

کشف اجتماعات در شبکه تشکیل شده از گفتگوها پیرامون واکسن کرونا در توئیتر

محمد بیات^۱، زینب الهدی حشمتی^{۲*}

^۱دانشجوی دکتری، دانشکده‌ی علوم و فنون نوین، دانشگاه تهران، تهران، ایران.

^۲عضو هیئت علمی، دانشکده‌ی علوم و فنون نوین، دانشگاه تهران، تهران، ایران.

چکیده

شبکه‌های اجتماعی به طور فزاینده‌ای میزبان جامعه مدنی هستند، مکانی که در آن اشتراک دانش، بحث‌های عمومی و گفتگوهای شخصی انجام می‌شود. گفتگوهای شبکه‌های اجتماعی به عنوان میدان عمومی جدید به اندازه سایر اجتماعات بزرگ مردمی برای مستند سازی مهم هستند. گراف شبکه‌های مربوط به گفتگوها در شبکه‌های اجتماعی عمومی همچون توئیتر می‌تواند به بینش ما از نقش شبکه‌های اجتماعی در روابط بین انسان‌ها کمک کند. یافتن اجتماعات یکی از مهم‌ترین زمینه‌های تحقیقاتی در حوزه تحلیل شبکه‌های پیچیده است و با استفاده از آن می‌توان ساختار شبکه‌های پیچیده را بهتر درک کرد. در این تحقیق انواع اجتماعات در شبکه اجتماعی توئیتر مورد بررسی قرار خواهد گرفت و پس از آن به الگوریتم‌های مربوط به کشف این اجتماعات اشاره خواهد شد و در نهایت اجتماعات شبکه حاصل از گفتگوها پیرامون واکسن کرونا مورد بررسی و تجزیه و تحلیل قرار خواهد گرفت.

کلمات کلیدی: شبکه‌های پیچیده، شبکه‌های اجتماعی، کشف اجتماعات، توئیتر، واکسن، ویروس کرونا.

تاریخچه مقاله:

تاریخ ارسال: ۱۳۹۹/۰۵/۰۱

تاریخ اصلاحات: ۱۳۹۹/۰۶/۰۲

تاریخ پذیرش: ۱۳۹۹/۰۶/۳۰

تاریخ انتشار: ۱۳۹۹/۰۷/۰۲

Keywords:

Complex Networks
Social Networks
Community Detection
Twitter
Coronavirus Vaccine

*ایمیل نویسنده مسئول:

zheshmati@ut.ac.ir

Community Detection in the Network of Conversations about Coronavirus Vaccination in Twitter

Muhammad Bayat¹, Zainabohoda Heshmati^{2*}

^{1,2}Faculty of Modern Science & Technology, University of Tehran, Tehran, Iran.

Abstract

Social networks increasingly host civil society, a place for sharing knowledge, discussing public matters, and conversing private affairs. Social network conversations, as new public spaces, are as important as other large communities for documentation. Network graphs which are related to conversations in public social networks like Twitter, can help us ponder over the role of social networks in human relations. Detection of communities is one of the most important research areas in the field of complex network analysis and applied to better understand the structure of complex networks. In this study, first varieties of communities will be examined in the social network of Twitter, then algorithms which are related to discovery of these communities will be discussed, finally the network of communities resulting from conversations about Coronavirus vaccination will be scrutinized.

۱ - مقدمه

شبکه‌ها یا گراف‌ها یک دید انتزاعی از سیستم‌های پیچیده همچون شبکه‌های اجتماعی، شبکه‌های کامپیوتری، شبکه‌های بیولوژیکی و غیره به ما می‌دهند. شبکه‌ها متشکل از مجموعه‌ای از گره‌ها و یال‌هایی هستند که یالها، گره‌ها را به یکدیگر متصل می‌کنند. یکی از روش‌های متداول مطالعه شبکه، تمرکز بر روی خصوصیات گره‌ها به صورت تکی و یا بصورت گروهی است. برای مثال یافتن مهمترین گره در شبکه و یا اینکه کدام گره ارتباطات بیشتری دارد. این رویکرد در مطالعه، اطلاعات اندکی از شبکه‌های مقیاس آزاد^۱ به ما می‌دهد. شبکه‌های مقیاس آزاد به طور معمول دارای ویژگی‌های مشترکی هستند. برای مثال میانگین فاصله بین دو گره در شبکه در مقایسه با اندازه شبکه خیلی کمتر است (نظریه شش درجه جدایی). ویژگی دوم این شبکه‌ها این است که در این شبکه‌ها تعداد زیادی گره وجود دارند که همسایه‌های کمی دارند و تعداد کمی گره وجود دارد که کانون^۲ هستند و همسایه‌های زیادی دارند. ویژگی سوم شبکه‌های مقیاس آزاد چگالی کم آنها است. این تراکم کم معمولاً مربوط به محدودیت‌های فردی هر گره است که نمی‌تواند روابط زیادی برقرار کند. بنابراین، درجه متوسط گراف برخلاف ضریب خوشه بندی^۳ که زیاد است، بسیار کم است. وقتی ضریب خوشه بندی زیاد است یعنی چگالی محلی در شبکه زیاد است. تفاوت چگالی در کل شبکه و چگالی محلی به طور کلی با حضور گروهی از گره-ها که به شدت به یکدیگر مرتبط هستند و به آنها اجتماعات گفته می‌شود تعریف می‌شود [۱]. مقدار چگالی یک اجتماع برابر است با نسبت تعداد اتصالات در یک اجتماع به کل اتصالات ممکن در آن اجتماع. این مقدار بین صفر (عدم وجود یال بین گره‌ها) و یک (همه گره‌ها با همه گره‌ها متصل هستند) است. وقتی اندازه یک اجتماع بزرگ می‌شود (تعداد گره‌ها زیاد می‌شود) شکل‌گیری اتصال و تشکیل یال بین همه گره‌ها بسیار دشوار خواهد بود بنابراین می‌توان نتیجه گرفت هرچه تعداد افراد در یک شبکه اجتماعی زیاد شود چگالی شبکه کاهش می‌یابد.

