش (Engagement) تعریف شد که به صورت جمع مؤلفه‌های در دسترس هر پلتفرم محاسبه می‌شود:

$$E = L + C + F,$$

که در آن L تعداد like، C تعداد forward/retweet/repost و F comment/reply (بسته به پلتفرم) است. در مواردی که یکی از مؤلفه‌ها در داده موجود نبود، مقدار آن صفر در نظر گرفته شد.

برای استخراج «رخدادهای قابل گزارش»، ابتدا پست‌های یزد-محور هر روز بر اساس E مرتب شدند. سپس از خروجی خوشبندی رخدادها (فایل event_candidates_by_day.csv) و فایل‌های روزانه event_clusters_day*.csv برای شناسایی روایت‌های پر تکرار استفاده شد. نمایش متنی پست‌ها پس از نرمال‌سازی، با وزن دهنی TF-IDF و خوشبندی KMeans به چند خوشة موضوعی تبدیل شده است تا روایت‌های مشابه در قالب یک رخداد تجمعی شوند.

۲.۷ نمای کلی حجم پست‌های یزد-محور در خروجی «پست‌های برتر»

جدول ۵۹ نشان می‌دهد که در مجموعه «پست‌های برتر یزد-محور»، سهم Instagram از نظر تعداد بیشتر است؛ با این حال، بخش عمدات از محتوای رخدادمحور (اعتراض/درگیری/آتش‌سوزی) در Twitter و Telegram برجسته‌تر ظاهر می‌شود، به‌ویژه در روزهای پایانی بازه.

جمع	Twitter	Telegram	Instagram	روز
30	1	6	23	15
30	0	8	22	16
11	3	1	7	17
19	1	1	17	18
30	7	7	16	19

جدول ۵۹: تعداد پست‌های یزد-محور در مجموعه «پست‌های برتر» هر روز.

۳.۷ رخدادهای شاخصِ مرتبط با یزد در بازهٔ زمانی

در ادامه، رخدادهای شاخص استخراج شده از خوش‌های رخدادی و پست‌های برتر به تفکیک روز گزارش می‌شوند. در هر مورد، «شواهد درونداده‌ای» (تعداد و پراکندگی پست‌های هم‌روایت در خوش‌ها) و نیز «راستی آزمایی بیرونی» (وجود گزارش‌های عمومی همزمان، با تأکید بر اختلاف روایت‌ها) به صورت همزمان در نظر گرفته شده است.

15 دی 1404 : ادعای حمله با کوکتل به پایگاه بسیج در مردانه

در خروجی خوش‌بندی روز ۱۵، یک خوش‌ه با محوریت ادعای «پرتاب کوکتل/حمله موتورسواران به پایگاه بسیج مردانه یزد» مشاهده می‌شود (تعداد تکرار در خوش‌ه: حدود ۱۴ پست). این روایت عمدتاً در Telegram بازنثر شده و سپس بازتاب‌هایی در Twitter پیدا کرده است. از منظر راستی آزمایی، برای این رخداد گزارش‌های عمومی همزمان نیز دیده می‌شود، اما تمرکز آن عمدتاً در رسانه‌های مشخص و همسو با یک طیف سیاسی است؛ از این‌رو، این مورد به عنوان «رخداد ادعایی با تأیید مستقل محدود» ثبت می‌شود و از قطعی انگاری دربارهٔ جزئیات (میزان خسارت/تلفات/زمان دقیق) پرهیز خواهد شد.

16 دی 1404 : میزبانی مسابقهٔ فوتbal در یزد

در روز ۱۶، خوش‌ه غالب رخدادی به خبر «برگزاری/میزبانی مسابقهٔ فوتbal مرتبط با استقلال و تراکتور در ورزشگاه شهید نصیری یزد» مربوط است (خوش‌ه ورزشی با چند پست هم‌روایت، عمدتاً در Telegram). این مورد در منابع خبری عمومی نیز گزارش شده است؛ بنابراین، به عنوان رخداد «تأییدشده» در نظر گرفته می‌شود. از نظر دینامیک انتشار، کانال‌ها و صفحات هواداری و خبری Telegram نقش اصلی را در انتشار اولیه داشته‌اند و سپس بخشی از محتوا در Instagram بازنشر شده است.

17 دی 1404 اعتصاب/تعطیلی کسبه و بازار در یزد (روایت متعارض)

در روز 17، خوشهای با کلیدوازه‌های «اعتصاب/بازار/کسبه/یزد» استخراج شد (حدود 5 پست هم روایت در خوشه) و در پست‌های برتر نیز چند نمونه با همین مضمون مشاهده می‌شود. با این حال، راستی‌آزمایی بیرونی با وضعیت «روایت متعارض» همراه است: بخشی از منابع بیرونی تعطیلی/اعتصاب را در قالب موج گسترده‌تر گزارش می‌کنند و بخشی دیگر بر جداسازی «صف بازاریان» از «اغتشاشگران» یا تکذیب تعمیم روایت تأکید دارند. ازین‌رو، این رخداد به صورت «گزارش شده در شبکه‌های اجتماعی با شواهد بیرونی غیرهمسو» ثبت می‌شود و در تحلیل شبکه‌ای، به عنوان نمونه‌ای از دوگانه تقویت روایت توسط بازنشر در برابر مدیریت روایت توسط تکذیب/چارچوب‌بندی رسمی در نظر گرفته می‌شود.

18 دی 1404 : بازتاب ویدیویی از کنش اعتراضی در یزد (صفاییه/بلوار صدوqi)

در روز 18، اگرچه به دلیل محدودیت تعداد پست‌های یزد-محور در داده، حجم «پست‌های برتر» کمتر از 30 است، اما یک روایت رخدادمحور با موقعیت مکانی مشخص بر جسته می‌شود: اشاره به صفاییه و نیز «بلوار صدوqi» و «پایگاه بزرگ بسیج». این رخداد در قالب ویدیو/گزارش در منابع عمومی نیز بازتاب یافته است؛ بنابراین از نظر بیرونی امکان تطبیق کلی وجود دارد. در داده‌های موجود نیز این محتوا عمدتاً از Twitter به عنوان مسیر خبررسانی سریع وارد شده و سپس در Instagram به صورت بازنشر صفحه‌های خبری مشاهده می‌شود.

19 دی 1404 (مطابق 9 ژانویه 2026): اوج گیری بازتاب اعتراضات و رخدادهای امنیتی (آتش‌سوزی شهرداری/مقابل استانداری)

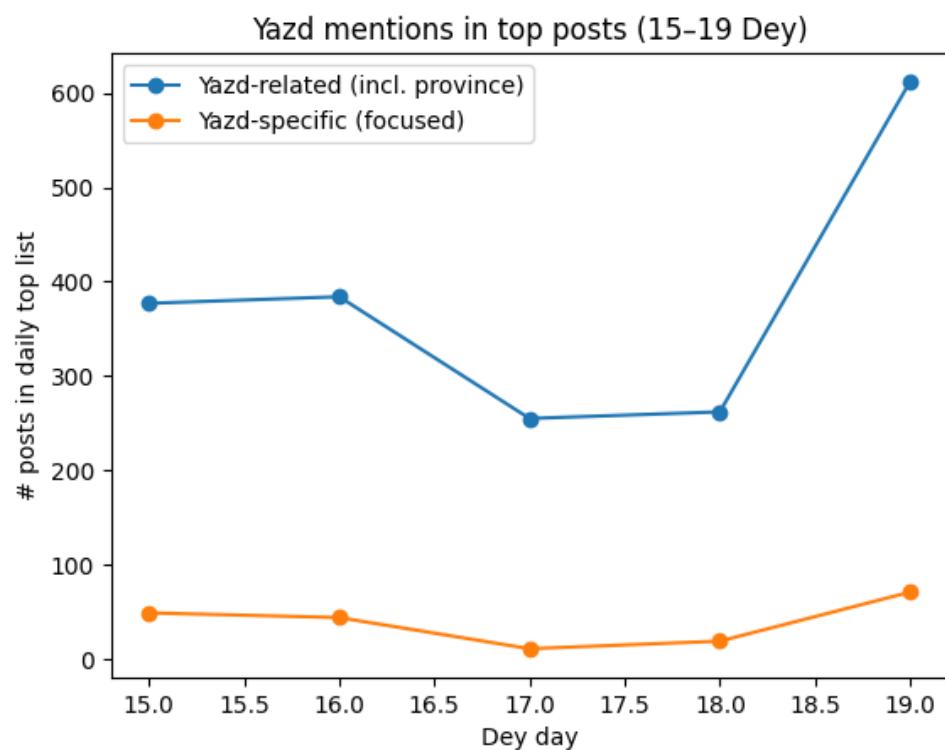
بالاترین تراکم محتوای رخدادمحور در داده‌های یزد-محور مربوط به روز 19 است. در «پست‌های برتر» این روز، تعداد پست‌های مرتبط با اعتراض/شعار/حضور خیابانی به حدود 15 پست و تعداد پست‌های مرتبط با آتش‌سوزی/تخريب/درگیری به حدود 6 پست می‌رسد (بر اساس دسته‌بندی محتوایی انجام شده روی متن نرمال شده). در خروجی خوش‌بندی نیز چند خوش‌مشخص مشاهده می‌شود که به مکان‌هایی مانند «بلوار پاکنژاد» و «میدان/خیابان امام حسین» اشاره دارند و در کنار آن، روایت «آتش گرفتن ساختمان شهرداری یزد» پررنگ است.

