رویکردی برای تحلیل آسیب‌پذیری شبکه‌های اجتماعی مبتنی بر عملکرد با استفاده از ضریب خوشه‌بندی

منصوره میرزایی[۱]، استادیار؛ مریم نورائی آباده[۲]، دکتری نرم‌افزار

۱- دانشکده مهندسی برق و کامپیوتر - دانشکده فنی و مهندسی گلپایگان - گلپایگان- ایران – mirzaie@gut.ac.ir
۲- گروه مهندسی کامپیوتر - واحد آبادان - دانشگاه آزاد اسلامی- آبادان- ایران - mnooraei@iauabadan.ac.ir

چکیده: توانمندی در واکنش به رویدادهای غیرمنتظره همواره برای شبکه‌های دنیای واقعی مطلوب است. به‌منظور بهبود توانمندی هر نوع سیستم شبکه، تجزیه و تحلیل آسیب‌پذیری برای اختلالات خارجی از قبیل نقص تصادفی یا حملات دفاعی که به عناصر شبکه وارد می‌شوند حائز اهمیت است. در این مقاله، یک مسئله نوظهور در ارزیابی توانمندی شبکه‌های پیچیده را بررسی می‌کنیم: آسیب‌پذیری خوشه‌بندی شبکه‌های مبتنی بر عملکرد در برابر فقدان عناصر شبکه. هدف اصلی شناسایی رئوسی است که فقدان آن‌ها به‌واسطه تضعیف خوشه‌بندی، به‌طور قابل‌توجهی به شبکه آسیب می‌رساند که از طریق میانگین ضریب خوشه‌بندی مورد ارزیابی قرار می‌گیرد. این مسئله به این دلیل حائز اهمیت است که هر تغییر قابل‌ملاحظه‌ای ناشی از نقص عناصر که منجر به تغییر خوشه‌بندی می‌شود می‌تواند عملکرد شبکه، مانند توانایی انتشار اطلاعات در یک شبکه اجتماعی را کاهش دهد. ما این تحلیل آسیب‌پذیری را به‌عنوان یک مسئله بهینه‌سازی تنظیم می‌کنیم و کامل بودن و عدم یکنواختی آن را نشان می‌دهیم. درنهایت، آزمایش‌های جامعی را در شبکه‌های اجتماعی ساختگی و واقعی که توسط مدل‌های شناخته‌شده تولیدشده‌اند، انجام می‌دهیم. نتایج تجربی در مقایسه با استراتژی‌های مختلف در شبکه‌های ترکیبی و واقعی نشان می‌دهد که میانگین ضریب خوشه‌بندی در تحلیل نقص گره‌های شبکه بسیار کارآمد است . همچنین نتایج به‌دست‌آمده تائید می‌کند که تکنیک حذف گره‌های پراهمیت به ویژه از نظر مقدار مرکزیت نزدیکی، در تجزیه و تحلیل آسیب‌پذیری خوشه‌بندی بسیار مؤثر است.

واژه‌های کلیدی: شبکه‌های اجتماعی، ضریب خوشه‌بندی، آسیب‌پذیری، معیارهای مرکزیت.

An Approach to Analyze the Vulnerability of Function-Based Social Networks Using Clustering Coefficient

Mansooreh Mirzaie[1], Assistant Professor; MaryamNooraei Abadeh[2], PhD in Software Engineering[2]

1- Faculty of Electrical and Computer Engineering, Golpayegan University of Technology,Golpayegan, Iran, Email: mirzaie@gut.ac.ir
2- Department of Computer Engineering, Abadan Branch, Islamic Azad University, Abadan, Iran, Email: mnooraei@iauabadan.ac.ir

Abstract: Robustness in response to unexpected events is always ideal for real-world networks. In order to improve the robustness of any network system, it is important to analyze the vulnerability to external interference such as accidental fault or attacks that are introduced into the network elements. In this paper, a novel study investigates the robustness of complex network using the clustering coefficient of networks against loss of elements. In particular, we can identify vertexes that can damage the network due to the weakening of its clustering, which is evaluated by means of the average of clustering coefficient. This is important because any tangible change that leads to a clustering is caused by defects that can reduce network functionality, such as the ability to spread information on a social network. We present this risk analysis as an optimization method and demonstrate the completeness and uncertainty of our approach to identify main vertices for clustering. Finally, we perform comprehensive experiments in the combined social networks that are generated by different models. The experimental results show that the average clustering coefficient is very efficient in analyzing the fault of network node failure. Also, the results confirm that the technique of removing the important nodes, especially in terms of the closeness centrality, is very effective in the analysis of clustering vulnerability.

Keywords: Social Network, Clustering Coefficient, Vulnerability, Centrality Measure.

تاریخ ارسال مقاله: ۱۳۹۷/۰۶/۱۴
تاریخ اصلاح مقاله: ۱۳۹۷/۰۹/۰۶
تاریخ پذیرش مقاله: ۱۳۹۷/۱۲/۱۴
نام نویسنده مسئول: منصوره میرزایی
نشانی نویسنده مسئول: ایران - گلپایگان - دانشگاه گلپایگان - دانشکده مهندسی برق و کامپیوتر.

۱- مقدمه

انعطاف‌پذیری شبکه در برابر حملات و نواقص، در زمان‌های اخیر دغدغه‌ای روزافزون بوده است. توانمندی شاید یکی از ویژگی‌های مطلوب برای شبکه‌های پیچیده ذاتی مانند شبکه جهانی وب، شبکه‌های حمل و نقل، شبکه‌های ارتباطی، شبکه‌های زیستی و شبکه‌های اطلاعاتی اجتماعی است. به طور خلاصه، توانمندی یک شبکه بررسی می‌کند که تا چه میزان عملکرد طبیعی شبکه در صورت بروز اختلالات خارجی تحت تأثیر قرار می‌گیرد، یعنی انعطاف‌پذیری شبکه را در واکنش به رویدادهای غیرمنتظره مانند حملات دفاعی و نواقص تصادفی اندازه‌گیری می‌کند.

برخی از پژوهش‌ها مسئله آسیب‌پذیری را با رویکردی متفاوت بررسی کرده‌اند. به عنوان مثال نویسندگان در [۱] تشکیل گروه در شبکه‌های اجتماعی با هدف پیدا کردن بهترین رهبر و تیمی از متخصصان به طوری که بتوانند در جهت رسیدن به هدفی مشترک، با یکدیگر همکاری داشته باشند راب با استفاده از توابع چندهدفه تعریف کردند.

مفهوم آسیب‌پذیری به‌طورکلی برای درک و توصیف عدم توانمندی و انعطاف‌پذیری سامانه‌های پیچیده مورداستفاده قرارگرفته است [۲] و [۳]. به‌منظور بهبود توانمندی سیستم‌های دنیای واقعی، لازم است که بینش عمده‌ای نسبت به آسیب‌پذیری‌های ساختاری شبکه‌هایی که آن‌ها را ارائه می‌دهند، به دست آید. یکی از جنبه‌های اصلی آسیب‌پذیری، تجزیه‌وتحلیل و درک تأثیر نقص (به‌صورت عمدی یا تصادفی) عناصر شبکه بر میزان خوشه‌بندی در شبکه است.

سیستم‌های پیچیده بسیاری برای بررسی میزان آسیب‌پذیری در برابر خطاها طراحی شده‌اند. میزان تحمل آسیب‌پذیری در شبکه‌های ارتباطی[۲] به‌ویژه شبکه‌های اجتماعی از اهمیت ویژه‌ای در تداوم ساختار شبکه برخوردار است. هرچند که اجزای اصلی ممکن به‌طور مرتب دچار خرابی شوند، اما خرابی گره‌های محلی بندرت سبب از دست رفتن اطلاعات می‌شود [۴]. توانمندی این سیستم‌ها و دیگر سیستم‌های پیچیده اغلب به ساختار شبکه و چگونگی اتصال و وابستگی اجزای شبکه است. هنگامی‌که گره‌ها از یک شبکه حذف می‌شوند، خوشه‌هایی از گره‌ها که پیوندهایشان با سیستم مورد تأثیر قرارگرفته، ممکن است از خوشه اصلی جدا شوند.