یافتن اجتماعات یکی از مهم‌ترین زمینه‌های تحقیقاتی در حوزه تحلیل شبکه‌های پیچیده است و با استفاده از آن می‌توان ساختار شبکه‌های پیچیده را بهتر درک کرد. کشف اجتماعات در

جامعه شناسی، زیست شناسی و علوم کامپیوتر از اهمیت زیادی برخوردار است. اجتماعات به دو دسته تقسیم می‌شوند: اجتماعات مبتنی بر ساختار شبکه که در آنها خوشه‌هایی از گره‌ها وجود دارند که بین خودشان ارتباطات زیادی دارند و با بقیه گره‌های شبکه ارتباط کمی دارند و اجتماعات مبتنی بر معنا^۴ که در آنها گره‌هایی وجود دارند که زمینه معنایی یکسانی دارند و یا اینکه علایق مشترکی دارند. توصیف همزمان هر دو دسته بخاطر ویژگی‌های متفاوت ساختار توپولوژیکی و زمینه معنایی بسیار چالش برانگیز است. محاسبات مربوط به یافتن اجتماعات در شبکه‌ها به دو دلیل پیچیده است. اول اینکه اندازه شبکه‌ها مخصوصاً شبکه‌های اجتماعی بسیار بزرگ است و در برخی از آنها تعداد گره‌ها چند صد میلیون تا چند میلیارد است. دومین دلیل مربوط به ذات پویایی شبکه‌هاست و اینکه شبکه‌ها بطور مداوم در حال تغییر هستند [۲]. از دیدگاه نظری یافتن اجتماعات از نوع مسائل NP-کامل^۵ در بررسی گراف‌ها است؛ بدین مفهوم که یافتن الگوریتمی کارا که بتواند همیشه جواب بهینه این مسئله را در زمان چند جمله‌ای بیابد بعید به نظر می‌رسد [۳].

شبکه‌های اجتماعی شکلی از ارتباطات اجتماعی است که از طریق آن، کاربران به خلق ارتباطات برخط جهت به اشتراک گذاری اطلاعات، ایده‌ها، پیام‌های شخصی و محتوای متنوع می‌پردازند. یک شبکه اجتماعی ساختاری اجتماعی است، متشکل از گره‌هایی که به صورت منفرد یا گروهی به یکدیگر مرتبط شده‌اند و روابطی همچون مبادلات مالی، دوستی، بازرگانی، پیوندهای وب و سرگرمی دارند. استفاده از این شبکه‌ها فقط محدود به گروه‌های دوستی نبوده و امروزه شرکت‌ها به منظور کمک به کارمندان و یا کمک به مشتریان و دست یافتن به اطلاعات، به استفاده از شبکه‌های اجتماعی روی آورده‌اند [۴].

گراف شبکه‌های مربوط به گفتگوها در شبکه‌های اجتماعی عمومی همچون توییتر می‌تواند به بینش ما از نقش شبکه‌های اجتماعی در روابط بین انسان‌ها کمک کند. گفتگوها در توییتر باعث ایجاد شبکه‌هایی با حدود مشخص می‌شود زیرا مردم به یکدیگر پاسخ می‌دهند، در پیام‌های خود به یکدیگر اشاره می‌کنند و یا از هشتگ‌ها در پیام‌های خود استفاده می‌کنند. این ساختارهای گفتگو بسته به موضوع و افرادی که گفتگو را بر عهده دارند متفاوت است.

¹ Scale-Free Networks

² Hub

³ Clustering Coefficient

⁴ Context

⁵ NP-Complete

۱-۱ بیان مسئله

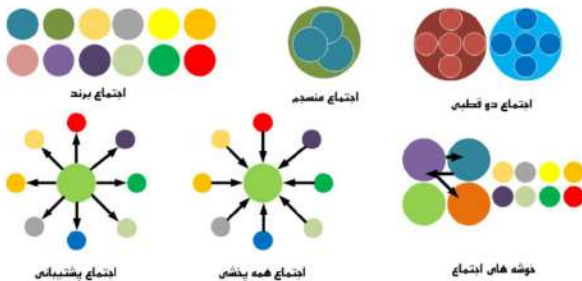
در حال حاضر یکی از مباحث مهم در سطح جامعه، واکسن مربوط به ویروس کرونا است. اینکه این واکسن خارجی باشد و یا ایرانی، اینکه واکسن خارجی چینی یا روسی باشد و یا اینکه واکسنی از یک شرکت معتبر اروپایی و یا امریکایی با تاییدیه سازمان بهداشت جهانی. اینها مباحث داغ حال حاضر جامعه هستند. هدف از این تحقیق یافتن اجتماعاتی است که روی موضوع واکسن کرونا گفتگو می کنند. آنچه مشخص است اغلب افراد جامعه از دولت انتظار دارند که واکسن های مربوط به شرکت های معتبر اروپایی و امریکایی که مورد تایید سازمان بهداشت جهانی است خریداری کند ولی رویکرد حاکمیت خلاف نظر مردم است و این نظر مردم در گفتگوهای توئیتر قابل مشاهده است. برای همین انتظار داریم با توجه به آراء مختلف در این حوزه جوامع مختلفی را در بحث واکسن مشاهده کنیم. حال سوال اصلی این تحقیق این است که آیا اجتماعات متفاوتی حول محور این موضوع در شبکه های اجتماعی وجود دارد؟ با بررسی گفتگوها پیرامون واکسنیاسیون کرونا ما می توانیم اجتماعاتی که پیرامون این موضوع شکل گرفته است را شناسایی کنیم. کشف اجتماعات پیرامون مباحث این چینی از این جهت اهمیت دارند که به حاکمیت و جامعه شناسان کمک می کند تا نظرات جامعه را پیرامون موضوع خاصی دریافت کنند و در تصمیم گیری های خود لحاظ نمایند. ادامه ی مقاله به شرح زیر می باشد:

در بخش ۲ پیشینه تحقیق مورد بررسی قرار خواهد گرفت. در بخش ۳ بطور خلاصه الگوریتم های کشف اجتماعات مخصوصاً الگوریتم Louvain بررسی خواهند شد. بخش ۴ شامل کشف و بررسی اجتماعات در شبکه پیرامون گفتگوها درباره واکسن کرونا خواهد بود و در بخش ۵ به نتیجه گیری و کارهای آتی خواهیم پرداخت.