از نظر راستی‌آزمایی بیرونی، درباره آتش‌سوزی ساختمان شهرداری یزد گزارش‌های عمومی هم‌زمان وجود دارد، اما در عامل رخداد و چارچوب روایت اختلاف مشاهده می‌شود: برخی منابع آن را به رخدادهای اعتراضی نسبت می‌دهند و برخی دیگر با ادبیات امنیتی/اغتشاش، بر تخریب و خسارت تأکید می‌کنند. علاوه بر آن، گزارش‌هایی از پیگیری انتظامی و بازداشت افراد مرتبط

با آتش‌سوزی/راه‌بندان/اخلال در نظم عمومی در یزد منتشر شده است. در این تحلیل، همزمانی این سه مؤلفه (حجم بالای پست‌های رخدادمحور، اشاره به نقاط مکانی داخل شهر، و وجود گزارش‌های عمومی درباره آتش‌سوزی و برخورد انتظامی) موجب می‌شود روز ۱۹ به عنوان «روز کانونی رخداد» در بازه انتخاب شود.

٤.٧ الگوی انتشار میان‌پلتفرمی و نکات تحلیلی

- غلبۀ عددی Instagram در برابر غلبۀ رخدادمحور Telegram/Twitter: هرچند در جدول تعداد پست‌های Instagram بالاتر است، اما بخش مهمی از آن‌ها محتوای عمومی/غیررخدادی (گردشگری، صفحه‌های محلی، هشتگ‌های شهری و ...) را شامل می‌شود. در مقابل، روایت‌های رخدادی دارای مؤلفه‌های مکان/زمان مشخص بیشتر از مسیر Twitter و Telegram وارد می‌شوند و سپس در Instagram بازنثر می‌گردند.
- تشدید هم‌پوشانی روایت‌ها در انتهای بازه: در روزهای ۱۸ و بهویژه ۱۹، چند روایت به‌طور هم‌زمان در یک قاب مشترک قرار می‌گیرند (حضور خیابانی، شعارها، آتش‌سوزی و برخورد انتظامی). این هم‌پوشانی احتمال دیده‌شدن رخداد در «پست‌های برتر» را افزایش می‌دهد و به شکل‌گیری خوش‌های متین واضح‌تر منجر می‌شود.
- عدم قطعیت در رخدادهای امنیتی کم‌گزارش: رخدادهایی مانند «حمله کوکتل به پایگاه بسیج مردانه» در داده بازتاب دارند، اما سطح تأیید مستقل آن‌ها نسبت به رخدادهای پرگزارش‌تر (مانند آتش‌سوزی شهرداری) پایین‌تر است؛ بنابراین در نگارش نهایی، این موارد به صورت ادعاهای شبکه‌های اجتماعی با ذکر سطح اطمینان گزارش می‌شوند.



شکل ۱۰: جایگاه زمانی رخدادهای یزد در بازه ۱۵ تا ۱۹ دی ۱۴۰۴ بر اساس خوشه‌های رخدادی.

راستی آزمایی و تحلیل انجمن شایعات

۱.۸ صورت مسئله و داده‌های مورد استفاده

طبق صورت سؤال، لازم است دست کم ۵ مورد «خبر نادرست/اطلاعات غلط» پس از رخداد ۱۷ دی (با اتکا به منابع خبری یا منابع موثق) شناسایی شود و سپس درباره عمدی/سهوی بودن انتشار آنها تحلیل ارائه گردد. علاوه بر این، با تکیه بر نتایج «تحلیل انجمن»، انجمن‌ها و اکانت‌هایی که به احتمال زیاد در گردش این اخبار نقش داشته‌اند استخراج و تفسیر شوند.
برای این منظور از دو دسته داده استفاده شد:

- داده‌های «پست‌های پُرتعامل» در بازه روزهای دی (با تمرکز بر بازه ۱۵ تا ۱۹ دی) در سه پلتفرم Instagram، Twitter و Telegram.
- گراف بازنشر هر پلتفرم (فایل GDF) که در آن جهت یال‌ها از «اکانت بازنشرکننده» به «اکانت مرجع بازنشر» است و وزن یال تعداد دفعات بازنشر را نشان می‌دهد.

۲.۸ روش تحلیل شبکه: انجمن‌یابی و سنجش اثرگذاری

برای پیوند دادن «اخبار/شایعات منتخب» به ساختار شبکه، دو محاسبه انجام شد:

۱. انجمن‌یابی با **Louvain**: روی نسخه بی‌جهت و وزن‌دار گراف (رویه رایج برای بیشینه‌سازی مازولاریتی)، انجمن‌ها با روش Louvain استخراج شد. ایده اصلی این روش بیشینه‌سازی مازولاریتی است:

$$Q = \frac{1}{2m} \sum_{i,j} \left(A_{ij} - \frac{k_i k_j}{2m} \right) \delta(c_i, c_j),$$

که در آن A_{ij} وزن یال بین i و j ، k_i مجموع وزن یال‌های متصل به i ، m مجموع وزن کل یال‌ها و $\delta(c_i, c_j)$ نشانگر «هم‌انجمن بودن» دو رأس است.

۲. برای سنجش اثرگذاری: روی نسخه جهت‌دار و وزن‌دار گراف، PageRank محاسبه شد تا اثرگذاری هر حساب (به عنوان «مرجع بازنشر» که از دیگران ورودی می‌گیرد)

تقریب زده شود. فرم استاندارد وزن دار:

$$PR(v) = \frac{1-d}{N} + d \sum_{u \in In(v)} \frac{w_{uv}}{\sum_x w_{ux}} PR(u),$$

که در آن d ضریب میرایی، N تعداد گرهها، $In(v)$ همسایه‌های ورودی v و w_{uv} وزن یال $v \rightarrow u$ است.

برای پرهیز از خلط «انتشار شایعه» با «بازنشر تکذیبیه»، در مرحله تفسیر کیفی متن پست‌ها، نشانه‌های واژگانی مانند «تکذیب شد»، «صحت ندارد»، «جعلی است»، «نادرست» به عنوان سرخ به کار رفت تا اکانت‌ها به صورت عملیاتی در دو نقش تقویت شایعه و اصلاح/ختن‌سازی تفکیک شوند.

۳.۸ پنج مورد خبر/ادعای نادرست منتخب و راستی‌آزمایی

در جدول ۶۰ پنج ادعای منتخب، حضور آنها در پلتفرم‌ها، نتیجه راستی‌آزمایی و برآورد کیفی از «سهوی/عمدی بودن» انتشار خلاصه شده است (جزئیات هر مورد پس از جدول ارائه شده است).

کلید ادعا	پلتفرم غالب	نتیجه راستی آزمایی	تحلیل نوع انتشار (سهوی/عمدی)
khanneni_escape_mehrabad	Telegram/Instagram	نادرست (ویدیو/روایت «آماده باش مهرآباد برای فرار به مسکو» تأیید نشد)	با توجه به اتكای روایت به «ویدیو به ظاهر مستند»، الگوی رایج گمراه‌سازی برای برانگیختن هیجان و ناطمینانی مشاهده می‌شود؛ در عین حال، بازنشر ثانویه می‌تواند سهواً و ناشی از اعتماد به شواهد بصری باشد.
police_joined_protesters	Twitter/Telegram	نادرست (ویدیو مربوط به رخداد دیگری بوده است)	بازنشر «ویدیوی قدیمی با زمینه‌زدایی» معمولاً می‌تواند نشانگان گمراه‌سازی باشد؛ با این حال، در چرخه بازنشر ثانویه در شبکه‌ها، سوءبرداشت و انتشار سهوی نیز محتمل است.
sarpol_two_killed	Telegram	شایعه تکذیب شد	این مورد در داده مشاهده شده عمدتاً به شکل «تکذیبیه» بازتاب یافته است؛ منشأ اولیه می‌تواند عمدی باشد، اما شواهد درون داده‌ای موجود بیشتر ناظر به گردش تکذیب است.
turkey_evacuating	(Instagram/Twitter) مهراه با	ادعای «تخليه اتباع تركيه» تکذيب شد	ادعاهای «خروج اتباع/تخليه» در وضعیت بحران ظرفیت بالایی برای عملیات روانی و القای نامنی دارند (احتمال عمدی بودن منشأ)؛ در عین حال، انتشار گسترده تکذیب توسط رسانه‌ها چرخه اصلاح را نیز فعال کرده است.
us_tanker_over_iran	Instagram/Telegram	ادعای «ورود/پرواز سوخت‌رسان آمریکا در آسمان ایران» نادرست گزارش شد	این ادعا معمولاً با تصاویر/اسکرین‌شات‌های مبهم از رهگیرهای پرواز منتشر می‌شود؛ می‌تواند ترکیبی از سوءتعییر داده‌های ردیابی و تقویت عمدی روایت برای القای نزدیک بودن درگیری باشد.

جدول ۶۰: پنج ادعای نادرست منتخب پس از رخداد ۱۷ دی و تحلیل کیفی نوع انتشار آن‌ها.

۱.۳.۸ مورد ۱: «آماده باش اضطراری مهرآباد برای فرار خامنه‌ای به مسکو»

نمونه‌های پر تکرار/پر تعامل:

● پست Instagram sangar_newss با درگیری 232864 مشاهده «دو هواپیما indypersian در بخش نظامی مهرآباد»؛ همچنین saba.majlessi با 62769 و 33174.

● در انجمن C₂₄₃₇ چند کanal با مضامین مشابه فعال بودند؛ از جمله 2597 پست در این کلید ادعا با تعامل (جمعاً rodast_omiddana).

راستی آزمایی: گزارش‌های درستی سنجی منتشر شده نشان می‌دهد ویدیوی منتبه به مهرآباد (و نتیجه‌گیری «پرواز اضطراری برای فرار») با شواهد مکانی/بصری هم خوان نیست و روایت مبتنی بر آن تأیید نمی‌شود.
چگونه می‌شد همان لحظه تردید کرد؟

۱. توقف (Stop): ادعا به طور مستقیم احساسات (ترس/هیجان) را هدف می‌گیرد؛ بنابراین، واکنش اولیه باید توقف و پرهیز از بازنشر باشد.