خوشه‌بندی[۳] یک ویژگی پایه در شبکه است که می‌تواند مرتبط با طیف وسیعی از عناوین باشد. مثلاً انتشار اطلاعات در یک شبکه اجتماعی را در نظر بگیرید (مانند انتشار یک شایعه). هر چه ساختار خوشه‌بندی شبکه بهتر باشد، انتشار اطلاعات در آن راحت‌تر انجام می شود [۵]. اهمیت خوشه‌بندی محدود به شبکه‌های اجتماعی نیست و به‌عنوان نمونه در شبکه‌های حمل و نقل هوایی هم کاربرد دارد. خوشه‌بندی خوب برای این شبکه‌ها می‌تواند به مسافران پروازهای لغو شده کمک کند تا مسیرهای جایگزین را آسان‌تر انتخاب کنند. در این تحقیق، ما از میانگین ضریب خوشه‌بندی محلی به عنوان معیاری برای تعریف و اندازه‌گیری خوشه‌بندی در شبکه استفاده می‌کنیم. برای درک بهتر اثرات شکست و آسیب‌پذیری ساختار شبکه، در ادامه این روند آسیب‌پذیری را تحلیل خواهیم کرد. پرسش اصلی تحقیق در این مقاله این است که: در یک شبکه پیچیده که ضریب خوشه‌بندی آن قابل محاسبه است، مهم‌ترین رئوسی که خطای آن‌ها در حمله داخلی و یا اتفاقی، خوشه‌بندی شبکه را تنزل می‌بخشند کدام رئوس هستند؟

کارهای اصلی انجام‌شده در پژوهش عبارت‌اند از:

- تحلیل ساختاری شبکه‌های اجتماعی و ضرورت بررسی آسیب‌پذیری آن‌ها.
- تعریف ارزیابی آسیب‌پذیری خوشه‌بندی در شبکه‌های پیچیده.
- فرموله کردن معیار فوق به‌عنوان یک مسئله بهینه‌سازی به‌عنوان یک تابع هدف.
- انجام آزمایش‌های جامع بر روی شبکه‌های مصنوعی و شبکه‌های واقعی.

این مقاله از بخش‌های زیر تشکیل‌شده است. بخش ۲ ادبیات تحقیق را بررسی می‌کند. بخش ۳ به‌مرور کارهای مرتبط می‌پردازد. بخش ۴ نمادها، توابع اندازه‌گیری و تعریف مسئله را شرح می‌دهد. بخش ۵ تجزیه و تحلیل مسئله با استفاده از رفتار خوشه‌بندی در مورد حذف تصادفی و غیر تصادفی گره‌ها را ارائه می‌دهد و همچنین نتایج تجربی رویکرد خود را در مقایسه با سایر استراتژی‌ها گزارش می‌دهیم. در بخش ۶ به جمع‌بندی و بیان راهکارهای آتی می‌پردازیم.

۲- ادبیات پژوهش

۲-۱- مروری بر شبکه‌های اجتماعی

شبکه اجتماعی[۴] [۶] نوعی شبکه است که گره‌های آن افراد و یال‌های آن وابستگی متقابل بین آن‌ها را نشان می‌دهد، مانند شبکه دوستی افراد، همکاری نویسندگان مجله و همکاری میان احزاب مختلف. یک شبکه اجتماعی می‌تواند به عنوان یک شبکه پیچیده که از افراد یک جامعه و روابط میان اشخاص حقیقی تشکیل شده است، مورد استفاده قرار گیرد. اندازه این شبکه‌ها معمولاً بسیار بزرگ است. این نوع ساختار اجتماعی پیچیده نقش مهمی در دسترسی و انتشار اطلاعات دارد. شبکه‌های اجتماعی از لحاظ نوع کاربرد، به شبکه‌های اجتماعی مبتنی بر توپولوژی و شبکه‌های اجتماعی مبتنی بر عملکرد دسته‌بندی می‌شوند [۷].

۲-۱-۱- شبکه‌های اجتماعی مبتنی بر توپولوژی

بر اساس توپولوژی شبکه‌ها، می‌توان شبکه‌های اجتماعی را به شبکه‌های اجتماعی برخط و شبکه‌های اجتماعی موبایل[۵] طبقه‌بندی کرد:

(۱) شبکه‌های اجتماعی برخط: شبکه‌های اجتماعی برخط[۶]جوامع مجازی هستند که به افراد این امکان را می‌دهند که به یکدیگر متصل شده و پیرامون یک موضوع خاص و یا صرفاً "گذران وقت" به‌صورت برخط با هم ارتباط برقرار کنند. با توجه به اطلاعات جمع‌آوری‌شده در آوریل ۲۰۱۸ [۸]، ۲/۲۳ میلیارد کاربر به سایت فیس‌بوک دسترسی دارد.

(۲) شبکه‌های اجتماعی موبایل: با توسعه قابل‌توجه فن‌آوری‌های شبکه بیسیم مانند شبکه سلولی، وایفای و بلوتوث، برنامه‌های کاربردی شبکه برخط اینترنتی از شبکه‌های برخط به سمت شبکه‌های تلفن همراه سوق داده‌شده است. با توجه به مزایای غیر قابل انکار آن، ازجمله استقرار و اتصال آسان و ارزان بین مکان فیزیکی کاربران و تعاملات اجتماعی آن‌ها، برنامه‌های کاربردی اجتماعی موبایل بسیار جذاب هستند.

۲-۱-۲- شبکه‌های اجتماعی مبتنی بر عملکرد

شبکه‌های اجتماعی مبتنی بر عملکرد به شش دسته تقسیم می‌شوند.

(۱) اخبار اجتماعی: یک وب سایت خبری اجتماعی ویژگی‌های نوشته‌شده توسط کاربر است که بر اساس محبوبیت رتبه‌بندی شده است. کاربران می‌توانند برای هر پست اظهارنظر کنند و این نظرات نیز ممکن است رتبه‌بندی شوند، مانند Reddit.

(۲) پیام فوری[۷]: پیام فوری یک نوع سکو چت برخط است که انتقال متن در زمان واقعی را در اختیار شما قرار می‌دهد.

(۳) وبلاگ نویسی در ابعاد کوچک[۸] : این امکان را به کاربران می‌دهد تا عناصر جزئی محتوایی را مانند جملات کوتاه، تصاویر فردی، یا لینک‌های ویدئویی را با هم مبادله کنند. توییتر یک پلت فرم وبلاگ نویسی در ابعاد کوچک است که به کاربران ثبت شده اجازه می‌دهد تا متن کوتاهی از طریق واسط‌هایی مانند وب و پیامک منتشر شود.

(۴) نشانه‌گذاری اجتماعی[۹]: یک سرویس برخط متمرکز است که به کاربران اجازه می‌دهد نشانه‌گذاری‌های اسناد وب را اضافه، ویرایش، حاشیه‌نویسی کرده و یا به اشتراک بگذارند .

(۵) انجمن اینترنتی: یک سایت مباحثه‌ای برخط است که مردم از تالارهای گفتمان متفاوت می‌توانند در قالب ارسال پیام مکاتبه داشته باشند.

(۶) وب سایت به اشتراک‌گذاری رسانه: اجازه می‌دهد تا کاربران برای ارسال، مشاهده، ارسال نظرات و به اشتراک‌گذاری اطلاعات مختلف چندگانه، مانند فایل‌ها، فیلم‌ها، عکس‌ها، موسیقی‌ها، فیلم‌ها.

۲-۲- آسیب‌پذیری در شبکه‌های اجتماعی

انعطاف‌پذیری شبکه در برابر حمله‌ها و خطاها امروزه به موضوعی قابل‌بحث تبدیل‌شده است. شناسایی گره‌های آسیب‌پذیر نقش بسیار مهمی در آنالیز شبکه‌های پیچیده دارد. بررسی گره‌های آسیب‌پذیر در شبکه‌های اجتماعی باهدف یک استراتژی کاهشی کارا و مؤثر جهت کاهش تأثیرات شکست‌های آبشاری و حفظ اتصالات شبکه‌های اجتماعی امری ضروری است. توانمندی شاید مهم‌ترین ویژگی مطلوب برای شبکه‌های پیچیده مثل www، شبکه‌های حمل‌ونقل، شبکه‌های ارتباطی، شبکه‌های بیولوژیکی و شبکه‌های اطلاعاتی و اجتماعی باشد. منظور از انعطاف‌پذیری این است که شبکه تا چه اندازه در برابر آشفتگی های خارجی و رویدادهای غیرمنتظره مثل حمله‌های دشمن و خطاهای تصادفی، به‌صورت نرمال عمل می‌کند [۹].