۲- پیشینه تحقیق

یکی از تحقیقاتی که در حوزه کشف اجتماعات در شبکه اجتماعی توئیتر انجام گرفت مربوط به همکاری بین پروژه اینترنتی مرکز تحقیقات پیو^۶ و بنیاد تحقیقات رسانه های اجتماعی^۷ می باشد [۵]. هنگامی که کاربران توئیتر یکدیگر را دنبال می کنند، به یکدیگر پاسخ می دهند، در توئیتهای خود از یکدیگر نام می برند، درباره یک موضوع مشخص بحث می کنند و یا اینکه از

هشتگی در پیام های خود استفاده می کنند شبکه ای بینشان شکل می گیرد. تمامی شبکه های اجتماعی یک ساختار کلی دارند در حالیکه افرادی که درون این شبکه ها هستند خود نیز با توجه به ارتباطاتی که با یکدیگر دارند، شبکه هایی با ساختارهای متفاوت شکل می دهند و افراد کلیدی درون این شبکه های محلی موقعیت های استراتژیک مثل کانون و یا مرجع دارند. گراف حاصل از شبکه ی اجتماعی توئیتر می تواند افراد کلیدی در یک شبکه تشکیل شده از یک گفتگو را برجسته کند. برخی از کاربران در این شبکه ها بسیار بیشتر از کاربران دیگر یال ورودی و یا یال خروجی دارند. گراف این شبکه ها افراد کلیدی در یک بحث را که کانون بحث هستند و بیش از بقیه افراد، گفتگوهای آنها مورد توجه قرار می گیرد و توسط افراد دیگر بازنشر می شود را به خوبی نشان می دهد. گفتگوها در توئیتر ممکن است گروه ها و اجتماعات مختلفی بوجود بیاورند. برخی از افراد در این شبکه ها در چندین اجتماع حضور دارند که به آنها پل^۸ گفته می شود. پل ها نقش مهمی در جابه جایی اطلاعات بین گروه ها و اجتماعات دارند. با توجه به تحقیقات انجام شده در [۵] شش نوع اجتماع در شبکه های اجتماعی همچون توئیتر وجود دارد. ممکن است



(شکل-۱): ساختار اجتماعات در توئیتر

اجتماعات دیگری هم وجود داشته باشند که در این تحقیق به آنها اشاره نشده است. برای دستیابی به مجموعه داده مورد نیاز از شبکه اجتماعی توئیتر در این تحقیق، از کلمات کلیدی مورد نظر و هشتگ های مشخص استفاده شده است. یال های شبکه از محتوای هر پیام که از نتایج جستجوی توئیتر استخراج شده است، شکل می گیرد. هر یال یا اتصال در این شبکه حاصل یک پاسخ به یک پیام و یا نام بردن از شخصی در یک پیام است. همچنین وقتی افراد در این شبکه یکدیگر را دنبال می کنند، این نیز یکی از

⁶ Pew Research Center Internet Project

⁷ Social Media Research Foundation

⁸ Bridge

زیر گروه‌هایی که در این اجتماع شکل می‌گیرد ارتباطات بسیار زیادی درون خود و با زیر گروه‌های دیگر دارند. وبسایت‌هایی که افراد گروه‌های مختلف به آنها ارجاع می‌دهند و یا هشتک‌هایی که در پیام‌هایشان استفاده می‌کنند در بین گروه‌ها همپوشانی دارند و معمولاً در بین اغلب اعضای این اجتماع مشترک هستند که نشان دهنده ترجیحات و علایق مشترک آنها است و مانند اجتماع دو قطبی نمی‌توان وبسایت‌ها و یا هشتک‌هایی یافت که در تضاد با یکدیگر باشند و فقط توسط یک گروه مورد استفاده قرار گیرند. این اشتراکات در محتوا نشان می‌دهد که گروه‌هایی که این شبکه‌ها را شکل می‌دهند اختلافات کمی با یکدیگر دارند و در تعامل و ارتباط تنگاتنگ با یکدیگر هستند.

۳-۲ اجتماع برند

افرادی که در مورد برندهای تجاری و یا سایر موضوعات عمومی توییت می‌کنند اغلب ساختار شبکه‌ای را تشکیل می‌دهند که با ساختار اجتماعات دو قطبی و منسجم متفاوت است. وقتی در توییت درباره محصولات یا خدمات شناخته شده و یا موضوعات داغ مانند افراد مشهور بحث می‌شود، غالباً پیام‌ها از افرادی است که هیچ ارتباطی با یکدیگر ندارند. برندهای معروف و سایر سوژه‌های محبوب می‌توانند تعداد زیادی از افراد پراکنده در توییت را که در مورد آن موضوع توییت می‌کنند، جذب کنند اما هیچ رابطه‌ای بین این افراد وجود ندارد. هرچه جمعیتی که درباره یک نام تجاری صحبت می‌کنند بیشتر باشد احتمال ارتباط شرکت کنندگان با یکدیگر کمتر است بنابراین این اجتماعات دارای تراکم کمی هستند و اغلب افراد، منزوی^{۱۰} هستند. این شرکت کنندگان در مباحث حول نام تجاری تمایلی به ارتباط با یکدیگر ندارند.

در این اجتماع معمولاً یک جمعیت بزرگ غیر متصل و منزوی و تعدادی خوشه‌های کوچک از افراد متصل وجود دارند که پیرامون یک برند شکل گرفته‌اند. این خوشه‌های کوچک ارتباط کمی با خوشه‌های دیگر دارند و از هشتک‌ها و یا URL‌های متفاوتی در پیام‌های خود استفاده می‌کنند. وبسایت‌هایی که هر یک از این گروه‌ها به آنها ارجاع می‌دهند از یکدیگر متمایز هستند و همپوشانی کمی با هم دارند و این نشان دهنده این است که هر

منابع شکل‌گیری پال‌ها در شبکه توییت است. شکل ۱ ساختار شش اجتماع در شبکه اجتماعی توییت را نشان می‌دهد.