۲. بررسی منبع (Investigate the Source): توجه به سابقه صفحه، سبک محتوا و این‌که آیا پیش‌تر سابقه انتشار اطلاعات تأییدنشده داشته است یا نه.

۳. یافتن پوشش بهتر (Find Better Coverage): جست‌وجوی هم‌زمان در چند منبع معتبر خبری/راستی‌آزما و پرهیز از اتکا به یک ویدیو به عنوان تنها شاهد.

۴. ردگیری به منشأ (Trace): بررسی محل انتشار اولیه ویدیو، وجود نسخه‌های قدیمی‌تر و نیز نشانه‌های ساخت/دست‌کاری (ناهمخوانی پرسپکتیو، صداگذاری نامعمول و ...).

۲.۳.۸ مورد ۲: «پیوستن نیروی انتظامی به معترضان»

نمونه‌ها:

● پست Twitter limondar0 با تعامل 3170 در انجمن C₄₀₀₂ و چند حساب دیگر با بازنشر محدودتر.

● Telegram تمرکز اصلی در انجمن C₂₄₃₇ (تجمیع ۵ پست) و حضور حساب‌هایی که «توضیح/بازنشر» انجام داده‌اند.

• Instagram: حضور factnameh به عنوان منبع اصلاحی (نقش خنثی‌سازی).

راستی آزمایی: ویدیویی مورد استناد مربوط به زمان و زمینه دیگری بوده و با زمینه‌زدایی به رخداد دی نسبت داده شده است؛ علاوه بر این، سرنخ‌های تصویری/متنی داخل ویدیو با روایت «پیوستن به معتبرضان» سازگار نیست.

نشانه‌های فوری برای تردید:

• نبود مکان/زمان دقیق و اتکا به «شنیده‌ها».

• امکان اجرای سریع reverse image/video search روی فریم‌های کلیدی.

• ناسازگاری پوشش/آب‌وهوا و شعارهای داخل تصویر با ادعای مطرح شده.

۳.۳.۸ مورد ۳: «کشته شدن دو نفر در سرپل ذهاب»

در داده‌های مشاهده شده، این کلید ادعا عمدتاً از مسیر «تکذیبیه» بازتاب یافته است:

• Telegram: پست 19 kermanshah1404 با تعامل 226 در انجمن C₅₄₈₇ با تعامل 25 در انجمن C₁₅₈₂₈ که هر دو صراحتاً «تکذیب» را بازنثر کرده‌اند.

تفسیر: این الگو نشان می‌دهد شایعه احتمالاً بیرون از زیرنمونه‌های «پرتعامل» نیز مطرح بوده است، اما آنچه در خروجی حاضر برجسته شده، کanal‌های خبری/ محلی با نقش غالب اصلاح است. از منظر «نوع انتشار»، منشأ اولیه می‌تواند عمدى باشد، ولی در داده مشاهده شده نقش غالب گردش تکذیبیه است.

۴.۳.۸ مورد ۴: «تخلیه اتباع ترکیه از ایران»

نمونه‌ها:

• Telegram: بیشترین تمرکز در انجمن C₁₅₈₂₈ (تجمیع ۵ پست) با حضور کanal‌های خبری مانند entekhab_ir (معامل 212)، isna94 (معامل 94 و PageRank ۹۴) و yjcnewschannel (معامل ۸۷).

• Twitter: نمونه‌هایی از بازنثر ادعا (بدون اتکا به منبع رسمی) مشاهده شد.

• Instagram: چند صفحه با روایت‌های «خروج فوری اتباع خارجی» بدون سند قابل اتکا.

راستی آزمایی: پیگیری رسمی از سفارت ترکیه در تهران، روایت «تخلیه» را تکذیب کرده است. بر این اساس، هر روایت قطعی درباره «شروع خروج اتباع» بدون ارائه سند رسمی باید مشکوک تلقی شود.

چراً احتمال عمدی بودن منشأ: در بحران‌ها، روایت‌هایی از جنس «خروج اتباع/تخلیه سفارت» به طور کلاسیک برای القای نزدیک بودن فروپاشی یا تشدید ناامنی استفاده می‌شوند؛ با این حال، در این مورد حضور پرنگ کانال‌های خبری اصلاحی نشان می‌دهد مکانیسم تصحیح نیز فعال بوده است.

۵.۳.۸ مورد ۵: «ورود/پرواز هوایی سوخت‌رسان آمریکا به آسمان ایران»

نمونه‌ها:

● Telegram: الگوی بسیار مت مرکز در انجمن C₂₄₃₇ با ۱۵ پست و تعامل تجمیعی ۱۸۸۳۰؛ حساب rodast_omiddana به تنایی ۱۲ پست مرتبط و تعامل ۱۸۳۳۵ داشته است. همچنین bisimchimedia با تعامل ۲۶۳۴ در انجمن C₁₅₈₂₈ مشاهده شد.

● Instagram: پست project_leon3 با تعامل ۶۰۱۰۶ (نمایش «اسکرین‌شات/روایت فوری»).

راستی آزمایی: بر اساس گزارش‌های منتشر شده، ادعای ورود هوایی سوخت‌رسان به خاک/آسمان ایران «کذب» اعلام شده و مسیر پرواز ادعایی با داده‌های ردیابی عمومی هم‌خوان گزارش نشده است.

چگونه می‌شد تردید کرد؟

- اسکرین‌شات‌های رهگیری پرواز معمولاً قابل برش/دست‌کاری هستند؛ ارجاع به صفحه اصلی رهگیر و شناسه پرواز برای اعتبارسنجی ضروری است.
- نبود شناسه پرواز، زمان دقیق و منبع مستقل نشانه هشدار است.
- بررسی پوشش چند منبعی و سپس «ردگیری به منشأ» (این‌که نخستین بار کدام کانال/صفحه ادعا را ساخته است) اهمیت دارد.

۴.۸ تحلیل انجمن‌ها: کدام خوش‌ها در چند شایعه تکرار شده‌اند؟

برای نزدیک شدن به مفهوم «سازمان‌یافتنگی»، از یک اکتشافی ساده استفاده شد: شمارش تعداد «کلید ادعای متمایز» که هر انجمن در آن ظاهر شده و نیز تعداد کل پست‌های هم‌خوان با الگوها. نتایج نشان داد:

- انجمن C₁₅₈₂₈ در هر 5 کلید ادعا حضور دارد (جمعاً 18 پست هم خوان). این انجمن ترکیبی از کانال‌های خبری/بازنشرکننده را شامل می‌شود؛ در برخی ادعاهای نقش غالب آن «تکذیب/اصلاح» (مانند ادعای تخلیه اتباع ترکیه) و در برخی دیگر «بازنشر روایت‌های هیجانی» است.
 - انجمن C₂₄₃₇ در 3 کلید ادعا حضور دارد، اما از نظر «حجم پست» پررنگ‌تر است (جمعاً 25 پست). مرکز بالا روی ادعای «هوایپمای سوخت‌رسان» و نیز حضور در دو ادعای دیگر نشان می‌دهد این خوشه می‌تواند نقش «هاب بازنشر» برای روایت‌های بحران‌محور داشته باشد.
 - انجمن C₁₁₂₄₆ در 2 کلید ادعا دیده می‌شود (جمعاً 6 پست) و در مواردی مانند بازنشر روایت‌های مبنی بر «منابع ثانویه» (برای نمونه نقل قول‌های رسانه‌ای) فعال بوده است.
- تعابیر احتیاطی:** تکرار یک انجمن در چند ادعا به‌نهایی اثبات هماهنگی نیست؛ با این حال، زمانی که این تکرار با (الف) تراکم پست بالا، (ب) چند حساب پر تکرار و (ج) اتکا به قالب‌های محتوا‌ای مشابه (اسکرین‌شات‌های فوری، ویدیوهای زمینه‌زدایی شده، تیترهای هیجانی) همراه شود، می‌تواند نشانه‌ای از «زنگیره‌های بازنشر پایدار» باشد که در وضعیت بحران چند شایعه را به‌طور هم‌زمان تقویت می‌کنند.
- ## ۵.۸ جمع‌بندی تحلیلی
- با کنار هم گذاشتن راستی‌آزمایی و تحلیل انجمن‌ها، نتایج زیر به‌دست آمد:
۱. بخش قابل توجهی از شایعات پس از رخداد 17 دی با محتوا‌ی بصری/شبیه مستند (ویدیو/اسکرین‌شات رهگیر پرواز) منتشر شده‌اند؛ این نوع محتوا به‌دلیل اقناع‌پذیری بالا، هم برای گمراه‌سازی عمدی مناسب است و هم زمینه بازنشر سهولی را تقویت می‌کند.
 ۲. در سطح شبکه‌ای، دو انجمن C₂₄₃₇ و C₁₅₈₂₈ «چندادعایی» بودن را نشان می‌دهند؛ اولی بیشتر نقش هاب بازنشر پُر حجم (به‌خصوص در ادعای سوخت‌رسان) و دومی نقش ترکیبی بازنشر خبر/تکذیبیه (به‌خصوص در ادعای تخلیه اتباع ترکیه).
 ۳. حضور منابع راستی‌آزمایی مانند factnameh و نیز کانال‌های خبری رسمی در برخی انجمن‌ها نشان می‌دهد «اکوسیستم اصلاح» نیز فعال بوده است؛ بنابراین، تحلیل نقش اکانت‌ها باید دو وجه «تقویت شایعه» و «اصلاح/تکذیب» را هم‌زمان در نظر بگیرد.

۴. از منظر سواد رسانه‌ای، اجرای نظاممند چهار گام SIFT (توقف، بررسی منبع، پوشش بهتر و ردگیری منشأ) به همراه ابزارهایی مانند reverse image search می‌تواند احتمال بازنثر سه‌های را به‌طور معناداری کاهش دهد.