تشخیص گره‌های آسیب‌پذیر در شبکه‌های مبتنی بر عملکرد از اهمیت ویژه‌ای برخوردار است. اگرچه تعداد گره‌های تأثیرگذار زیاد نخواهد بود، اما تأثیر آن‌ها می‌تواند به‌سرعت در سراسر شبکه گسترش یابد و تأثیرات ناشی از فقدان عناصر کلیدی می‌تواند عملکرد شبکه، مانند توانایی انتشار اطلاعات در یک شبکه اجتماعی را کاهش دهد. همان‌طور که گفته شد در این پژوهش سعی داریم تا روش جدیدی برای شناسایی گره‌های تأثیرگذار در شبکه‌های اجتماعی مبتنی بر عملکرد با استفاده از ضریب خوشه‌بندی به‌عنوان یک مسئله بهینه سازی ارائه دهیم. برای این منظور نیاز است معیارهای مرکزیت شبکه معرفی شوند که در بخش بعدی مورد بررسی قرار می‌گیرند.

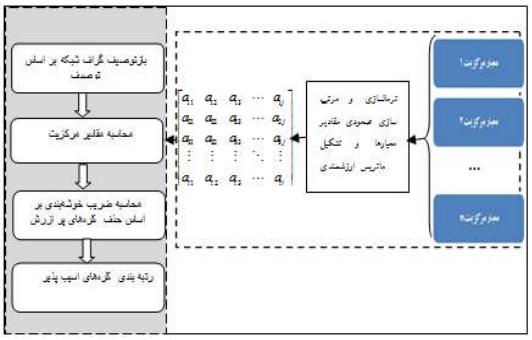
۳- کارهای مرتبط

ارزیابی آسیب‌پذیری توجه زیادی را از جامعه علمی شبکه به خود جلب کرده است. تحقیق در پیشینه ادبی می‌تواند به دودسته تقسیم شود: اندازه‌گیری توانمندی و دست‌کاری آن در یک شبکه. در اندازه‌گیری توانمندی، اندازه‌گیری‌ها و معیارهای مختلفی مانند اتصال گراف [۱۰]، قطر، اندازه نسبی بزرگترین عناصر و اندازه متوسط خوشه جداشده[۱۱]پیشنهاد شده‌اند. ازنظر دست‌کاری توانمندی، استراتژی های مختلفی نظیر حذف گره‌ها بر مبنای درجه [۱۲] یا مدل‌سازی حذف گره بانفوذ گراف [۱۳]پیشنهاد شده‌اند. مطالعات دیگر به حذف گره‌ها با اندازه‌گیری‌های مرکزی [۱۴]یا کوتاه‌ترین مسیر بین جفت گره‌ها [۱۵]یا بررسی کل ارتباط دو جهته [۱۶]پرداخته‌اند. برای کسب اطلاعات بیشتر در مورد ارزیابی آسیب‌پذیری‌های شبکه، خواننده متن به مطالعات مروری چن [۴]، گومز و همکاران [۱۷]و منابع موجود در آن ارجاع داده می‌شود.

بررسی آسیب‌پذیری بر اساس میانگین خوشه‌بندی محلی (ALCC)[۱۰] [۵] برای یک شبکه پیچیده، حوزه‌ای است که بندرت مورد بررسی قرار گرفته است. در یک تحقیق مرتبط [۱۸]، نویسندگان آسیب‌پذیری ساختار جامعه را با استفاده از رویکردهای اکتشافی برای پیدا کردن عناصر حیاتی در ساختار اجتماعی تجزیه و تحلیل کنند. آلیم و همکاران [۱۹] روشی مبتنی بر ایجاد مرزهای یک جامعه پیشنهاد کردند تا عناصر حیاتی را پیدا کنند. آن‌ها مسئله شکستن همه جوامع مبتنی بر تراکم در شبکه را مورد بررسی قراردادند که سختی آن را اثبات کردند و راه‌حل‌های اکتشافی را پیشنهاد کردند. همچنین در[۹] مطالعاتی بر اساس تجزیه و تحلیل آسیب‌پذیری مبتنی بر جامعه مطرح شده است که به دلیل ابهام تعریف جامعه در یک شبکه، محدودیت‌های اساسی دارند. تحقیق ما این کمبود خاص را برطرف می‌کند، زیرا ALCC به خوبی تعریف شده است و به‌طورکلی برای اندازه‌گیری خوشه‌بندی شبکه پذیرفته شده است. ارتم و همکاران [۲۰] مسئله چگونگی تشخیص گروه‌های وابسته یک گره در یک شبکه اجتماعی را با استفاده از ضریب خوشه‌بندی بالا بررسی کردند. با این وجود، تحقیق آن‌ها آسیب‌پذیری میانگین ضرایب خوشه‌بندی محلی یک شبکه را در نظر نمی‌گیرد.

۴- رویکرد پیشنهادی

در این بخش به تعریف نمادها و تعریف مسئله و ارائه رویکرد پیشنهادی می‌پردازیم. در جدول ۱، فهرستی از نمادهای مورد استفاده در این تحقیق آورده شده است. فلوچارت رویکرد پیشنهادی در شکل ۱ آورده شده است. همان‌طور که در شکل نشان داده است ابتدا بر اساس معیارهای مرکزیت مطلوب ماتریس مرتب نرمال شده مقادیر ارزشمندی محاسبه شده و در نهایت با استفاده از ضریب خوشه‌بندی به تحلیل آسیب‌پذیری در شبکه‌های اجتماعی مبتنی بر عملکرد می‌پردازیم.



شکل ۱: رویکرد پیشنهادی

۴-۱- توصیف ساختار گرافی شبکه

ابتدا گراف شبکه بر اساس تعاریف ۱ و ۲ بازتعریف می‌شوند. گره‌ها و ارتباط بین آن‌ها شناسایی شده و ماتریس مجاورت گراف ساخته می شود. در این پژوهش از گراف‌های ساختگی و گراف‌های واقعی مبتنی بر عملکرد استفاده شده است. نمادهای اصلی روش پیشنهادی در جدول ۱ معرفی شده‌اند.

۴-۲- محاسبه ارزش مرکزیت هر گره با اندازه‌گیری‌های مرکزیت مختلف

۴-۲-۱- معیارهای مرکزیت گره‌های یک شبکه

نحوه اتصال گره‌های یک شبکه اجتماعی می‌تواند اطلاعات ارزشمندی در خصوص اهمیت آن‌ها بیان کند. برای تشخیص اهمیت گره‌ها از معیارهای مرکزیت متفاوتی استفاده می‌شود. همان‌طور که قبلاً گفته شد در این کار، سه معیار مهم درجه مرکزیت (C_D) [۱۱]، میانوندی[۱۲] (C_B) و نزدیکی[۱۳] (C_C) [۲۱] مورد بررسی قرارگرفته‌اند و نحوه محاسبه مرکزیت C_D ، C_B و C_C [۲۲] در ادامه بیان می‌شوند. البته رویکرد پیشنهادی قابل‌تعمیم به معیارهای دیگر نیز هست که در کارهای آتی بررسی می‌شوند.

تعریف ۱. گراف: یک گراف را با دوتایی $G=(V,E)$ نشان می‌دهند که V مجموعه‌ای غیر تهی و متناهی از عناصری است که گره نامیده می شوند و E مجموعه‌ای متناهی از دوتایی‌هایی از عناصر V هستند (که از ضرب دکارتی V در خودش حاصل می‌شود) که یال نامیده می‌شود.