۱-۲ اجتماع دو قطبی

اجتماعات دو قطبی شامل دو گروه بزرگ و متراکم است که ارتباط کمی بین آنها وجود دارد. موضوعات مورد بحث اغلب موضوعات سیاسی بسیار تفرقه‌انگیز و داغی است، برای مثال بحث انتخابات ریاست جمهوری. در اجتماع دو قطبی گفتگوی کمی بین دو گروه صورت می‌گیرد مگر اینکه هر دو گروه روی یک موضوع گفتگو کنند. این اجتماعات به ما نشان می‌دهند که هر یک از گروه‌ها فقط به منابع خاصی که مورد تایید خودشان است ارجاع می‌دهند. اعضای هر گروه به پیام‌های یکدیگر پاسخ می‌دهند و یا در پیام‌های خود از یکدیگر نام می‌برند ولی ارتباط خیلی کمی بین اعضای دو گروه دیده می‌شود. کلمات، هشتک‌ها و وبسایت‌هایی که دو گروه از آنها استفاده می‌کنند و یا به آنها ارجاع می‌دهند نیز کاملاً بین هر دو گروه متفاوت است. در بین هر یک از این دو اجتماع افراد کمی هستند که کانون هستند و ارتباطات بسیار زیادی با دیگر افراد دارند. افراد بسیار اندکی نیز وجود دارند که با اعضای هر دو گروه در ارتباط هستند و نقش پل را در این شبکه دارند.

۲-۲ اجتماع منسجم^۹

برخلاف اجتماع دو قطبی، افرادی که این اجتماعات را شکل می‌دهند ارتباطات بسیار قوی با یکدیگر دارند. مشخصه این گفتگوها افراد بسیار بهم پیوسته‌ای هستند و تعداد کمی از افراد در انزوا هستند. بسیاری از کنفرانس‌ها، موضوعات حرفه‌ای، گروه‌های سرگرمی و سایر موضوعاتی که مردم را به خود جلب می‌کنند از این شکل اجتماعات هستند. در این اجتماعات افراد اغلب یکدیگر را می‌شناسند و با یکدیگر گفتگو می‌کنند، یکدیگر را در شبکه‌های اجتماعی دنبال می‌کنند، به پیام یکدیگر پاسخ می‌دهند و یا در پیام‌های خود از یکدیگر نام می‌برند. اینها افرادی هستند که علایق و یا گرایش‌های مشترکی دارند. برخلاف اجتماع دو قطبی گروه‌های مخالف هم در این شبکه‌ها وجود ندارد و ارتباطات بسیار زیادی بین گروه‌هایی که شکل می‌گیرد وجود دارد.

¹⁰ Isolated

⁹ Tight

یک از این گروه‌ها به یک موضوع خاص در رابطه با برند مورد نظر علاقه دارند.

(جدول-۱): مقایسه بین انواع اجتماعات در شبکه اجتماعی توئیتر

مثال	تعداد گره‌های منزوی	میزان ارتباط بین گروه‌ها	تعداد و اندازه گروه‌ها	ساختار
مباحث سیاسی	کم	عدم ارتباط	۲ گروه بزرگ	اجتماع دو قطبی
موضوعات پیرامون کنفرانس‌ها	کم	ارتباطات زیاد	۲-۶ گروه متوسط	اجتماع منسجم
گفتگوهای پیرامون یک برند تجاری	زیاد	ارتباطات کم	تعداد زیادی گروه کوچک	اجتماع برند
موضوعات رسانهای	کم	ارتباطات متوسط	تعداد زیادی گروه کوچک	خوشه‌های اجتماع
مباحثی که یک آدم مشهور شروع می‌کند	متوسط	ارتباطات کم	یک گروه بزرگ و چندین گروه کوچک	اجتماع همه پخش
خدمات پشتیبانی شرکت‌ها	متوسط	ارتباطات متوسط	یک گروه بزرگ و چندین گروه کوچک	اجتماع پشتیبانی

۲-۴ خوشه‌های اجتماع

برخی از موضوعات پرتعداد ممکن است چندین گروه کوچکتر ایجاد کنند که اغلب در اطراف چند کانون تشکیل می‌شوند، هرکدام با مخاطبان، تأثیرگذاران و منابع اطلاعاتی خاص خود. داستان‌های خبری بین‌المللی اغلب از سوی خبرگزاری‌های مختلف مبتنی بر علاقه مندی آنها پوشش داده می‌شوند و هر کدام دنبال کنندگان مخصوص به خود را دارند که این باعث تشکیل گروه‌هایی با اندازه‌های مختلف می‌گردد. برخی از موضوعات، گفتگوهای متعددی را برانگیخته و هرکدام مخاطبان و جامعه خود را گسترش می‌دهند. این گروه‌ها می‌توانند زوایای متنوعی از یک موضوع را بر اساس ارتباط آن با مخاطبان مختلف و تنوع نظرات و دیدگاه‌ها، در مورد یک موضوع در شبکه‌های اجتماعی نشان دهند.

زیر گروه‌های تشکیل شده در این اجتماع اندازه‌های تقریباً مشابهی دارند. در این اجتماع همانند اجتماع برند، تعداد زیادی از افراد با یکدیگر ارتباطی ندارند و بصورت تکی در این اجتماع دیده می‌شوند. گروه‌هایی که در این اجتماع وجود دارند به یک جنبه خاص پیرامون موضوع گفتگو علاقه‌مند هستند و بزرگتر از گروه‌های اجتماع برند هستند. همچنین ارتباط بین گروه‌ها نیز زیادتر است. هشتگ‌ها و وبسایت‌هایی که در هر یک از این گروه‌ها مورد استفاده قرار می‌گیرد با یکدیگر همپوشانی دارند.

۲-۵ اجتماع همه پخش^{۱۱}

بسیاری از افراد در توئیتر آنچه را که خبرگزاری‌ها، رسانه‌ها و یا افراد مشهور منتشر می‌کنند، بازنشر می‌کنند و یا اینکه روی پیام آنها اظهار نظر می‌کنند. این افراد معمولاً با یکدیگر در ارتباط نیستند و فقط به خبرگزاری که در شبکه یک کانون است متصل هستند. اجتماعات همه پخش معمولاً شامل یک یا دو کانون با اتصالات بسیار زیاد در اطراف آنها و چندین گروه کوچک و متراکم است. این گروه‌ها بطور معمول فقط پیرامون یک موضوع مشخص که برای مثال یک خبرگزاری آغاز کرده است گفتگو می‌کنند.