تعادل ساختاری و تحلیل خوشه‌ها در شبکه بازنشر

۱.۹ تعریف شبکه و مسئله

شبکه‌های ارائه شده در قالب GDF، شبکه بازنشر جهت‌دار و وزن‌دار هستند؛ جهت هر یال از «اکانت بازنشرکننده/ارجاع‌دهنده» به «اکانت مرجع بازنشر» ثبت شده و وزن یال، تعداد دفعات بازنشر از ارجاع‌دهنده به مرجع را نشان می‌دهد. در پلتفرم Instagram به دلیل نامشخص بودن مسیر بازنشر، یال‌ها به صورت «ارجاع اکانت‌های بعدی به اولین منتشرکننده پست» ساخته شده‌اند؛ بنابراین بخشی از روابط در این پلتفرم بیش از آن‌که بازتاب مسیر واقعی انتشار باشد، ناظر به «ارجاع به منبع اولیه» است.

هدف این سؤال، انتساب علامت مثبت/منفی به اکانت‌ها بر اساس موضع نسبت به جمهوری اسلامی (موافق/مخالف)، ساخت یک گراف علامت‌دار (Signed Graph)، سنجش میزان تعادل ساختاری، و درنهایت تحلیل خوشه‌ها (به‌ویژه زیرخوشه‌های مخالفان) و ارائه یک تصویر احتمالی از الگوی روابط آینده میان آن‌ها است.

۲.۹ برچسب‌گذاری موضع و تولید گروه‌های تحلیل

برای کنترل پیچیدگی محاسبات و تمرکز بر بخش اثرگذار شبکه، در هر پلتفرم 8000 نод منتخب (بر اساس شاخص اهمیت/فیلدهای مشابه در داده) وارد پایپ‌لاین شد. سپس مراحل زیر اجرا شد:

۱. **برچسب اولیه (Seed):** با تکیه بر محتوای مطالب پر تکرار و مجموعه‌ای از کلیدواژه‌ها/عبارات شاخص موافق و مخالف، برای بخشی از نودها برچسب اولیه $+1$ (موافق) یا -1 (مخالف) تولید شد و برای سایر نودها مقدار 0 (نامعلوم) در نظر گرفته شد.

۲. **گسترش برچسب:** برای افزایش پوشش برچسب‌گذاری، یک فرآیند انتشار برچسب از خانواده Label Propagation روی گراف وزن‌دار (با استفاده از همسایگی و وزن یال‌ها) اجرا شد و برای هر نod یک نمره پیوسته در بازه $[1, -1]$ به دست آمد.

۳. **سه‌حالت‌سازی موضع:** نودهایی که نمره نهایی آن‌ها در بازه $[0.2, -0.2]$ قرار داشتند

به عنوان «نامعلوم/خنثی» نگه داشته شدند و سایر نودها به صورت دودویی به «موافق» یا «مخالف» نگاشت شدند.

۴. زیرخوشه‌بندی مخالفان: برای تحلیل دقیق‌تر جریان‌های درون مخالفان، زیرگراف القایی نودهای مخالف ساخته شد و با الگوریتم Louvain به ۶ خوشه مخالف (به علاوه یک بخش «متفرقه/ایزوله» برای نودهای کم اتصال) تقسیم شد. گروه «موافقان» به عنوان خوشه ۰ تعریف شد.

پلتفرم	کل نودهای تحلیل	Seed	موافق	مخالف	نامعلوم	موافق نهایی	نامعلوم	(3.2%)	256
Twitter	8000	(11.4%)	154	(1.9%)	5131	(64.1%)	2613	(32.7%)	(3.2%)
Telegram	8000	(7.0%)	147	(1.8%)	4736	(59.2%)	3046	(38.1%)	(2.7%)
Instagram	8000	(8.3%)	81	(1.0%)	1726	(21.6%)	6151	(76.9%)	(1.5%)

جدول ۶۱: خلاصه برچسب‌گذاری موضع (Seed و خروجی نهایی) برای هر پلتفرم.

۳.۹ ساخت گراف علامت‌دار و تعریف تعادل ساختاری

پس از تعیین گروه‌ها، به هر رابطه (در سطح گروهی) علامت نسبت داده شد. تعریف عملیاتی به این صورت بود که روابط درون‌گروهی «ثبت» و روابط بین‌گروهی «منفی» در نظر گرفته شد. سپس تعادل ساختاری بر اساس مثلث‌ها سنجیده شد. برای یک سه‌تایی از نودها $\{u, v, w\}$ ، اگر یال‌های بین آن‌ها موجود باشد، علامت مثلث با ضرب علائم یال‌ها تعریف می‌شود:

$$\Delta(u, v, w) = s_{uv} s_{vw} s_{wu}, \quad s_{ab} \in \{+1, -1\}.$$

یک مثلث «متعادل» است اگر $\Delta(u, v, w) = +1$ باشد (یعنی الگوهای $+++$ یا $-+-$ رخ دهد) و «نامتعادل» است اگر $\Delta(u, v, w) = -1$ باشد. به دلیل بزرگی شبکه، شمارش روی مجموعه‌ای از مثلث‌های استخراج شده/نمونه برداری شده از زیرگراف نودهای برچسب‌خورده انجام شد (نودهای «نامعلوم» در این مرحله کنار گذاشته شدند).

پلتفرم	تعداد مثلث‌های بررسی شده	متعادل	نامتعادل	نسبت تعادل
Twitter	2442	2342	100	0.959
Telegram	7020	6796	224	0.968
Instagram	4291	4051	240	0.944

جدول ۶۲: نتیجه تعادل ساختاری (بر پایه مثلث‌ها) در هر پلتفرم.

نسبت‌های جدول ۶۲ نشان می‌دهد که در هر سه پلتفرم، الگوی تعاملات (پس از امضاسازی) به‌طور محسوسی به سمت وضعیت‌های سازگار با تعادل ساختاری میل دارد. از منظر تفسیر اجتماعی، چنین الگویی معمولاً با «قطبی شدن» و شکل‌گیری بلوک‌های نسبتاً منسجم درون‌گروهی همراه است، به‌ویژه در حالتی که روابط بین بلوک‌ها عمدتاً منفی تلقی شوند.

۴.۹ تحلیل زیرخوشهای مخالفان و شدت تعامل بین گروه‌ها

برای مخالفان، 6 خوše اصلی استخراج شد و باقی نودهای کماتصال/ایزوله در گروه «متفرقه» نگهداری شدند (این گروه در نتایج خوشبندی مخالفان با شناسه 1 – مشخص است). اندازه هر زیرخوشه در جدول ۶۳ گزارش شده است.

F6	F5	F4	F3	F2	F1	متفرقه	کل مخالف	پلتفرم
172	519	667	1095	1096	1129	453	5131	Twitter
207	255	317	594	692	1923	748	4736	Telegram
104	112	133	186	220	237	734	1726	Instagram

جدول ۶۳: اندازه زیرخوشهای مخالفان (F6 تا F1) و گروه متفرقه/ایزوله در هر پلتفرم.

برای سنجش «نzdیکی» یا «درهم‌تنیدگی» زیرخوشهای، ماتریس شدت تعامل بین گروه‌ها محاسبه شد. در این ماتریس، مقدار خانه (j, i) برابر مجموع وزن یال‌های بازنشر بین نودهای گروه i و گروه j است (برای خوانایی به صورت تجمعی دوسویه گزارش شد). گروه 0 نماینده «موافقان» و گروه‌های 1 تا 6 نماینده زیرخوشهای «مخالفان» هستند. به عنوان شاخص تکمیلی، نسبت وزن درون‌گروهی به کل وزن بین‌گروهی نیز سنجیده شد؛ در Telegram و Instagram سهم درون‌گروهی بزرگتر از Twitter است که با انسجام بیشتر خوشهای سازگار است. از میان تعامل‌های بروزنگروهی، سه پیوند قوی‌تر (بر حسب وزن) در هر پلتفرم در جدول ۶۴ گزارش شده است.

جدول ۶۴: سه تعامل برونگروهی قوی‌تر (بر حسب وزن بازنشر) در هر پلتفرم.

پلتفرم	قوی‌ترین تعامل‌ها	وزن	تفسیر عملیاتی
Twitter	(F4, F3)	7398	هم‌پوشانی پیام/بازنشر قابل توجه بین دو زیرخوشه مخالف
	(F3, F2)	5984	نزدیکی گفتمانی/انتشاری بین دو زیرخوشه مخالف
Telegram	(F4, F2)	3828	F2, F3, F4 پیوند سه‌گانه پرقدرت بین دو زیرخوشه مخالف
	(F4, F3)	6657	اتصال پرقدرت بین دو زیرخوشه مخالف
Instagram	(F3, F2)	3762	تعامل بالا بین زیرخوشه‌های مخالف
	(F4, 0)	2992	جريان بازنشر بین موافقان و یک زیرخوشه مخالف
	(F3, 0)	299	تعامل بین موافقان و یک زیرخوشه مخالف
	(F3, F1)	279	نزدیکی نسبی دو زیرخوشه مخالف
	(F5, F1)	180	تعامل متوسط بین دو زیرخوشه مخالف

۵.۹ پیش‌بینی کیفی روابط آینده میان خوشه‌ها و زیرخوشه‌ها

با توجه به نسبت تعادل بالا در جدول ۶۲، انتظار می‌رود ساختار کلی «دو قطب اصلی» (موافق/مخالف) در کوتاه‌مدت پایدار بماند؛ زیرا سهم مثبت‌های نامتعادل اندک است و شبکه از منظر تعادل، فشار محدودی برای بازارایی ریشه‌ای نشان می‌دهد. در سطح خردتر مخالفان نیز چند الگوی قابل اتکا مشاهده می‌شود:

- هسته هم‌بازنشر میان مخالفان: در Twitter و Telegram، سه زیرخوشه F2, F3, F4 بیشترین وزن تعامل برونگروهی را با یکدیگر دارند (جدول ۶۴). این الگو با نوعی همگرایی تاکتیکی یا ائتلاف مؤقت سازگار است: حتی با وجود تفاوت‌های هویتی/سیاسی، اشتراک هدف (مخالفت با حکومت) می‌تواند سطح بالای اشتراک پیام و بازنشر میان این زیرخوشه‌ها را توضیح دهد. بنابراین به صورت کیفی محتمل است در رخدادهای مشابه آینده نیز این زیرخوشه‌ها سریع‌تر به محور مشترک پیام‌رسانی برسند و در عمل «هم‌پیمان‌انتشاری» باقی بمانند.