جدول ۱: تعریف نمادهای اصلی مسئله

معنا	نشانه‌گذاری
تعداد گره‌های شبکه	N
گراف G با مجموعه رئوس V و لبه‌های E	G(V,E)
مرکزیت درجه	$\mathbf{C_D}(i)$
مرکزیت میانوندی	$\mathbf{C}_B(i)$
مرکزیت نزدیکی	$\mathbf{C}_C(i)$
تعداد کوتاه‌ترین مسیرهای دودویی بین گره i و j	g_{ij}
یال بین گره i و j	x_{ij}
فاصله میان گره‌های i و j	d_{ij}
درایه‌های ماتریس مجاورت	e_{ij}
درایه‌های ماتریس مرکزیت	a_{ij}
مجموعه همسایگان u	N(u)
تعداد مثلث‌های حاوی u	T(u)
ضرایب خوشه‌بندی u و G	C(u), C(G)
تعداد مثلث‌های موجود بین u و v	T(u,v)
زیر گراف $S \in V$ از گراف G	G[S]

تعریف ۲. زیر گراف: گراف $G1 = (V1, E1)$ را یک زیر گراف گراف $G= (V, E)$ نامیم اگر $V1 \subseteq V$ و $E1 \subseteq E$.

تعریف ۳. مرکزیت درجه: اندازه‌گیری مرکزیت درجه گره i که به صورت k_i به صورت زیر تعریف شده است :

$$\mathbf{C_D}(i) = \sum_{j=0}^{n} x_{ij} \qquad (۱)$$

که در آن i گره اصلی (مرکزی) است، درحالی‌که j نشان‌دهنده گره‌هایی است که به گره i متصل هستند. x_{ij} یال میان گره‌های i و j است و یک مقدار باینری ۰ و یا ۱ دارد. مقدار x_{ij} برابر با ۱ است اگر گره i به گره j متصل باشد ، در غیر این صورت برابر با صفر خواهد بود. مرکزیت درجه یک گره تعداد گره‌هایی است که در همسایگی آن گره وجود دارد. گره‌هایی که لبه‌های مجاورت بیشتری را دارند دارای مرکزیت درجه بیشتری است.

تعریف۴. مرکزیت میانوندی: اندازه‌گیری مرکزیت میانوندی گره i که به‌صورت b_i نشان داده شد به‌صورت زیر تعریف شده است:

$$C_B(i) = \sum_{j<k} \frac{g_{jk}(i)}{g_{jk}} \qquad (۲)$$

که در آن g_{jk} تعداد کوتاه‌ترین مسیرهای بین گره j و k است و $g_{jk}(i)$ تعداد مسیرهایی است که از گره i عبور می‌کند. معیار میانوندی نیز قابلیت عملکرد یک گره به‌عنوان واسطه بین گره‌های دیگر را نشان می دهد. به عبارت دیگر این پارامتر تعداد دفعاتی که یک گره در مسیر کوتاه‌ترین فاصله بین گره‌های دیگر قرار می‌گیرد را مشخص می‌کند.

تعریف۵ . مرکزیت نزدیکی: اندازه‌گیری مرکزیت نزدیکی گره i با c_i نشان داده‌شده که به‌صورت زیر تعریف می‌شود :

$$C_C(n_i) = \left[\sum_{j=1}^{n} d_{ij}\right]^{-1} \quad (۳)$$

که در آن d_{ij} نشان‌دهنده فاصله میان گره‌های i و j است. معیار نزدیکی بر اساس میانگین طول مسیرهایی بین یک گره با سایر گره‌ها تعریف می شود و قابلیت دسترسی‌پذیری یک گره را مشخص می‌کند.

همان‌طور که توضیح داده شد، اطلاعات در شبکه‌های مختلف با روش‌های مختلف گسترش می‌یابد. اندازه‌گیری‌های مرکزیت مناسب برای محاسبه ارزش مرکزیت هر گره بر اساس خواص شبکه ساخته‌شده در مرحله ۱ انتخاب می‌شوند. تعداد گره‌ها به‌صورت n در شبکه G محاسبه می‌شوند و اندازه‌گیری‌های انتخاب شده مرکزی به صورت $C = \{C_1, C_2, \ldots, C_m\}$ توضیح داد ه‌شده است. n گره در شبکه G بر اساس m اندازه‌گیری مرکزیت انتخابی مرتب شده و ماتریس ارزش مرکزیت به شرح زیر نشان داده شده است. که در آن a_{ij} ارزش مرکزیت گره i بر اساس اندازه‌گیری مرکزیت C_j است.

$$D = \begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1j} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2j} \\ a_{31} & a_{32} & a_{33} & \cdots & a_{3j} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{i1} & a_{i2} & a_{i3} & \cdots & a_{ij} \end{bmatrix}$$

به دلیل تفاوت روش‌های مرکزیت متفاوت برای به دست آوردن مقادیر مرکزیت، ماتریس D باید برای ابعاد یکپارچه بر اساس رابطه (۴) نرمال‌سازی شود.

$$a_{ij} = \frac{a_{ij}}{\sqrt{\sum_{i=1}^{m} {a_{ij}}^2}} \qquad i \in 1,2 \ldots m \ ; \ j \in 1,2 \ldots n \quad (۴)$$

مقادیر ماتریس به دست آمده برای به دست آوردن ضریب خوشه‌بندی جدید و در نتیجه آن تحلیل آسیب‌پذیری مورد استفاده قرار خواهد گرفت که در ادامه شرح داده می‌شود.

۴-۳- تحلیل آسیب‌پذیری خوشه‌بندی بر اساس ضریب خوشه‌بندی و خطای گره

اغلب مطالعات در رابطه با آسیب‌پذیری شبکه بر این تمرکز دارند که شبکه در زمانی که عناصرش (گره‌ها و یال‌ها) حذف شوند چه رفتاری نشان خواهد داد. برای این منظور از ارتباطات متقابل، ارتباطات طبیعی و یا پارامترهای مرکزیت مثل مرکزیت درجه و میانوندی استفاده شده است. هیچ‌کدام از مطالعات موجود ضریب خوشه‌بندی میانگین را از منظر آسیب‌پذیری مورد بررسی قرار نداده‌اند (همان‌طور که گفته شد خطای یک گره می‌تواند تأثیرات شدیدی بر عملکرد شبکه داشته باشد). بنابراین نیاز به تحلیل و بررسی آسیب‌پذیری خوشه‌بندی در شبکه‌های پیچیده احساس می‌شود. این مسئله با توجه به‌اندازه شبکه‌های واقعی یک مسئله NP-Complete است[۹]. در این کار ما از ضریب خوشه‌بندی میانگین استفاده می‌کنیم و خوشه‌بندی یک شبکه را ارزیابی می‌کنیم. تشخیص عناصری که خوشه‌بندی را تحت تأثیر قرار می‌دهند یک نتیجه بزرگ دارد. مثلاً در مورد مسئله امنیت ملی کشور، عناصر بحرانی و مهم در شبکه‌های ارتباطی باید منابع بیشتری برای حمایت داشته باشند و در مقابل تشخیص عناصر بحرانی در یک شبکه اجتماعی مرتبط با دشمنان می‌تواند گسترش اطلاعات در چنین شبکه‌ای را محدود کند.

تعریف ۶. ارزیابی ساختار خوشه‌بندی[۱۴] (CSA): برای یک شبکه (گراف) $G(V,E)$ و یک عدد صحیح مثبت $k \leq N$، یک زیرمجموعه S^* از V با کاردینالیتی حداکثر k (تعداد اعضا) را پیدا کنید که کاهش ضریب خوشه‌بندی را بیشینه کنند:

$$\Delta C(S) = C(G) - C(G[V \backslash S]) \quad (۵)$$

$$S^* = argmax\, \Delta C(S), \qquad S \subseteq V, |S| \leq k \quad (۶)$$

هدف مسئله CSA، شناسایی رئوس حیاتی شبکه با نسبت میانگین ضریب خوشه‌بندی محلی است. پارامتر ورودی k می‌تواند به‌عنوان حداکثر تعداد نقص گره‌ها تفسیر شود که عملکرد نرمال شبکه را در برابر حملات دفاعی یا تصادفی مقاوم می‌کند. بر این اساس، وضعیت $|S| = k$ به‌طور دقیق رئوس حیاتی را شناسایی می‌کند و بدترین سناریوهایی که ممکن است اتفاق بیفتد را بررسی می‌کند.