¹¹ Broadcast

بینابینی^۲ استفاده می‌کند. روش کار بدین صورت است که مرکزیت بینابینی برای تک‌تک یالها محاسبه می‌شود و یال‌هایی که بیشترین مقدار را دارند به عنوان یال‌های بین اجتماعات شناسایی شده و حذف می‌شوند و این کار تا تشخیص کل اجتماعات ادامه می‌یابد. مشکل این الگوریتم هزینه بالای محاسباتی آن است.

- **الگوریتم‌های تجمعی^۳:** که به صورت بازگشتی گروه‌ها و یا اجتماعات مشابه را با یکدیگر ادغام می‌کنند [۹،۱۰]. این الگوریتم‌ها مبتنی بر این واقعیت هستند که گره‌های یک اجتماع ویژگی‌های مشترکی دارند و می‌توان از این ویژگی‌های مشترک برای گروه‌بندی استفاده کرد. بر خلاف الگوریتم‌های تقسیم‌کننده، روش تجمعی در ابتدا همه گره‌ها را جدا از هم و غیر متصل در نظر می‌گیرد و آنها را بر اساس ویژگی‌های مشترک به هم متصل می‌کند تا به اجتماعات برسد. نحوه عملکرد این رویکرد بدین صورت است که یال‌های بین اجتماعات خود به خود حذف شده و تنها یال‌های داخل اجتماع باقی می‌مانند.

- **الگوریتم‌های پیشینه‌سازی پیمان‌بندی^۴:** که بر اساس حداکثر سازی یک تابع هدف است [۱۱]. نیومن معیار پیمان‌بندی را برای ارزیابی کیفیت اجتماعات معرفی کرد. میزان دقت این الگوریتم‌ها در پارتیشن‌بندی شبکه با توجه به معیار پیمان‌بندی پارتیشن‌ها اندازه‌گیری می‌شود. با داشتن یک افراز از شبکه، معیار پیمان‌بندی یال‌های درون اجتماعات را با گراف تصادفی فرضی با توزیع درجه یکسان با شبکه مورد نظر مقایسه و از این طریق کیفیت خوشه‌ها را ارزیابی می‌کند. پیمان‌بندی یک پارتیشن یک مقدار اسکالر بین ۱- و ۱ است که نشان‌دهنده میزان تراکم یال‌ها درون یک اجتماع در مقایسه با تراکم یال‌های بین اجتماعات است. برای شبکه‌های وزن دار مقدار پیمان‌بندی از رابطه ۱ محاسبه می‌شود:

$$Q = \frac{1}{2m} \sum_{i,j} [A_{i,j} - \frac{k_i k_j}{2m}] \delta(c_i, c_j) \quad (\text{رابطه-۱})$$

که در این رابطه $A_{i,j}$ وزن یال بین گره i و j است، $k_i = \sum_j A_{i,j}$ مجموع وزن یال‌هایی است که به گره i متصل شده‌اند، c_i اجتماعی است که گره i عضو آن است، تابع δ برابر با

هشتگ‌ها و وسایط‌هایی که این گروه‌ها در پیام‌هایشان استفاده می‌کنند معمولاً مشترک هستند.

۶-۲ اجتماع پشتیبانی

شرکت‌های بزرگ برای ارائه خدمات پشتیبانی به مشتریان خود از حساب توئیتری استفاده می‌کنند. از این حساب کاربری شرکت‌ها برای پاسخ‌گویی به مشتریان در رابطه با محصولات و خدمات خود استفاده می‌کنند. در این مدل یک کانون با اتصالات زیاد در اطراف آن خواهیم داشت اما متفاوت از ساختار اجتماع همه‌پخشی. در این مدل، کانون پاسخگوی تعداد زیادی از افرادی است که هیچ ارتباطی با یکدیگر ندارند و این افراد با یال خروجی به کانون‌ها متصل هستند، بر خلاف مدل همه پخشی که افراد با باز نشر توئیتهای کانون از طریق یال ورودی به کانون متصل‌اند. بررسی اجتماع پشتیبانی پیرامون حساب کاربری و یا هشتگ مربوط به پشتیبانی شرکت‌ها نشان می‌دهد که یک کانون در این اجتماع وجود دارد که به درخواست‌ها و سوالات افراد زیادی که هیچ ارتباطی با یکدیگر ندارند پاسخ می‌دهد. این گروه‌ها معمولاً از هشتگ‌ها و وسایط‌های مشترک و مرتبط با محصولات و خدمات شرکت در پیام‌های خود استفاده می‌کنند. جدول ۱ ویژگی‌های اجتماعات در شبکه‌ی اجتماعی توئیتر را بطور خلاصه نشان می‌دهد.

۳- الگوریتم‌های تشخیص اجتماعات

برای پیدا کردن اجتماعات در یک شبکه باید شبکه را به گروه‌های کوچکتر تقسیم کرد به گونه‌ای که در داخل هر یک از این گروه‌ها ارتباطات مترادفی بین گره‌ها وجود دارد و گره‌هایی که در دو اجتماع متفاوت قرار دارند به ندرت با یکدیگر در ارتباط هستند. الگوریتم‌های متعددی برای پارتیشن‌بندی شبکه و یافتن اجتماعات ارائه شده است که با توجه به مرتبه اجرای الگوریتم و میزان حافظه مورد نیاز با یکدیگر متفاوت هستند. الگوریتم‌های کشف اجتماعات را می‌توان به سه دسته تقسیم کرد:

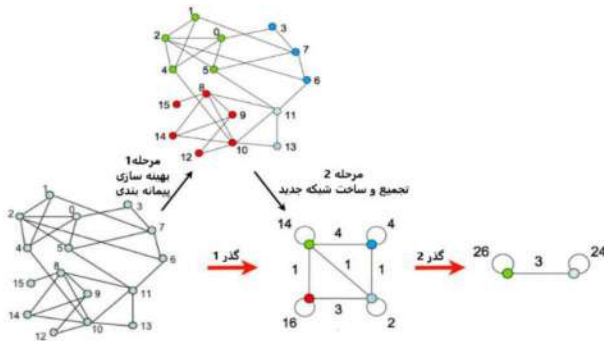
- **الگوریتم‌های تقسیم‌کننده^۱:** این الگوریتم‌ها پیوندهای بین اجتماعات را تشخیص داده و آنها را در یک فرایند تکراری از شبکه حذف می‌کنند [۶-۸]. چون اجتماعات مختلف توسط تعداد کمی یال به هم متصل شده‌اند، این روش گلوگاه‌هایی که در یالها قرار دارند و دو اجتماع را به هم متصل می‌کنند مورد توجه قرار می‌دهد. از طریق این یالها همه مسیرهای کوتاه عبور می‌کنند. این روش از مرکزیت