- زیرخوشه‌های کم اختلاط: در Twitter، زیرخوشه F5 عمدهاً وزن درونگروهی بالایی دارد و تعامل برونگروهی آن با سایر زیرخوشه‌ها محدود‌تر است. چنین الگویی معمولاً به «گفتمان متمایز» یا «منابع بازنشر متفاوت» اشاره دارد و می‌تواند نشانه استقلال نسبی

این جریان باشد؛ از این‌رو، انتظار می‌رود در آینده نیز به عنوان یک بلوک نسبتاً مستقل باقی بماند و صرفاً در نقاط اشتراک بزرگ با سایر مخالفان هم راستا شود.

● **تعامل موافقان با بخشی از مخالفان:** وجود وزن قابل توجه بین گروه ۰ (موافقان) و برخی زیرخوشهای مخالف (به خصوص در Telegram) لزوماً به معنای رابطه مثبت نیست؛ زیرا در بازنیش، امکان «انتشار برای نقد/پاسخ» یا «نظرارت و دیده‌بانی» وجود دارد. با این حال، همین پیوندها نشان می‌دهد برخی زیرخوشهای این‌گروهها در مرکز توجه/تنش قرار دارند و در آینده نیز احتمالاً کانون اصلی رویارویی روایت‌ها خواهند بود.

● **احتیاط در Instagram:** سهم بالای گروه متفرقه در مخالفان و ساختار خاص تولید یال‌ها در Instagram باعث می‌شود زیرخوشهای بندی و شدت تعامل بین گروه‌ها حساس‌تر باشد. در این پلتفرم، پیش‌بینی روابط آینده بهتر است با بررسی محتوایی اکانت‌های کانونی هر زیرخوشه (برای نمونه ۲۰ اکانت پرتکرار/پربازنشر در هر خوش) تکمیل شود.

٦.٩ ملاحظات روش‌شناسی

● علامت‌گذاری یال‌ها در این تحلیل استنتاجی است (بر مبنای هم‌گروهی/بین‌گروهی) و علامت واقعی روابط به صورت مستقیم در داده وجود ندارد؛ بنابراین نتایج نیازمند اعتبارسنجی با تحلیل محتوایی نمونه‌ای از تعامل‌ها (به‌ویژه در توییتر نقل قولی) است.

● محاسبه تعادل بر اساس مجموعه‌ای از مثلث‌های استخراج شده/نمونه‌برداری شده انجام شد؛ از این‌رو، نسبت‌ها برآورده از تعادل شبکه علامت‌دار در زیرمجموعه تحلیل هستند.

چارچوب ارزش‌های خبری و تحلیل ۵ مطلب پر بازنشر هر روز

۱.۱۰ چارچوب نظری ارزش‌های خبری

برای تبیین این‌که «چرا یک مطلب خبری انتخاب یا بازنشر می‌شود»، از مفهوم News Values (یا Newsworthiness) استفاده شد. در ادبیات کلاسیک، گالتونگ و روگه مجموعه‌ای از عوامل/ارزش‌ها را برای توضیح منطق انتخاب خبر معرفی می‌کنند و پژوهش‌های بعدی (از جمله بازبینی هارکاپ و اونیل) این فهرست را مناسب با فضای رسانه‌ای جدید و شبکه‌های اجتماعی بازتعریف و به روزرسانی کرده‌اند [؟، ？]. در این سؤال، به صورت عملیاتی، ارزش‌های زیر در سطح هر «پست» سنجیده شد (هر پست می‌تواند هم‌زمان چند ارزش را داشته باشد):

- **بهنگام‌بودن** (Timeliness): خبر/موضوع مربوط به بازه زمانی جاری.
- **منفی‌بودن** (Negativity): تأکید بر رخدادهای منفی، تهدید، خشونت، بحران یا سوگ.
- **عارض/درگیری** (Conflict): حضور دوگانه تقابل، نزاع، اعتراض، برخورد یا زبان دوقطبی.
- **اثر/بزرگی** (Impact / Magnitude): اشاره به پیامدهای گسترده، اعداد/آمار یا شدت رخداد.
- **برجستگی/نخبگان** (Elite / Prominence): ارجاع به افراد/نهادهای شناخته‌شده یا کنشگران اثرگذار.
- **مجاورت** (Proximity): نزدیکی جغرافیایی/فرهنگی (برای این مسئله، اشاره به ایران و شهرهای ایران).
- **علاقه انسانی** (Human Interest): برجسته‌سازی تجربه فردی، قربانی، روایت شخصی یا مؤلفه‌های احساسی.
- **تصویر/ویدئو** (Visuals): نشانه‌هایی از محتوای دیداری (عکس/ویدئو/کلیپ).
- **قابلیت اشتراک‌گذاری** (Shareability): هشتگ، دعوت به انتشار یا الگوهای متین تسهیل‌کننده بازنشر.

- **غافلگیری** (Surprise / Unexpectedness): نشانه‌های «غیرمنتظره/عجیب/شوك‌آور» (در نتایج این داده‌ها مشاهده نشد).

۲.۱۰ عملیاتی‌سازی «پربازنشر بودن» و استخراج ۵ مطلب برترِ هر روز

داده ورودی شامل فایل‌های روزانه (روزهای ۱۵ تا ۱۹ دی) از هر پلتفرم بود که برای هر روز، ۵۰۰۰ پستِ پربازنشر همان روز را در بر می‌گرفت. به‌منظور تعریف یک معیار قابل بازتولید برای «پراکنش/انتشار»، برای هر پست s_i یک امتیاز انتشار به‌صورت زیر تعریف شد:

$$s_i = \begin{cases} \text{forward}_i & \text{اگر } \text{forward}_i > 0, \\ \text{engagement}_i & \text{در غیر این صورت} \end{cases}$$

این انتخاب از آن‌جا توجیه می‌شود که در برخی پلتفرم‌ها (به‌ویژه Instagram) مسیر بازنشر به‌صورت صریح در داده قابل مشاهده نیست و ستون forward یا در دسترس نیست یا عمدتاً مقدار صفر دارد؛ در نتیجه، engagement به‌عنوان جانشینِ قابل مشاهده به کار گرفته شد.

برای استخراج «۵ مطلب برتر هر روز» به‌صورت کلی (میان همه پلتفرم‌ها)، از آن‌جا که مقیاس s_i میان پلتفرم‌ها هم مقیاس نیست، ابتدا در هر روز d و هر پلتفرم p ، امتیاز انتشار به نمره استاندارد (z-score) تبدیل شد:

$$z_{i,p,d} = \frac{s_i - \mu_{p,d}}{\sigma_{p,d}},$$

که در آن $\mu_{p,d}$ و $\sigma_{p,d}$ به‌ترتیب میانگین و انحراف معیار spread_score در همان (p, d) هستند. سپس برای هر روز، ۵ پست با بیشترین مقدار $z_{i,p,d}$ انتخاب شد (و در حالت تساوی، به‌عنوان معیار شکستن تساوی استفاده گردید). علاوه بر این، برای کنترل و مقایسه درون‌پلتفرمی، «۵ مطلب برترِ هر (پلتفرم، روز)» نیز بر اساس s_i (بدون نرم‌السازی) استخراج و ذخیره شد.

۳.۱۰ برچسب‌گذاری ارزش‌های خبری در سطح پست

برای هر یک از ۲۵ پست منتخب (در ۵ روز و هر روز ۵ پست برتر کلی)، مجموعه‌ای از قواعد بازتولید‌پذیر (قاعدۀ محور) تعریف شد که با اتکا به ستون sentiment و واژه‌نامه‌های کوچک کلیدواژه (به‌همراه الگوهای ساده مانند وجود عدد یا هشتگ) برچسب دودویی هر ارزش را تولید می‌کند. نمونه‌ای از عملیاتی‌سازی قواعد به شکل زیر است:

اگر sentiment برابر «منفی» باشد یا متن حاوی واژه‌های منفی/سوگ/خشونت باشد. •

وجود کلیدواژه‌های تقابل (اعتراض، درگیری، برخورد، جنگ، حمله، ...) یا زبان قطبی. •

وجود عدد/آمار در متن یا واژه‌های شدت/پیامد (کشته، زخمی، گسترد، ...). •

اشاره به چهره‌ها/نهادهای شناخته شده یا عبارات مرتبط با ساختار قدرت/حاکمیت/اپوزیسیون. •

اشاره صریح به «ایران» یا شهرهای ایران. •

کلیدواژه‌های روایت فردی/احساسی (مردم، خانواده، کودک، هموطن، (...).

نشانه‌های رایج محتوای دیداری (عکس، ویدئو، کلیپ، emoji‌های رایج عکس). •

وجود هشتگ یا دعوت به اشتراک‌گذاری/بازنشر. •

این قواعد جایگزین تحلیل کیفی دقیق متن نیستند، اما برای افزودن یک لایه کمی قابل بازنگشته شدن بر روی خروجی «۵ پست برتر هر روز» کارکرد عملی دارند.