تعریف ۷. ALCC: در تئوری گراف، ALCC میانگین برای $C(G)$، پارامتری است که نشان می‌دهد چقدر رئوس G تمایل دارند که با یکدیگر خوشه‌بندی شوند. این پارامتر به‌عنوان میانگین LCC روی همه رئوس شبکه تعریف می‌شود. $C(G)$ به‌صورت زیر تعریف می‌شود:

$$C(G) = \frac{1}{N}\sum_{u \in V} C(u) \quad (۷)$$

چون به ازای همه گره‌های u زیرمجموعه V، رابطه $0 \leq C(u) \leq 1$ برقرار است، $C(G)$ فقط می‌تواند در بازه $[0,1]$مقدار بگیرد. به عنوان مثال $C(G)=0$ است زمانی که G یک گراف خالی از مثلث (triangle-free) است و $C(G)=1$ زمانی که G یک کلونی یا مجموعه‌ای از کلونی ها است. بیشتر بودن ضریب خوشه‌بندی G به معنی نزدیکی بیشتر به یک کلونی است. . به بیان دیگر هر چه ضریب خوشه‌بندی G بیشتر باشد، گراف به‌صورت محلی بیشتر به یک دسته یا مجموعه‌ای از دسته‌های جدا شده شباهت دارد [۹].

تعریف ۸- LCC: ضریب خوشه‌بندی محلی. برای یک گره $u \in V$، d_u رأس مجاور در G تعریف‌شده است و $\frac{du(du-1)}{2}$ یال ممکن بین همه همسایه‌های u وجود دارد. $C(u)$، ضریب خوشه‌بندی محلی و این احتمال را بیان می‌کند که دو همسایه u به هم متصل باشند. به عبارتی دیگر، $C(u)$ محاسبه می‌کند که گراف استخراجی تا چه اندازه به یک کلونی نزدیک است. ضریب محلی خوشه‌بندی $C(u)$ به صورت زیر تعریف می‌شود:

$$\begin{cases} C(u) = \begin{cases} \dfrac{2T(u)}{du(du-1)} & du > 1 \\ 0 & \text{درغیراینصورت} \end{cases} \end{cases} \quad (۸)$$

واضح است که $0 \le C(u) \le 1$ به ازای تمام $u \in V$. برای هر گره u، $C_v(u)$ به ضریب خوشه‌بندی u در گراف $G[V\backslash\{v\}]$ اشاره دارد. در نهایت $T(u)$ به‌عنوان تعداد مثلث‌های حاوی هر دو رأس u و v تعریف می‌شود. هر ماتریس مجاورت G دارای درایه‌های e_{ij} است که:

$$2T(u) = \sum\sum e_{ui} e_{uj} e_{ij} \quad i, v \in V \quad (۹)$$

هنگامی‌که نواقص تصادفی رخ می‌دهد، مقدار ALCC به دلیل عدم یکنواختی ALCC، غیر قابل پیش بینی است. به این معنی که حذف گره‌ها می‌تواند به ALCC بالاتر یا پایین‌تری از گراف باقی مانده منجر شود. در این کار نشان داده می‌شود که در صورت نقص تصادفی یکنواخت، ALCC مورد انتظار قابل تحلیل است. این نتیجه همچنین نشان می‌دهد که با توجه به شبکه G، دنباله‌ای از زیر گراف Gi از G وجود دارد که مقادیر ALCC آن یک دنباله کاهشی را تشکیل می‌دهد.

در گراف G، گزاره‌های زیر صدق می‌کند:

الف) $C(G) = 0$ است اگر G یک شبکه بدون مثلث باشد.

ب) $C(G) = 1$ است اگر G یک دسته با حداقل ۳ رأس باشد یا فقط حاوی دسته‌های جدا باشد که هرکدام حداقل حاوی ۳ رأس است [۹].

۵- ارزیابی

همان‌طور که گفته شد، مزایای زیادی در مورد ALCC نسبت به سایر اندازه‌گیری‌های ساختاری وجود دارد و یکی از معیارهای معروف برای ارزیابی خوشه‌بندی شبکه است. هر چه ALCC یک شبکه بالاتر باشد، خوشه‌بندی بهتری را نشان می‌دهد. ALCC ویژگی‌های متعدد شبکه مدولار مانند پدیده‌های کوچک در مقیاس جهانی، ساختار مدولار و قطر کوچک (یا ساختار جامعه) را نشان می‌دهد و در نمودارهای متصل و غیر متصل همانند نمودارهای متراکم و پراکنده معنی‌دار است: شبکه‌های کوچک انتظار می‌رود که ضریب خوشه‌بندی کمی داشته باشند، درحالی که شبکه‌های پیچیده موجود، دارای ضریب خوشه‌بندی بالایی هستند.

به منظور تحلیل آسیب‌پذیری خوشه‌بندی بر اساس ضریب خوشه‌بندی و برای ارزیابی پارامترهای مرکزیت انتخاب‌شده، ۵ مجموعه داده ساختگی و دو مجموعه داده واقعی در نظر گرفته شده است که در ادامه بررسی می‌شوند.

۵-۱- تحلیل مجموعه داده‌ها

برای آزمایش رابطه بین آسیب‌پذیری شبکه و خوشه‌بندی، ازلحاظ تجربی از تعمیم ابعادی نمودارهای واتس-استروگاتز استفاده شده است [۵]. تحلیل عملکرد رویکرد پیشنهادی در دو بخش مجموعه داده‌های ساختگی و مجموعه داده‌های شبکه‌های واقعی مبتنی بر عملکرد انجام شده است. مشخصه‌های مجموعه داده‌های ساختگی در جدول‌های ۲ تا ۴ نشان داده شده‌اند. تولید گراف‌های ساختگی با استفاده از ویژگی‌های ماتریس مجاورت گراف‌های ساده انجام‌گرفته است:

۱-ماتریس مجاورت هر گراف ساده یک ماتریسی متقارن است که تمام درایه‌های روی قطر اصلی هر ماتریس مجاورت صفر می‌باشند.

۲-تعداد یک‌های موجود روی هر سطر (یا ستون) با درجه رأس متناظر با آن سطر (یا ستون) برابر است.

۳-همچنین تعداد کل یک‌های موجود در ماتریس مجاورت گراف ساده Gبرابر است با مجموع درجات رئوس گراف G و برابر است با 2q.

۳-تعداد یک‌های واقع بر ماتریس مجاورت متناظر با گراف ساده G از مرتبهp و اندازه q برابر است با2q و تعداد صفرهای آن برابر است p^2-2q.

جدول۲: مشخصه‌های مجموعه داده‌های ساختگی با درنظرگرفتن پارامتر مرکزیت C_D

N	تعداد یال‌ها	Max C_D	Min C_D	Max C_G	Min C_G
۵۰۰	۶۱۳۰	۱۶۰	۱۷	۰/۲۶۱	۰/۲۳۵
۱۰۰۰	۱۳۷۲۲	۲۳۵	۶۱	۰/۵۳۲	۰/۴۶۱
۱۵۰۰	۲۵۱۳۳	۲۷۲	۳	۰/۳۸۲	۰/۳۲۱
۲۰۰۰	۱۸۶۹۱۶	۵۸۲	۲۲	۰/۱	۰/۰۸
۳۰۰۰	۷۵۲۳۴۰	۷۶۹	۲۹	۰/۱۵۰	۰/۱۴۵

جدول۳: مشخصه‌های مجموعه داده‌های ساختگی با درنظرگرفتن پارامتر مرکزیت C_B

N	Max C_B	Min C_B	Max C_G	Min C_G
۵۰۰	۰/۰۴	۰/۰۰۰۹	۰/۲۶۲	۰/۲۲۸
۱۰۰۰	۰/۰۱	۰/۰۰۰۱	۰/۵۳۱	۰/۴۴۸
۱۵۰۰	۰/۳۸	۰/۰۰۳	۰/۳۸۳	۰/۳۱۱
۲۰۰۰	۰/۶۵	۰/۴۳	۰/۰۹۹	۰/۰۸۲
۳۰۰۰	۰/۷۶	۰/۳۸	۰/۱۵۱	۰/۱۴۴