² Betweenness Centrality

³ Agglomerative

⁴ Modularity Maximization

¹ Divisive



(شکل-۲): مراحل اجرای الگوریتم Louvain

۴- کشف اجتماعات پیرامون واکسیناسیون کرونا

در این بخش به دنبال کشف و بررسی اجتماعات در شبکه تشکیل شده از گفتگوها پیرامون واکسیناسیون کرونا در شبکه اجتماعی توییتر هستیم. نرم افزار مورد استفاده در این بخش Gephi ورژن ۰.۹.۲ است که روی سیستم عامل ویندوز ۱۰ نصب شده است [۱۳]. برای جمع‌آوری داده‌های شبکه مربوط به این گفتگوها از افزونه Twitter Streaming Importer در نرم‌افزار Gephi استفاده شده است. برای ارتباط با توییتر و دریافت مجموعه داده مورد نیاز یک API از توییتر تهیه گردید. از #واکسن، #واکسن_کرونا و #واکسن_بخرید برای بدست آوردن مجموعه داده مورد نظر استفاده شده است. اینها هشتگ‌هایی هستند که پیرامون واکسیناسیون کرونا به وفور مورد استفاده قرار گرفته‌اند.

گراف جهت‌داری که از این مجموعه داده بدست آمد شامل ۱۰ هزار گره و بیش از ۲۵ هزار یال است که در مدت ۱۰ روز غیر متوالی جمع‌آوری شده است. نودهای این شبکه نام‌های کاربری در توییتر هستند. اگر کاربران در پیام خود از هشتگ‌های مذکور استفاده کرده باشند و با یکدیگر در ارتباط باشند، یالی بینشان وجود دارد. اگر کاربری پیامی که یکی از هشتگ‌ها در آن است را بازنشر کرده باشد و یا به آن پاسخ داده باشد و یا اینکه پیام را همراه با اظهار نظر^۳ خود بازنشر کرده باشد، یالی بین او و کاربر نویسنده پیام وجود دارد.

برای کشف اجتماعات در این گراف جهت‌دار از الگوریتم پیمانه‌بندی Louvain [۱۲] استفاده شده است. مقدار پیمانه‌بندی پس از اجرای این الگوریتم $Q = 0.557$ بدست آمد و خروجی الگوریتم چهارده اجتماع مجزا است و در شکل ۳ هر اجتماع با یک رنگ مشخص شده است. تقریباً ۳۸ درصد گره‌ها در اجتماع صورتی، ۱۶ درصد گره‌ها در اجتماع سبز و ۹ درصد در اجتماع آبی رنگ مشاهده می‌شوند. اجتماعات مشکی و نارنجی هر

یک است اگر $C_i = C_j$ ، در غیر اینصورت مقدار آن صفر است و $m = \frac{1}{2} \sum_{i,j} A_{i,j}$. اگر کل گراف شامل تنها یک اجتماع باشد یا گره‌ها به صورت تصادفی در بین اجتماعات قرار گرفته باشند $Q = 0$ خواهد بود. هر چه مقدار Q به ۱ نزدیکتر باشد اجتماعات بهتر جدا شده‌اند ولی هیچ گاه مقدار آن یک نمی‌شود. در عمل مقدار $Q > 0$ ساختار گروهی مناسبی را نشان می‌دهد. این معیار می‌تواند مقادیر منفی را هم بپذیرد.

الگوریتم Louvain در زمان کمی پارتیشن‌های با پیمانه‌بندی بالا را در شبکه‌های بزرگ پیدا می‌کند که خروجی آن اجتماعات موجود در شبکه است [۱۲]. این الگوریتم از دو مرحله تشکیل شده است که تکرار می‌شوند. برای شبکه‌های وزن دار با N گره، در مرحله اول ابتدا هر یک از گره‌ها به یک اجتماع نگاشت می‌شوند در نتیجه به تعداد گره‌ها در شبکه اجتماع وجود دارد. سپس برای هر گره i با همسایه‌های i ، با حذف گره i از اجتماع خودش و اضافه کردن آن به اجتماع j مقدار پیمانه‌بندی محاسبه می‌گردد. بنابراین گره i به اجتماعی اضافه می‌شود که مقدار پیمانه‌بندی حداکثر شود. اگر مقدار بدست آمده برای گره i به هیچ‌وجه مثبت نباشد در آنصورت گره i در اجتماع خودش باقی می‌ماند. این فرآیند به طور مکرر و متوالی برای همه گره‌ها اعمال می‌شود تا زمانی که دیگر پیشرفتی حاصل نشود، یعنی هیچ جابه‌جایی دیگری باعث بهبود مقدار پیمانه‌بندی نشود و در اینجا مرحله اول تمام می‌شود.

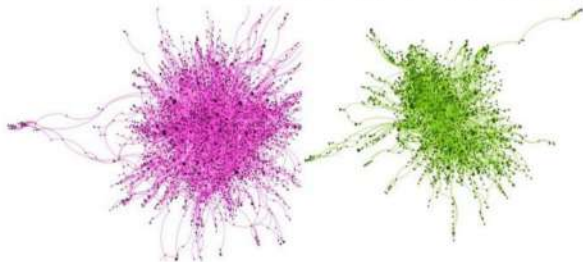
مرحله دوم الگوریتم شامل ساخت یک شبکه جدید است که گره‌های آن اکنون اجتماعات ایجاد شده در مرحله اول هستند. برای انجام این کار وزن یالها بین گره‌های جدید از مجموع وزن یالهای بین گره‌ها در دو اجتماع متناظر بدست می‌آید. یال‌های بین گره‌های عضو یک اجتماع در شبکه جدید، یال به خودی^۱ برای آن اجتماع در نظر گرفته خواهند شد. پس از اتمام مرحله دوم، مرحله اول الگوریتم دوباره روی این شبکه جدید اعمال خواهد شد. این دو مرحله را یک گذر^۲ در نظر می‌گیریم. این مراحل آنقدر تکرار می‌شوند تا اینکه تغییر بیشتری قابل اعمال نباشد و بیشترین مقدار پیمانه‌بندی حاصل گردد. شکل ۲ مراحل اجرای این الگوریتم را نشان می‌دهد. این الگوریتم از مرتبه n است و سرعت بالایی در کشف اجتماعات در شبکه‌های بسیار بزرگ دارد.