۴.۱۰ نتایج کمی ارزش‌های خبری در ۵ پست برتر هر روز

جدول ۶۵ تعداد وقوع هر ارزش خبری را در میان ۵ پست برتر کلی هر روز نشان می‌دهد (برای هر روز، حداقل هر ستون ۵ است). در این مجموعه، Timeliness برای همه موارد برقرار در نظر گرفته شد و Surprise با قواعد فعلی در هیچ‌یک از ۲۵ پست فعال نشد.

جدول ۶۵: شمارش ارزش‌های خبری در میان ۵ پست برتر کلی هر روز (روزهای ۱۵ تا ۱۹ دی).

روز دی	منفی	تعارض	اثر/بزرگی	برجستگی	مجاورت	علاقه انسانی	دیداری	اشتراک‌پذیری
1	0	3	5	5	3	1	3	15
1	2	0	4	2	4	3	5	16
3	0	3	5	5	5	5	5	17
1	1	0	5	2	4	3	3	18
1	0	1	5	5	4	2	2	19
7	3	7	24	19	20	14	18	جمع

۵.۱۰ تحلیل کیفی ۵ مطلب برتر هر روز

در این بخش، برای هر روز، ۵ پستِ برتر کلی (بر اساس نرمال‌سازی درون‌پلتفرمی و سپس انتخاب کلی) مرور شد. در هر مورد، «امتیاز انتشار» همان s است و برچسب‌های ارزش خبری مطابق قواعد بخش قبل استخراج شده‌اند.

روز ۱۵ دی.

pahlavireza/Twitter ●
 $s = 13248$: برجسته‌سازی رخدادهای منفی و قربانی محور (منفی، تعارض، اثر/بزرگی، مجاورت، علاقه انسانی) در قالب خطاب عمومی.

amirallii.ir/Instagram ●
 $s = 777622$: تمرکز بر مجاورت (اشاره به شهرها) و تقویت اشتراک‌پذیری از طریق هشتگ‌ها؛ مؤلفه‌های منفی/تعارض در متن کوتاه کمتر برجسته است (مجاورت، اشتراک‌پذیری، علاقه انسانی).

officialrezapahlevi/Instagram ●
 $s = 770609$: یادآوری/سوگواری جمعی با تأکید بر منفی‌بودن و علاقه انسانی، همراه با برجستگی منبع و مجاورت (منفی، برجستگی، مجاورت، علاقه انسانی).

moltafchannel/Telegram ●
 $s = 7458$: محتوای فرهنگی/اشتراکی (ارجاع به آلبوم/لينک) که در قواعد فعلی بیشتر از منظر «اثر/بزرگی» (به علت وجود الگوی عددی/اطلاعاتی در متن) و مجاورت برچسب خورده است؛ از نظر کارکرد خبری، این مورد به «تنوع محتوا» در میان پست‌های برتر نزدیک است.

pahlavicommss/Twitter ●
 $s = 6426$: محتوای سیاسی مبتنی بر چهره/منبع برجسته و تأکید بر اثر/پیامد (برجستگی، اثر/بزرگی، مجاورت).

روز ۱۶ دی.

trendspersian/Instagram ●
 $s = 529628$: بازنثر واکنش یک چهره شناخته شده و برانگیختن مشارکت مخاطب، همراه با نشانه‌های دیداری/کلیپ و هشتگ (برجستگی، دیداری، اشتراک‌پذیری، مجاورت، منفی).

akhbarefori/Telegram ●
 $s = 6648$: خبر خدماتی/اقتصادی درباره یارانه/کالا برگ که هم‌زمان باز منفی و مؤلفه تعارض (خطاب به نهاد/سیاست) را حمل می‌کند (منفی، تعارض، اثر/بزرگی، مجاورت).

کنشگر و زبان تقابلی (منفی، تعارض، برجستگی، مجاورت). ●
alighazizade/Twitter با $s = 2772$: تحلیل/ادعای سیاسی با اتکا به برجستگی

فعالی عمدتاً در مؤلفه‌های منفی و اثر/بزرگی و نیز نشانه دیداری فعال شده است (منفی، اثر/بزرگی، دیداری). ●
jenabesali46146/Twitter با $s = 2638$: متن افشاگرانه/اتهامی که در قواعد

بیشتر به قالب دیداری (نماد emoji دوربین) و مؤلفه‌های منفی/تعارض (منفی، تعارض، اثر/بزرگی، مجاورت). ●
akharinkhabar/Telegram با $s = 5701$: خبر مشابه درباره کالا برگ با اتکای

روز 17 دی.

نیروهای امنیتی/نظمی با او جگیری هم‌زمان منفی‌بودن، تعارض، اثر/بزرگی و برجستگی منبع؛ در عین حال، نشانه‌هایی از دعوت ضمنی به بازنشر نیز دیده می‌شود (منفی، تعارض، اثر/بزرگی، برجستگی، مجاورت، علاقه انسانی، اشتراک‌پذیری). ●
officialrezapahlavi/Instagram با $s = 2015069$: پیام مستقیم خطاب به

مشابه ارزش‌های خبری (منفی، تعارض، اثر/بزرگی، برجستگی، مجاورت، علاقه انسانی، اشتراک‌پذیری). ●
pahlavireza/Twitter با $s = 13402$: بازتاب همان مضمون در Twitter با الگوی

فعال شدن مجموعه مشابه ارزش‌ها (منفی، تعارض، اثر/بزرگی، برجستگی، مجاورت، علاقه انسانی، اشتراک‌پذیری). ●
officialrezapahlavi/Telegram با $s = 9813$: بازنشر همان پیام در تلگرام و

آن تعارض، اثر/بزرگی و منفی‌بودن پررنگ است (منفی، تعارض، اثر/بزرگی، برجستگی، مجاورت). ●
mamlekate/Telegram با $s = 7053$: روایت مبتنی بر «افشا/پیش‌بینی/حمله» که در

رخدادهای روزهای اخیر با زبان تقابلی و ارجاع زمانی/عددی (منفی، تعارض، اثر/بزرگی، برجستگی، مجاورت). ●
hichkasofficial/Twitter با $s = 7413$: متن زمان‌مند و فراخوان‌گونه درباره

روز 18 دی.

روایت موقیت/گستردگی $s = 2290283$ با officialrezapahhlavi/Instagram ●
رخدادهای مرتبط با روز 17 دی که در قواعد فعلی عمدتاً بر «اثر/بزرگی»، «برجستگی» و
«مجاورت» متمرکز می‌شود (اثر/بزرگی، برجستگی، مجاورت).

با زتاب همان مضمون با الگوی مشابه ارزش‌ها $s = 13084$ با pahlavireza/Twitter ●
(اثر/بزرگی، برجستگی، مجاورت).

گزارش رخداد خیابانی با بار منفی و تعارض، اما $s = 3304$ با ir_javan/Telegram ●
بدون فعال شدن برجستگی/علاقة انسانی در قواعد فعلی (منفی، تعارض، مجاورت).

ارائه گزارش همراه با تصویر/ویدئو از $s = 4038$ با farshidbabaei18/Twitter ●
اعتراضات و اشاره به عناصر عددی/توصیفی؛ در نتیجه، دیداری و اشتراک‌پذیری نیز
فعال شده است (منفی، تعارض، اثر/بزرگی، دیداری، اشتراک‌پذیری، مجاورت).

گزارش رخداد محلی (شهر/استان) با تأکید $s = 3757$ با pahlavicoms/Twitter ●
بر تعارض، اثر و مجاورت (منفی، تعارض، اثر/بزرگی، مجاورت).

روز 19 دی.

پیام/جمع‌بندی درباره رخدادهای $s = 291892$ با officialrezapahhlavi/Twitter ●
خیابانی با مؤلفه‌های منفی بودن، اثر/بزرگی و برجستگی منع (منفی، تعارض، اثر/بزرگی،
برجستگی، مجاورت).

روایت همزمان رخدادها با $s = 256549$ با officialrezapahhlavi/Twitter ●
تمرکز بر برجستگی و مجاورت (برجستگی، مجاورت).

نمادپردازی (پرچم) و ارجاع به کنش $s = 4335$ با pahlavicoms/Instagram ●
جمعی که در قواعد فعلی بیشتر به اثر/بزرگی، برجستگی و مجاورت نگاشت می‌شود
(اثر/بزرگی، برجستگی، مجاورت).

فراخوان آشکار برای تولید و بازنشر $s = 3602$ با hichkasofficial/Instagram ●
محظوظ به زبان‌های مختلف؛ در نتیجه علاقه انسانی و اشتراک‌پذیری نیز فعال شده است
(منفی، تعارض، اثر/بزرگی، برجستگی، مجاورت، علاقه انسانی، اشتراک‌پذیری).

• اطلاع‌رسانی درباره مسیرهای پروازی با حضور $s = 1301$ mamlekate/Telegram
عدد/جزئیات و تمرکز بر اثر/بزرگی و مجاورت (اثر/بزرگی، برجستگی، مجاورت).

کشف پیامرسانی هماهنگ و ارتباط آن با نفوذ شبکه‌ای

در این سؤال یک تحلیل روی داده‌های متنی انجام شد تا «پیامرسانی هماهنگ» به عنوان یک الگوی رفتاری قابل اندازه‌گیری استخراج شود و سپس با «نفوذ شبکه‌ای» (بر مبنای PageRank روی گراف بازنیش) مقایسه گردد. ایده اصلی این است که اگر تعداد زیادی اکانت/کanal در یک بازه زمانی کوتاه و در یک روز مشخص، متن‌هایی بسیار شبیه (یا تقریباً یکسان) منتشر کنند، می‌توان آن را به عنوان نشانه‌ای از هماهنگی (چه هماهنگی سازمان یافته و چه هماهنگی خودجوش) در نظر گرفت.