جدول۴: مشخصه‌های مجموعه داده‌های ساختگی با درنظرگرفتن پارامتر مرکزیت CC

N	Max C_C	Min C_C	Max C_G	Min C_G
۵۰۰	۰/۷۳۱	۰/۶۱۲	۰/۲۴۸	۰/۲۱۱
۱۰۰۰	۰/۶۸	۰/۶۲	۰/۵۰۳	۰/۴۱۵
۱۵۰۰	۳/۴	۲	۰/۳۷۶	۰/۳۰۲
۲۰۰۰	۳۴	۲۴/۵	۰/۰۹۹	۰/۰۸۵
۳۰۰۰	۴۴	۱۴	۰/۱۵۰	۰/۱۴۳

در فاز تحلیل برای همه گره‌های شبکه مقادیر C_Dو C_B و C_C محاسبه و نرمال‌سازی شدند. سپس برای هر شبکه k گره با بزرگترین مقادیر هر یک از این ویژگی‌ها از شبکه خارج می‌شوند. به‌منظور تحلیل آسیب‌پذیری استراتژی‌های زیر در نظر گرفته شده است:

Random fail: حذف گره‌ها به‌صورت تصادفی

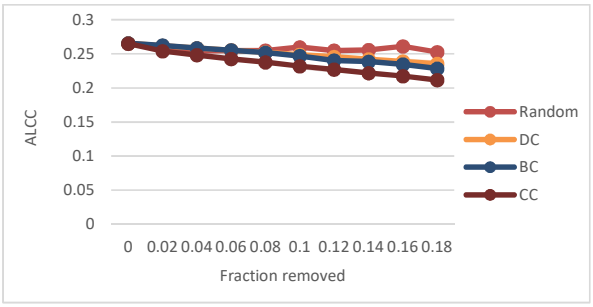
BC_greedy: حذف گره‌ها با یک روش حریصانه با توجه به بیشترین مقادیر معیار ارزشمندی C_B

DC_greedy: حذف گره‌ها با یک روش حریصانه با توجه به بیشترین مقادیر ارزشمندی C_D

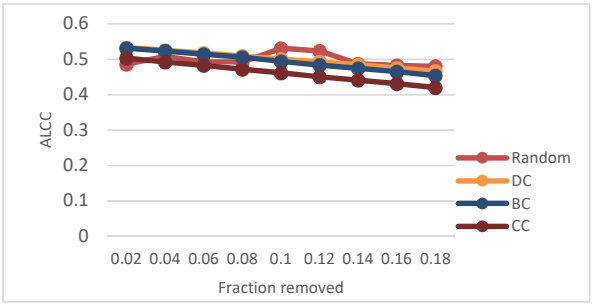
CC_greedy: حذف گره‌ها با یک روش حریصانه با توجه به بیشترین مقادیر ارزشمندی C_C

نتایج ارزیابی برای داده‌های ساختگی در نمودارهای شکل ۲ تا شکل ۶ نمایش داده شده است.

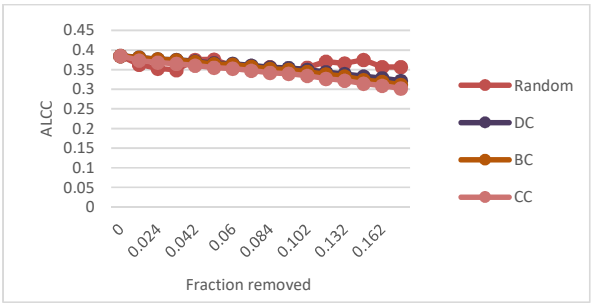
مقادیر افقی در نمودارها نشان دهنده درصد حذف‌شده از کل گره‌های شبکه و مقادیر عمودی بیانگر مقدار میانگین ضریب خوشه‌بندی است.



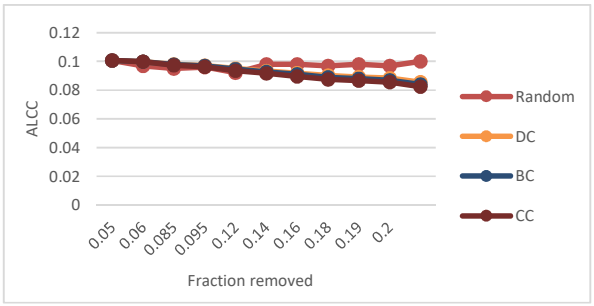
شکل۲. مقادیر ضریب خوشه‌بندی برای N= ۵۰۰



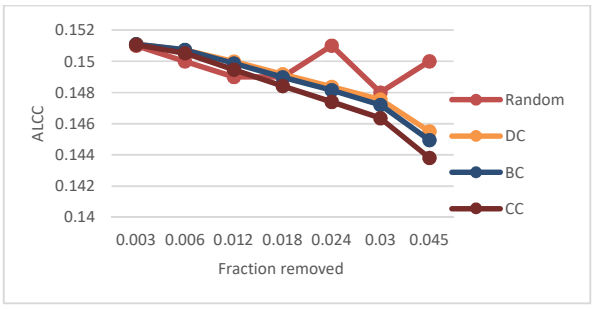
شکل۳. مقادیر ضریب خوشه‌بندی برای N=۱۰۰۰



شکل ۴. مقادیر ضریب خوشه‌بندی برای N=۱۵۰۰



شکل۵. مقادیر ضریب خوشه‌بندی برای N=۲۰۰۰

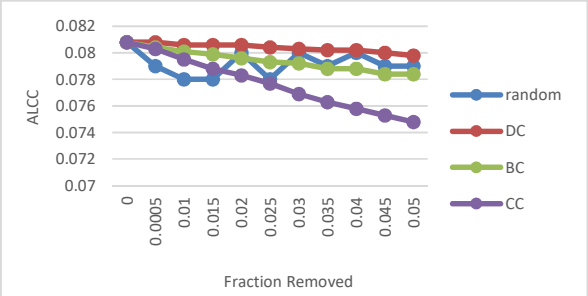


شکل۶. مقادیر ضریب خوشه‌بندی برای N=۳۰۰۰

نمودارهای فوق تغییر مقادیر ضریب خوشه‌بندی را بر اساس حذف گره‌های پر اهمیت با توجه به چهار استراتژی تعریف شده نمایش می‌دهند. نتایج نشان می‌دهد با حذف گره‌های دارای بیشترین مقدار C_c، ALCC افت ناگهانی را تجربه می‌کند و نمودار سریع‌تر به مقدار کمتر متمایل می‌شود. این افت نسبت به دیگر استراتژی‌ها به‌ویژه برای ابعاد بزرگ‌تر داده، بیشتر قابل احساس است. مثلاً در شکل ۶، پس از حذف ۰/۰۱۲ از نودها، نمودار CC مقدار ۰/۱۴۹ را برای ALCC نشان می‌دهد که نسبت به سایر استراتژی‌ها کمتر است. این اختلاف کاهشی با حذف درصد بیشتری از نودها حفظ می‌شود. با توجه به اینکه ALCC، میانگین ضریب را محاسبه می‌کند حذف نودهای پر اهمیت در شبکه‌های کوچک‌تر، تأثیر بیشتری بر این مقدار خواهد داشت. این تفاوت در کاهش برای داده‌های با ابعاد کوچک‌تر ۵۰۰ و ۱۰۰۰ تا ۰/۰۱ برای سایر داده‌ها با حجم بزرگ‌تر بسیار کمتر است (با توجه به مقدار اولیه ضریب خوشه‌بندی). به عبارت دیگر حذف گره‌ها بر اساس استراتژی CC_greedy بیشترین کاهش را در ضریب خوشه‌بندی و هم‌زمان با کمترین حذف تعداد نودها ایجاد می‌کند و آسیب‌پذیری شبکه را بیشتر تحت تأثیر قرار می‌دهد.