¹ Self-loop

² Pass

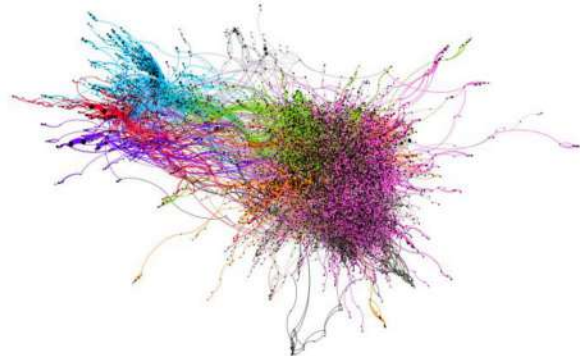
³ Quote

هر اجتماعی به یک جنبه از موضوع پرداخته است. برای مثال در اجتماع نارنجی پیام‌ها پیرامون واکسن زدن شهاب حسینی است و در اجتماع بنفش پیام‌ها پیرامون مطالبه مردم از دولت برای تهیه واکسن. همچنین بررسی این شبکه نشان می‌دهد در پیام‌های زیادی از هشتک‌های سیاسی در کنار هشتک‌های مربوط به واکسیناسیون استفاده شده است؛ به عبارت دیگر جامعه موضوع واکسن کرونا را یک موضوع سیاسی می‌بیند.



(شکل-۴): اجتماعات صورتی و سبز

کدام حدود ۷ درصد از گره‌ها را شامل می‌شوند و مابقی گره‌ها تقریباً بطور یکسان در بقیه اجتماعات پخش شده‌اند.



(شکل-۳): اجتماعات شبکه حاصل از گفتگوها پیرامون واکسیناسیون کرونا

گره‌هایی که اندازه آنها بزرگتر است درجه بیشتری دارند. در اجتماع صورتی که بزرگترین اجتماع است، گفت‌گوهای کاربران در کنار واکسیناسیون بیشتر اعتراض به حاکمیت است. در این اجتماع #تغییر، #پیمان_نوین و #رای_بی_رای به وفور در پیام کاربران مورد استفاده قرار گرفته است. در این اجتماع کاربرانی همچون @fariborzabdoll8، @masood202000 و @yin87ml چند کاربر دیگر نقش کانون دارند و پیام‌های آنها زیاد بازنشر شده است. در اجتماع سبز #ایران، #نوروز و #نوروز_بازندانان به کرات در پیام کاربران دیده می‌شود و @parvizkhazaei1، @victorious_20 و @sonofthebeach33 نقش کانون دارند. شکل ۴ دو اجتماع صورتی و سبز را نشان می‌دهد. در اجتماع آبی از #کرونا زیاد استفاده شده است و گفتگوها پیرامون ویروس کرونا بوده است. در اجتماع مشکی کاربر @pahlavireza کانون است و غالب پیام‌ها در این اجتماع مربوط به طرفداری از خاندان پهلوی است. در اجتماع نارنجی #شهاب_حسینی به دلیل مباحثی که پیرامون واکسن زدن ایشان پیش آمد، زیاد مورد استفاده قرار گرفته است. در این اجتماع #سیاسی_کاری، #تبعیض و #سلبریتی توسط کاربران زیادی استفاده شده است. در اجتماع قرمز #واکسن و در اجتماع بنفش #واکسن_کرونا در پیام‌های کاربران زیاد دیده می‌شود. در این اجتماع #واکسن_کرونا_مطالبه_ملی و #واکسن_تأییدشده_بخیرید نیز در پیام کاربران به وفور وجود دارد. با بررسی اجتماعات این شبکه مشاهده می‌شود که اجتماعات از نوع دو قطبی، منسجم، همه‌پخشی و یا پشتیبانی نیستند. اجتماعات این شبکه بیشتر شبیه دسته‌بندی خوشه‌های اجتماع است که در آن گفتگوها پیرامون واکسن کرونا است ولی

۵- نتیجه‌گیری

در این تحقیق ابتدا انواع اجتماعات در شبکه اجتماعی توییتر مورد بررسی قرار گرفت. این اجتماعات شامل اجتماع دو قطبی، اجتماع منسجم، اجتماع برند، خوشه‌های اجتماع، اجتماع همه‌پخشی و اجتماع پشتیبانی بودند. سپس سه دسته الگوریتم تشخیص اجتماعات مورد بررسی اجمالی قرار گرفت. این الگوریتم‌ها عبارت بودند از الگوریتم‌های تقسیم کننده، الگوریتم‌های تجمعی و الگوریتم‌های بیشینه‌سازی پیمان‌بندی. در نهایت به بررسی اجتماعات در شبکه حاصل از گفتگوها حول واکسن کرونا در توییتر پرداخته شد.