۱.۱۱ داده‌های استفاده شده

برای هر پلتفرم Instagram/Twitter، داده‌های متنی «5000 پست پُرانشار» در ۵ روز ۱۵ تا ۱۹ دی به صورت روزانه بررسی شد. همچنین برای سنجش نفوذ شبکه‌ای، PageRank وزن‌دار روی گراف جهت‌دار بازنیش همان پلتفرم محاسبه شد و مقدار PageRank هر اکانت (در حد امکان) به اکانت‌های حاضر در داده‌های متنی متصل گردید.

۲.۱۱ روش پیشنهادی برای کشف هماهنگی متنی

پیش‌پردازش متن. برای کاهش نویز، متن نرمال‌سازی شد (یکنواخت‌سازی فاصله‌ها، حذف URL و علائم کم‌اثر، و یکدست‌سازی حروف/کوچک‌سازی در حد ممکن) و سپس توکن‌سازی ساده انجام گرفت. هدف این مرحله آن بود که «کپی‌های نزدیک به هم» (با تغییرات جزئی مانند فاصله و علائم) همچنان مشابه تشخیص داده شوند.

اثر انگشت SimHash و معیار نزدیکی. به هر متن یک اثرانگشت 64-بیتی تخصیص داده شد. به طور خلاصه، برای هر توکن t یک هش 64-بیتی محاسبه می‌شود و سپس با جمع/تفريق وزن‌های توکن‌ها روی هر بیت، بردار امتیاز ساخته می‌شود و علامت هر مؤلفه بیت نهایی را تعیین می‌کند. در نتیجه، متن‌های مشابه اثرانگشت‌های نزدیک تولید می‌کنند. نزدیکی دو متن با فاصله همینگ

روی اثranگشت‌ها سنجیده شد:

$$d_H(f(x), f(y)) = \text{popcount}(f(x) \oplus f(y)),$$

که در آن $(\cdot) f$ اثranگشت 64-بیتی و \oplus عملگر XOR است. در این پروژه آستانه نزدیکی

$$d_H \leq 3$$

در نظر گرفته شد تا تنها «کپی‌های بسیار نزدیک» در یک خوشه قرار گیرند.

کاهش هزینه با LSH (باندبندي). برای اجتناب از مقایسه‌های زوجی $O(n^2)$, از باندبندي LSH استفاده شد: اثranگشت 64-بیتی به $b = 4$ باند 16-بیتی شکسته شد. هر متن در هر باند وارد یک «سطل» می‌شود و فقط متن‌هایی نامزد مقایسه دقیق شدند که دست‌کم در یک باند برخورد داشته باشند. سپس روی نامزدها شرط $d_H \leq 3$ اعمال شد.

ساخت خوشه‌ها. یک گراف روی «پست‌ها» ساخته شد که یال بین دو پست زمانی ایجاد می‌شود که (با روش بالا) شبکه‌تکراری تشخیص داده شوند. سپس مؤلفه‌های همبند (با-Union) به عنوان خوشة شبکه‌تکراری در نظر گرفته شد. در نهایت، یک خوشه «هماهنگ» محسوب شد اگر حداقل 5 پست و حداقل 5 اکانت یکتا داشته باشد تا تکرارهای ساده یک اکانت یا چند بازنشر محدود، به اشتباه هماهنگی تلقی نشود.

۳.۱۱ تعریف شاخص‌های کمی

برای هر پلتفرم و هر روز، کمیت‌های زیر محاسبه شد:

- تعداد خوشه‌های هماهنگ: N_{cluster}

- سهم پست‌های هماهنگ از کل پست‌های آن روز:

$$\text{CoordPostFrac} = \frac{|\{\text{clusters coordinated in posts}\}|}{|\{\text{posts all}\}|}$$

- سهم «انتشار/بازنشر» مربوط به پست‌های هماهنگ از کل انتشار/بازنشر آن روز (با جمع‌زن

امتیاز انتشار موجود در داده):

$$\text{CoordSpreadFrac} = \frac{\sum_{i \in \text{coord}} s_i}{\sum_{i \in \text{all}} s_i}$$

۴.۱۱ نتایج: شدت هماهنگی در روزهای مختلف

روز (دی)	سهم Twitter	N_{clus}	سهم Telegram	N_{clus}	سهم Instagram	N_{clus}
15	13.34%	50	5.20%	31	1.24%	7
16	16.20%	39	8.18%	29	0.30%	3
17	0.12%	1	0.44%	4	0.12%	1
18	0.38%	3	1.22%	11	0.96%	6
19	0.48%	3	0.48%	4	10.68%	29

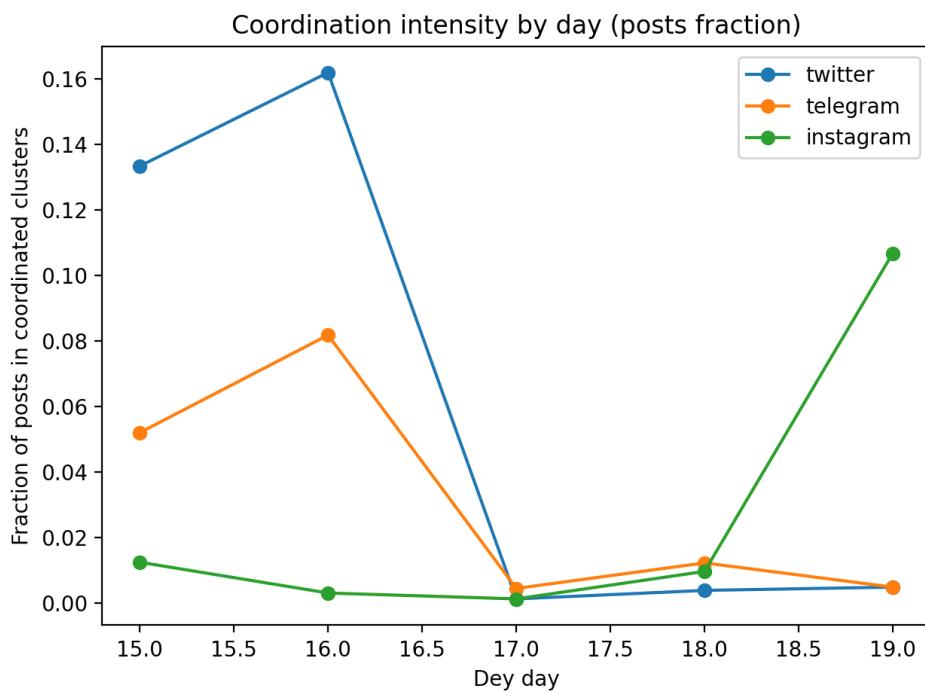
جدول ۶۶: شدت هماهنگی متنی بر حسب «سهم پست‌های هماهنگ» و «تعداد خوش‌های هماهنگ» در هر روز.

نتیجهٔ کلیدی جدول ۶۶ این است که هماهنگی متنی در Twitter و Telegram در روزهای ۱۵ و ۱۶ دی به مرتب پررنگ‌تر از روزهای بعدی است (به‌ویژه اوچ روز ۱۶ دی). در روز ۱۷ دی افت بسیار شدید مشاهده می‌شود و سپس در روزهای ۱۸ و ۱۹ دی، هماهنگی در Twitter و Telegram در سطح پایین باقی می‌ماند. در مقابل، Instagram در روز ۱۹ دی جهش قابل توجهی دارد (هم از نظر سهم پست هماهنگ و هم از نظر تعداد خوش‌ها).

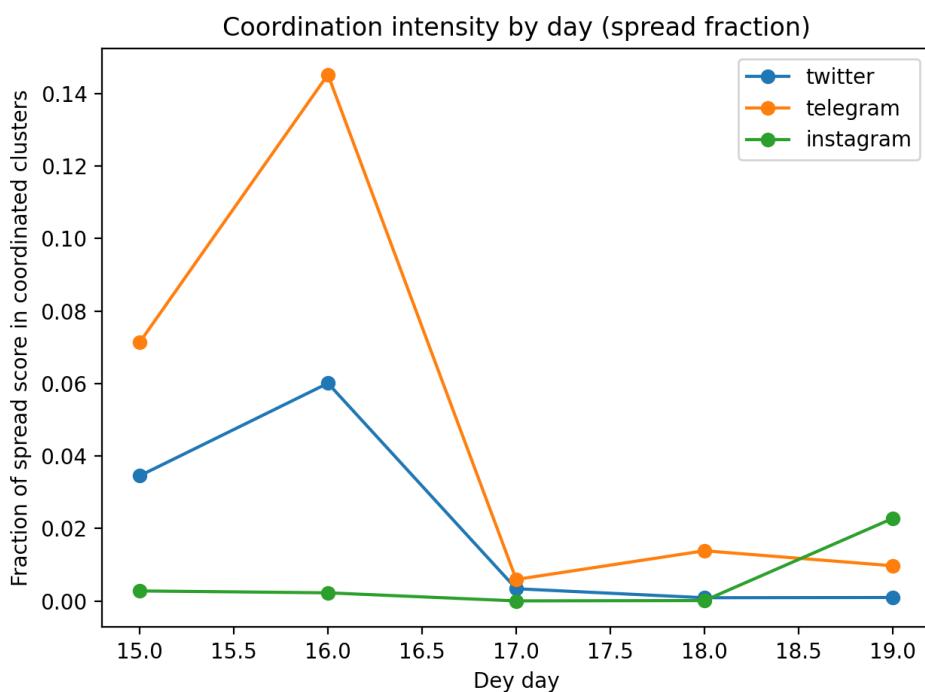
روز (دی)	سهم انتشار Twitter	سهم انتشار Telegram	سهم انتشار Instagram
15	3.453%	7.136%	0.277%
16	6.007%	14.525%	0.224%
17	0.337%	0.593%	0.001%
18	0.088%	1.385%	0.010%
19	0.095%	0.970%	2.278%

جدول ۶۷: شدت هماهنگی متنی بر حسب «سهم انتشار/بازنشر پست‌های هماهنگ» در هر روز.