در ادامه برای تکمیل ارزیابی از دو مجموعه داده واقعی یوتیوب و EU email به‌عنوان دو مجموعه داده پرکاربرد عملکردی در شبکه‌های اجتماعی استفاده خواهد شد. یوتیوب یک وب‌سایت اشتراک‌گذاری ویدئو است که حاوی یک شبکه اجتماعی است. در یوتیوب کاربران با یکدیگر رابطه دوستی ایجاد می‌کنند و می‌توانند گروههایی را ایجاد کنند که کاربران دیگر نیز می‌توانند به آن‌ها بپیوندند. این داده‌ها توسط آلن و همکاران [۲۳] ارائه شده است. این مجموعه داده حاوی ۱۱۳۴۹۰ گره

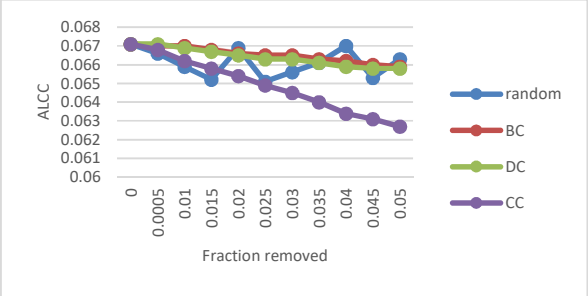
و ۲۹۸۷۶۲۴ یال است. نتایج ارزیابی میانگین ضریب خوشه‌بندی برای این مجموعه داده با توجه به چهار استراتژی تعریف شده، در شکل ۷ نمایش داده شده است.



شکل ۷. مقادیر ضریب خوشه‌بندی برای یوتیوب

مجموعه داده EU email، یک شبکه تولیدشده با داده‌های ایمیل از یک موسسه تحقیقاتی بزرگ اروپایی است. این داده‌ها در بازه اکتبر ۲۰۰۳ تا می ۲۰۱۵ (۱۸ ماه) توسط لکسوی و همکاران[۲۴]جمع‌آوری شده‌اند و اطلاعات زیادی در مورد تمام نامه‌های دریافتی و خروجی این موسسه تحقیقاتی وجود دارد.

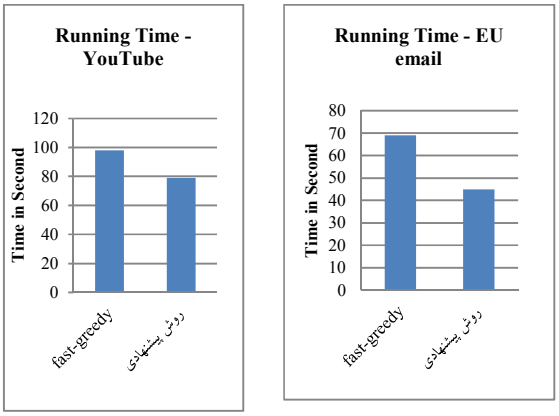
درکل، ۳۰۳۸۵۳۱ ایمیل بین ۲۸۷۷۵۵ کاربر مختلف وجود دارد. قابل‌ذکر است که یک گراف کامل ایمیل برای ۱۲۵۸ آدرس ایمیل وجود دارد که از مؤسسات تحقیقاتی می‌آیند. علاوه بر این ۳۴۲۰۳ آدرس ایمیل وجود دارند که در محدوده مجموعه داده ایمیل ارسال و دریافت می‌کنند. تمام آدرس‌های ایمیل دیگر mistyped یا اسپم هستند. بر اساس یک مجموعه از پیام‌های ایمیل، هر گره متناظر با یک آدرس ایمیل است. اگر i حداقل یک ایمیل به j ارسال کند یک یال جهت‌دار بین گره‌های i و j ایجاد می‌شود. تعداد گره‌های این مجموعه ۲۶۵۲۱۴و تعداد یال‌ها ۴۲۰۰۴۵ است. نتایج ارزیابی میانگین ضریب خوشه‌بندی برای این مجموعه داده در شکل ۸ نمایش داده‌شده است.



شکل۸ . مقادیر ضریب خوشه‌بندی برای EU email

نمودارها بیانگر رفتار ضریب خوشه‌بندی بر اساس حذف گره‌های پراهمیت با توجه به استراتژی‌های گفته شده است. نتایج نشان می‌دهد ضریب خوشه‌بندی نسبت به حذف گره‌ها به‌صورت تصادفی حساسیت زیادی نشان نمی‌دهد. اما از بین سایر پارامترها، حذف گره‌های شبکه بر اساس استراتژی CC_greedy بیشترین کاهش را در ضریب خوشه‌بندی همزمان با کمترین حذف تعداد نودها ایجاد می‌کند و آسیب‌پذیری شبکه را بیشتر تحت تأثیر قرار می‌دهد. به عبارت دیگر حذف آن‌ها، آسیب بیشتری به شبکه ازلحاظ کاهش ضریب خوشه‌بندی وارد می‌کند. بنابراین، پارامترهای CB و CD برخلاف محبوبیت آن‌ها در ارزیابی رفتار شبکه‌های اجتماعی، برای ارزیابی آسیب‌پذیری مناسب نمی‌باشند. به‌عبارت‌دیگر این دو معیار برخلاف CC، هنگامی کاهش قابل‌ملاحظه‌ای در ضریب خوشه‌بندی ایجاد می‌کنند که تعداد قابل‌توجهی از نودهای شبکه حذف شوند. حذف درصد زیادی از گره‌ها بمنظور بررسی رفتار ضریب خوشه‌بندی در تحلیل آسیب‌پذیری مناسب نیست. نکته قابل‌توجه دیگر پیچیدگی محاسباتی و زمان مورد نیاز برای به دست آوردن مقادیر جدید ضریب خوشه‌بندی است که برای CC_greedy زمان بیشتری را نیاز دارد این امر می‌تواند به خاطر روند محاسبه این معیار مرکزیت توجیه شود.

زمان اجرای روش پیشنهادی با یکی از روش‌های جدید تحلیل آسیب‌پذیری، الگوریتم fast_greedy [۹] برای دو مجموعه داده EU email و یوتیوب در شکل ۹ مقایسه شده است. در این نمودار زمان اجرای الگوریتم‌های fast_greedy و روش پیشنهادی با حذف گره‌های دارای بیشترین مقدار CC، به‌عنوان مهم‌ترین معیار آسیب‌پذیری حاصل از این پژوهش، نمایش داده شده است. همان‌طور که مشهود است روش پیشنهادی این مقاله، زمان اجرا برای مجموعه داده‌های EU email و یوتیوب را به ترتیب با میزان کاهش زمان ۲۴ و ۱۹ ثانیه‌ای به ۴۵ و ۷۹ ثانیه رسانده است.



شکل ۹. زمان اجرا برای دو الگوریتم fast_greedy [۹] و روش پیشنهادی

همان‌طور که گفته شد، هدف این پژوهش، تحلیل آسیب‌پذیری شبکه از طریق بررسی رفتار کاهشی مقادیر ضریب خوشه‌بندی و پس از حذف نودهای باارزش بیشتر است. در هر مرحله و با حذف بخشی از نودهای پراهمیت، نیاز به محاسبه مجدد ALCC است. بنابراین پیچیدگی روش‌های محاسبه ALCC در شبکه‌ها، نشان‌دهنده پیچیدگی محاسباتی رویکرد پیشنهادی است. با توجه به

اینکه روش‌های محاسبه ALCC هرکدام دارای مزایا و معایبی هستند در این کار از روش پیشنهادی ارائه شده در مرجع [۲۵] استفاده‌شده است که دارای پیچیدگی زمانی $O(N^w)$ که w کوچک‌تر یا مساوی ۲/۲۳۷ است.