همانطور که پیش از این اشاره شد در این شبکه هشتک‌های دیگری در کنار #واکسن_کرونا به وفور مورد استفاده قرار گرفته‌اند که نشان دهنده دغدغه جامعه در خصوص واکسیناسیون کرونا است. برای مثال #واکسن_کرونا_مطالبه_ملی و #واکسن_تأییدشده_بخیرید در پیام‌های زیادی مورد استفاده قرار گرفته‌اند؛ در نتیجه دولت باید به دغدغه جامعه در خصوص تهیه واکسن‌های تأییدشده توجه نماید. از دیگر حساسیت‌های جامعه که از تحلیل این شبکه بدست می‌آید در خصوص رفتار سلبریتی‌ها است. سلبریتی‌ها افراد مشهور و محبوب در جامعه هستند و رفتار آنها زیر ذره‌بین جامعه است. پس از واکسن زدن آقای شهاب حسینی در امریکا هشتک‌های #شهاب_حسینی، #سیاسی_کاری، #تبعیض و #سلبریتی در پیام‌های زیادی مشاهده شده‌اند. مشاهده برخورد سیاسی جامعه با مسئله واکسیناسیون یکی دیگر از خروجی‌های این تحلیل بود. در واقع جامعه مسئله واکسیناسیون

- national academy of sciences* 99, no. 12 (2002): 7821-7826.
- [7] Newman, Mark EJ, and Michelle Girvan. "Finding and evaluating community structure in networks." *Physical review E* 69, no. 2 (2004): 026113.
- [8] Radicchi, Filippo, Claudio Castellano, Federico Cecconi, Vittorio Loreto, and Domenico Parisi. "Defining and identifying communities in networks." *Proceedings of the national academy of sciences* 101, no. 9 (2004): 2658-2663.
- [9] Pons, Pascal, and Matthieu Latapy. "Post-processing hierarchical community structures: Quality improvements and multi-scale view." *Theoretical Computer Science* 412, no. 8-10 (2011): 892-900.
- [10] Newman, Mark EJ. "Fast algorithm for detecting community structure in networks." *Physical review E* 69, no. 6 (2004): 066133.
- [11] Clauset, Aaron, Mark EJ Newman, and Cristopher Moore. "Finding community structure in very large networks." *Physical review E* 70, no. 6 (2004): 066111.
- [12] Blondel, Vincent D., Jean-Loup Guillaume, Renaud Lambiotte, and Etienne Lefebvre. "Fast unfolding of communities in large networks." *Journal of statistical mechanics: theory and experiment* 2008, no. 10 (2008): P10008.
- [13] "Gephi - The Open Graph Viz Platform." <https://gephi.org/> (accessed Feb. 23, 2021).



محمد بیات مدرک کارشناسی خود را در رشته ادبیات انگلیسی در سال ۱۳۸۵ از دانشگاه قم و مدرک ارشناسی ارشد خود را در رشته مهندسی فناوری اطلاعات گرایش شبکه‌های کامپیوتری در سال ۱۳۹۴ از دانشگاه امیرکبیر اخذ کرده است. ایشان در حال حاضر

دانشجوی دکتری مهندسی فناوری اطلاعات گرایش شبکه‌های کامپیوتری در دانشگاه تهران هستند زمینه‌های پژوهشی مورد علاقه ایشان اینترنت اشیا، محاسبات ابری و محاسبات مه می‌باشد. نشانه رایا نامه ایشان عبارتند از: m.bayat110@ut.ac.ir

را جدای از مسائل سیاسی روز نمی‌بیند و می‌توان نتیجه گرفت که یکی از آفت‌های جامعه ما نگرش سیاسی به هر موضوعی است حتی موضوعاتی که در ارتباط مستقیم با سلامت جامعه هستند. با بررسی پیام‌ها و هشنگ‌ها می‌توان احساسات جامعه پیرامون یک موضوع مشخص را نیز کشف کرد. بررسی پیام‌ها و هشنگ‌ها در این اجتماعات نشان دهنده ناراحتی، نارضایتی و خشم جامعه نسبت به عملکرد دولت در خصوص واکنش‌های است. در صورتی که دولت‌ها به نارضایتی و خشم جامعه بی‌توجه باشند ممکن است جامعه دچار تنش شود که عواقب خطرناکی خواهد داشت.

مطالعه بر روی اجتماعات در شبکه‌های اجتماعی می‌تواند به جامعه شناسان در خصوص وضعیت و رویکرد جامعه در موضوعات مختلف کمک کند. ممکن است بتوان اجتماعات دیگری بجز شش اجتماع مذکور در شبکه اجتماعی توییتر کشف کرد که نیاز به تحقیقات بیشتری دارد. همچنین بررسی اجتماعات در شبکه‌های اجتماعی همچون اینستاگرام و فیسبوک که اغلب پست‌ها شامل عکس و فیلم است می‌تواند اجتماعات دیگری را نشان دهد.

۸ - مراجع

- [1] Albert, Réka, and Albert-László Barabási. "Statistical mechanics of complex networks." *Reviews of modern physics* 74, no. 1 (2002): 47.
- [2] Azaouzi, Mehdi, Delel Rhouma, and Lotfi Ben Romdhane. "Community detection in large-scale social networks: state-of-the-art and future directions." *Social Network Analysis and Mining* 9, no. 1 (2019): 1-32.
- [3] Brandes, Ulrik, Daniel Delling, Marco Gaertler, Robert Gorke, Martin Hoefer, Zoran Nikoloski, and Dorothea Wagner. "On modularity clustering." *IEEE transactions on knowledge and data engineering* 20, no. 2 (2007): 172-188.
- [4] Newman, Mark EJ. "Communities, modules and large-scale structure in networks." *Nature physics* 8, no. 1 (2012): 25-31.
- [5] Smith, M. A., L. Rainie, B. Shneiderman, and I. Himelboim. "Mapping twitter topic networks: From polarized crowds to community clusters. Pew Research Internet Project." (2014).
- [6] Girvan, Michelle, and Mark EJ Newman. "Community structure in social and biological networks." *Proceedings of the*



خانم دکتر حشمتی مدرک کارشناسی و کارشناسی ارشد خود را در رشته مهندسی الکترونیک، مخابرات و کامپیوتر از دانشگاه برادفورد و مدرک دکتری خود را در رشته مهندسی مخابرات از دانشگاه لیدز اخذ کرده‌اند. ایشان در حال حاضر عضو هیئت علمی

دانشکده علوم و فنون دانشگاه تهران می‌باشد. زمینه تحقیقاتی مورد علاقه ایشان تحلیل شبکه‌های پیچیده است.

روش ارجاع به مقاله : م. بیات، ز. الهدی حشمتی. کشف اجتماعات در شبکه تشکیل شده از گفتگوها پیرامون واکسن کرونا در توئیتر ، دوفصلنامه محاسبات و سامانه های توزیع شده، سال سوم، شماره اول، شماره پیاپی ۵، صفحه ۶۶ تا ۷۵، سال ۱۳۹۹