طبق جدول ۶۷، نه تنها تعداد پست‌های هماهنگ در روز ۱۶ دی بالا است، بلکه سهم انتشار/بازنشر هماهنگ در Telegram نیز به حدود ۱۴.۵۲۵٪ می‌رسد؛ به بیان دیگر، بخش معناداری از گردش انتشار روزانه می‌تواند توسط چند خوش‌شنبه‌تکراری توضیح داده شود.



شکل ۱۱: روند سهم پست‌های هماهنگ در روزهای ۱۵ تا ۱۹ دی برای هر پلتفرم.



شکل ۱۲: روند سهم انتشار/بازنشر هماهنگ در روزهای ۱۵ تا ۱۹ دی برای هر پلتفرم.

۵.۱۱ نمونه‌های کیفی از خوش‌های بزرگ

برای پرهیز از محدودشدن خروجی به شاخص‌های عددی، بزرگ‌ترین خوش‌های هماهنگ به صورت نمونه بررسی شد. جدول ۶۸ سه نمونه از بزرگ‌ترین خوش‌ها را نشان می‌دهد (متن نماینده به صورت بردیده آورده شده است).

پلتفرم	روز (دی)	تعداد پست	تعداد اکانت	متن نماینده (بریده) + لینک نمونه
Twitter	16	205	204	#فراخوان_عمومی اولین فراخوان عمومی شاهزاده رضا پهلوی؛ و عده‌ ما پنجشنبه ۱۸ دی و جمعه ۱۹ دی رأس ساعت ۸ شب ... https://twitter.com/chawshin_83/status/2008613074362830954
Instagram	19	119	119	استقبال چشمگیر مردم کرج از فراخوان شاهزاده رضا پهلوی: «این آخرین نبرده، پهلوی برمی‌گرده» و «سیدعلی سرنگونه» پنجشنبه ۱۸ دی ... https://twitter.com/Mohamma08709989/status/2009546033957831043
Telegram	16	103	102	افرادی که «یارانه» نمی‌گیرند چگونه برای دریافت «کالا برگ» ثبت‌نام کنند؟ سرپرستان خانوار برای شروع ... https://telegram.me/akhbarefori/611346

جدول ۶۸: نمونه خوش‌های هماهنگ متنی در پلتفرم‌های مختلف.

به صورت کیفی، این نمونه‌ها نشان می‌دهند که خوش‌های بزرگ می‌توانند حول «فراخوان‌ها»، «سیاست‌های حمایتی/اقتصادی» یا «بازتاب یک روایت خبری» شکل بگیرند؛ بنابراین «هماهنگی» الزاماً هم معنی «محتوای جعلی» نیست، اما به عنوان یک سیگنال رفتاری می‌تواند برای شناسایی کمپین‌ها، شبکه‌های بازنثر و الگوهای تقویت مصنوعی دیده شدن مفید باشد.

۶.۱۱ ارتباط هماهنگی با نفوذ شبکه‌ای (مقایسه با PageRank)

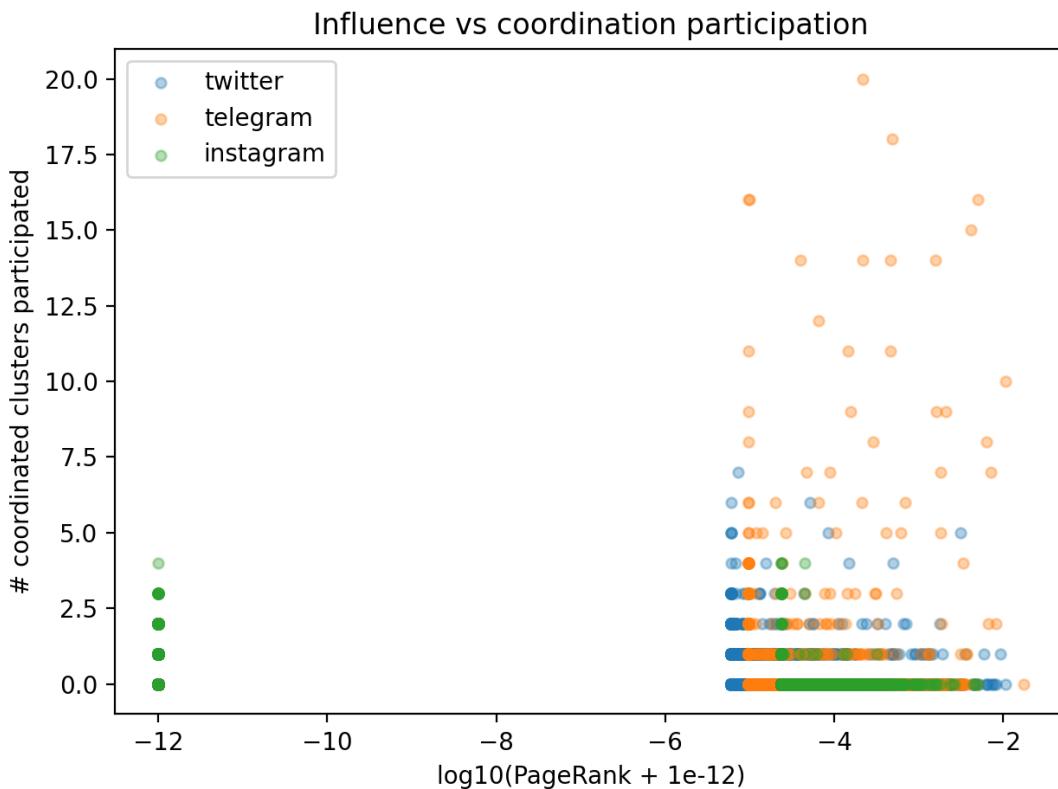
در گام بعد، برای هر اکانت حاضر در داده‌های متنی، شاخص‌هایی مانند «تعداد خوش‌های هماهنگ مشارکت‌کرده» و «سهم پست‌های هماهنگ همان اکانت» محاسبه شد و سپس با مقدار PageRank همان اکانت (محاسبه‌شده روی گراف بازنشر) مقایسه گردید. به‌شکل PageRank استاندارد روی گراف جهت‌دار، با درنظر گرفتن وزن یال‌ها و با $\alpha = 0.85$ محاسبه شد.

برای سنجش رابطه، از همبستگی رتبه‌ای اسپیرمن بین PageRank و «تعداد خوش‌های هماهنگ مشارکت‌کرده» استفاده شد. همچنین، در میان ۱۰۰۰ اکانت برتراز نظر PageRank (در اکانت‌هایی که در داده‌های متنی حضور داشتند)، سهم اکانت‌هایی محاسبه شد که دست‌کم در یک خوش‌های هماهنگ دیده می‌شوند.

سهم هماهنگ در ۱۰۰۰ اکانت برتراز	$\rho_s(\text{PageRank}, \#\text{Clusters})$	تعداد اکانت‌های دارای PageRank در داده متنی	پلتفرم
8.9%	-0.185	7304	Twitter
8.2%	-0.018	4403	Telegram
0.5%	-0.008	11713	Instagram

جدول ۶۹: رابطه هماهنگی متنی با نفوذ شبکه‌ای: همبستگی اسپیرمن و سهم اکانت‌های هماهنگ در میان ۱۰۰۰ اکانت برتراز PageRank.

طبق جدول ۶۹، در Twitter یک رابطه منفی قابل توجه مشاهده می‌شود: اکانت‌هایی که PageRank بالاتری دارند، به‌طور متوسط در خوش‌های شبکه‌تکراری کمتر ظاهر می‌شوند. تفسیر طبیعی این الگو آن است که «تقویت شبکه‌تکراری پیام» بیشتر توسعه اکانت‌های کم نفوذ‌تر انجام می‌شود و اکانت‌های مرکزی‌تر لزوماً برای دیده‌شدن به تکثیر شبکه‌تکراری متن متکی نیستند. در Instagram و Telegram این رابطه بسیار ضعیف است که می‌تواند ناشی از تفاوت سازوکار انتشار، تفاوت ماهیت کanal‌ها/اکانت‌ها، و نیز محدودیت داده (تمرکز روی ۵۰۰۰ پست پُرانشوار در هر روز) باشد.



شکل ۱۳: پراکنش PageRank در برابر تعداد خوشه‌های هماهنگ مشارکت‌کرده برای هر پلتفرم.

۷.۱۱ بررسی هماهنگی بین‌پلتفرمی

یک آزمون تکمیلی انجام شد تا بررسی شود آیا متن نماینده خوشه‌های بزرگ در یک روز مشخص، بین پلتفرم‌های مختلف نیز به صورت شبکه‌تکراری تکرار شده است یا خیر (با همان منطق LSH/SimHash و آستانه همینگ کمی سهل‌گیرانه‌تر). در خروجی این داده‌ها، خوشه بین‌پلتفرمی معناداری به دست نیامد. این نتیجه می‌تواند ناشی از تغییرات واژگانی/قالبی بین پلتفرم‌ها، تفاوت مخاطب، یا محدودیت‌تمرکز بر «خوشه‌های بزرگ هر روز» (و نه کل پست‌ها) باشد.

۸.۱۱ محدودیت‌ها و نکات دقت

- از آنجا که داده متنی فقط شامل 5000 پست پُرانشمار هر روز است، هماهنگی‌هایی که در لایه‌های کم‌انتشار رخ می‌دهند ممکن است مشاهده نشوند.
- آستانه $d_H \leq 3$ عمدتاً کپی‌های نزدیک را شناسایی می‌کند؛ بنابراین پارافرایزها یا نسخه‌های با تغییرات زیاد ممکن است از دید روش خارج بمانند.
- «هماهنگی» الزاماً «غیراصيل/مخرب» نیست؛ این تحلیل صرفاً یک سیگنال رفتاری ارائه

می‌کند و داوری محتوایی نیازمند تحلیل زمینه‌ای جداگانه است.