۶- نتیجه‌گیری و راهکارهای آتی

آسیب‌پذیری خوشه‌بندی در ارزیابی توانمندی شبکه‌های اجتماعی نقش مهمی را ایفا می‌کند، زیرا سطح خوشه‌بندی برای برنامه‌های مختلف حائز اهمیت است، ازجمله نقش برجسته‌ای در انتشار اطلاعات در یک شبکه اجتماعی دارد. در این پژوهش کشف گره‌های بحرانی را بر اساس ضریب خوشه‌بندی بررسی کرده و نتایج قابل‌توجهی را به دست آوردیم. نتایج تجربی در مقایسه با استراتژی‌های مختلف در شبکه‌های ترکیبی و واقعی نشان می‌دهد که میانگین ضریب خوشه‌بندی در تحلیل نقص گره‌های تصادفی قدرتمند است. همچنین نتایج تائید می‌کند که بررسی آسیب‌پذیری بر اساس حذف نودهای پراهمیت به‌ویژه ازنظر مقدار Cc در تجزیه و تحلیل آسیب‌پذیری خوشه‌بندی مؤثر است و در شبکه‌های بزرگ‌تر هم مقیاس‌پذیر است. در این کار، تأثیر حذف گره‌های یک شبکه بر اساس پارامترهای ارزشمندی مرکزیت C_D ، C_B و C_C بررسی شد و نتایج نشان داد حذف گره‌های دارای بیشترین مقدار C_C آسیب‌پذیری شبکه را بیشتر تحت تأثیر قرار می‌دهد و حذف آن‌ها آسیب بیشتری به شبکه ازلحاظ کاهش ضریب خوشه‌بندی وارد می‌کند.

بااین‌حال، برای بهبود کار پژوهشی فعلی، هنوز برخی از مشکلات بالقوه باید حل شود. به عنوان مثال، برای مرتب کردن گره‌ها در شبکه‌های با ابعاد بزرگ، پیچیدگی زمانی و فضای مورد نیاز قابل توجه است. همچنین ویژگی خاص شبکه در بررسی مسئله آسیب‌پذیری باعث می‌شود اتفاق‌هایی غیر منتظره، ناگزیر رخ دهند که نیاز به بررسی بیشتر دارند. نزدیکی نسبی به دست آمده از دو گره ممکن است با استفاده از تحلیل پیشنهادی یکسان باشد، پس چگونه می‌توان رتبه‌بندی آن‌ها را تعیین کرد؟ علاوه بر این، برای یک شبکه خاص، گاهی اوقات انتخاب معیارهای مرکزیت مناسب دشوار خواهد بود. بنابراین، در مطالعه آینده، چارچوب روش ارائه‌شده می‌تواند برای شناسایی گره‌های بحرانی بر اساس روشی مانند آنتروپی نسبی بهبود یابد. همچنین برای نشان دادن اثربخشی روش پیشنهادی، کاربردهای بیشتری در سیستم‌های پیچیده ازجمله مهار بیماری‌ها، تشخیص ساختار جوامع ویروسی می‌تواند بررسی شود.

به‌عنوان یک پژوهش آتی دیگر می‌توان به استفاده از گراف‌های علامت‌دار به‌جای گراف‌های ساده برای تحلیل آسیب‌پذیری بر اساس پارامترهای پیشنهادی این پژوهش پرداخت.

منابع

[۱] توکلی، سمیه. و فاطمی، افسانه. «تشکیل تیم دوهدفه در شبکه‌های اجتماعی». مجله مهندسی برق دانشگاه تبریز،(۲) ۴۷: ۴۲۳-۴۳۳، ۱۳۹۷.

[2] Fei, L., H. Mo, and Y. Deng, *A new method to identify influential nodes based on combining of existing centrality measures.* Modern Physics Letters B, 2017. **31**(26): p. 1750243.

[3] Lü, L., Chen, D., Ren, X. L., Zhang, Q. M., Zhang, Y. C., & Zhou, T., *Vital nodes identification in complex networks*. Physics Reports, 2016. **650**: p. 1-63.

[4] Chen, X., *System vulnerability assessment and critical nodes identification.* Expert Systems with Applications, 2016. **65**: p. 212-220.

[5] Watts, D.J. and S.H. Strogatz, *Collective dynamics of 'small-world'networks*. nature, 1998. **393**(6684): p. 440.

[6] Aggarwal, C.C., *Social Network Data Analytics*. 2011: Springer Publishing Company, Incorporated. 516.

[7] Peng, S., Zhou, Y., Cao, L., Yu, S., Niu, J., & Jia, W., *Influence analysis in social networks: A survey.* Journal of Network and Computer Applications, 2018.

[8] *Social media active users by network*. https://www.statista.com, 2018.

[9] Kuhnle, A., Nguyen, N. P., Dinh, T. N., & Thai, M. T., *Vulnerability of clustering under node failure in complex networks.* Social Network Analysis and Mining, 2017. **7**(1): p. 8.

[10] Dinh, T. N., Xuan, Y., Thai, M. T., Pardalos, P. M., & Znati, T., *On new approaches of assessing network vulnerability: hardness and approximation.* IEEE/ACM Transactions on Networking, 2012. **20**(2): p. 609-619.

[11] Crucitti, P., Latora, V., Marchiori, M., & Rapisarda, A., *Error and attack tolerance of complex networks.* Physica A: Statistical mechanics and its applications, 2004. **340**(1-3): p. 388-394.

[12] Peixoto, T.P. and S. Bornholdt, *Evolution of robust network topologies: Emergence of central backbones.* Physical review letters, 2012. **109**(11): p. 118703.

[13] Callaway, D. S., Newman, M. E., Strogatz, S. H., & Watts, D. J., *Network robustness and fragility: Percolation on random graphs.* Physical review letters, 2000. **85**(25): p. 5468.

[14] Holme, P., Kim, B. J., Yoon, C. N., & Han, S. K., *Attack vulnerability of complex networks*. Physical review E, 2002. **65**(5): p. 056109.

[15] Grubesic, T. H., Matisziw, T. C., Murray, A. T., & Snediker, D., *Comparative approaches for assessing network vulnerability.* International Regional Science Review, 2008. **31**(1): p. 88-112.

[16] Veremyev, A., O.A. Prokopyev, and E.L. Pasiliao, *Critical nodes for distance-based connectivity and related problems in graphs.* Networks, 2015. **66**(3): p. 170-195.

[17] Gomes, T., Tapolcai, J., Esposito, C., Hutchison, D., Kuipers, F., Rak, J., De Sousa, A., Iossifides, A., Travanca, R., André, J. and Jorge, L., *A survey of strategies for communication networks to protect against large-scale natural disasters*. in *Resilient Networks Design and Modeling (RNDM), 2016 8th International Workshop*. 2016. IEEE.

[18] Nguyen, N. P., Alim, M. A., Shen, Y., & Thai, M. T., *Assessing network vulnerability in a community structure point of view*. in *Advances in Social Networks Analysis and Mining (ASONAM), 2013 IEEE/ACM International Conference* .2013. IEEE.

[19] Alim, M. A., Nguyen, N. P., Dinh, T. N., & Thai, M. T.. *Structural vulnerability analysis of overlapping communities in complex networks*. in *Proceedings of the 2014 IEEE/WIC/ACM International Joint Conferences on Web Intelligence (WI) and*

Intelligent Agent Technologies (IAT)-Volume 01. 2014. IEEE Computer Society.

[20] Ertem, Z., A. Veremyev, and S. Butenko, Detecting large cohesive subgroups with high clustering coefficients in social networks. Social Networks, 2016 :۴۶ p. 1-10.

[21] Freeman, L.C., *Centrality in social networks conceptual clarification.* Social networks, 1978. **1**(3): p. 215-239.

[22] Liu, J., Xiong, Q., Shi, W., Shi, X., & Wang, K., Evaluating the importance of nodes in complex networks. Physica A: Statistical Mechanics and its Applications, 2016. 452: p. 209-219.

[23] https://snap.stanford.edu/data/com-Youtube.html.

[24] https://snap.stanford.edu/data/email-Eu-core.html.

[25] F. Le Gall, "Powers of tensors and fast matrix multiplication." pp. 296-303.

زیرنویس‌ها

[۱] Robustness

[۲] Communication Network

[۳] Clustering

[۴] Social Network (SN)

[۵] Mobile Social Network (MSN)

[۶] Online Social Network (OSN)

[۷] Instant Messaging

[۸] Robustness

[۹] Microblogging

[۱۰] Social Bookmarking

[۱۱] Average Clustering Coefficient

[۱۲] Degree Centrality

[۱۳] Betweenness Centrality

[۱۴] Closeness Centrality