

مقاله پژوهشی

مرکزیت در شبکه‌های اجتماعی مبتنی بر مدل آستانه خطی

Doi: 10.30508/KDIP.2022.361686.1049

علی سرآبادانی^۱ | خیرالله رهسپارفرد (نویسنده مسئول)^۲^۱- دانشجوی دکتری تخصصی مهندسی فناوری اطلاعات گرایش تجارت الکترونیک، دانشگاه قم، قم، ایران^۲- استادیار گروه مهندسی کامپیوتر و فناوری اطلاعات، دانشگاه قم، قم، ایران

تاریخ دریافت: ۱۴۰۱/۰۶/۲۰

تاریخ پذیرش: ۱۴۰۱/۱۰/۰۳

صفحه: ۴۲ - ۵۳

چکیده:

مرکزیت و انتشار نفوذ دو مؤلفه اساسی در تحلیل شبکه‌های اجتماعی هستند. در سال‌های اخیر، اندازه‌گیری مرکزیت توجه بسیاری از محققان را جلب کرده و همین امر باعث شده که بسیاری از مطالعات جدید در مورد تحلیل شبکه‌های اجتماعی و کاربردهای آن صورت بگیرد. مدل‌های متداول قدیمی از انتشار نفوذ به صورت گسترده برای اندازه‌گیری مرکزیت استفاده نمی‌کنند. بسیاری از مطالعات صورت گرفته بر اساس مدل آبشاری مستقل شکل گرفته‌اند. در این مقاله به بررسی ارزیابی و اندازه‌گیری مرکزیت بر اساس مدل آستانه خطی در شبکه‌های اجتماعی برچسب دار و وزن دار پرداخته شده است. یک روش برای اندازه‌گیری مرکزیت به منظور رتبه‌بندی اعضای شبکه در نظر گرفته شده است که رتبه آستانه خطی نامگذاری شده است. برای ارزیابی روش پیشنهادی چهار شبکه اجتماعی را با دو معیار اندازه‌گیری مرکزیت و یک روش آبشاری مستقل مقایسه شده‌اند. نتیجه نشان می‌دهد که روش پیشنهادی برای رتبه‌بندی اعضای شبکه و شبکه‌ها مناسب‌تر است. یک سنجه مرکزیت جدید مبتنی بر مدل آستانه خطی که برای هر عضو، به عنوان تعداد اعضای در نظر گرفته می‌شد که می‌توانست انتشار نفوذ صورت دهد، ارائه شد. این سنجه را با سه سنجه مرکزیت دیگر (مرکزیت کاتز، رتبه صفحه و رتبه‌بندی آبشاری مستقل) بر اساس نفوذ و تأثیر در چهار شبکه واقعی دو شبکه بزرگ (یکی جهت‌دار، دیگری بی‌جهت) و دو شبکه کوچک (یکی جهت‌دار، دیگری بی‌جهت) مقایسه شده است.

کلمات کلیدی: مرکزیت، مدل آستانه خطی، مدل آبشاری مستقل، گسترش نفوذ، شبکه اجتماعی

۱- مقدمه

مرکزیت یکی از مؤلفه‌های اساسی در تحلیل شبکه‌های اجتماعی است. بررسی این مؤلفه از سال ۱۹۴۸ آغاز شده است (باولاس^۱، ۱۹۴۸). یک شبکه اجتماعی را می‌توان مانند یک گراف نشان داد که گره‌ها نماینده اعضای شبکه هستند و یال‌ها نماینده روابط بین اعضا (سان و تانگ^۲، ۲۰۱۱؛ سدیمان^۳، ۱۹۸۳). گاهی یال‌ها وزن دار هستند که نشان‌دهنده قدرت بین هر کدام از روابط اعضا خواهند بود. در این میان سنجه‌های مرکزیت نشان‌دهنده ارتباط ساختاری هر عضو با این شبکه اجتماعی هستند. یکی از قدیمی‌ترین سنجه‌های مرکزیت، می‌توان به مرکزیت درجه^۴، مرکزیت بینایی^۵، مرکزیت نزدیکی^۶ اشاره کرد که به توپولوژی گراف نیز مرتبط بودند. در این سنجه‌ها، یک عضو بیشتر مرکزی در نظر گرفته می‌شود هنگامی که از درجه بالاتری برخوردار باشد، یا به دیگر اعضا نزدیک‌تر باشد، یا در بین کوتاه‌ترین مسیر بین اعضا قرار گرفته باشد (بالابس^۷، ۱۹۸۴؛ فریمن^۸، ۲۰۰۲). در سال‌های اخیر به دلیل افزایش حیرت‌برانگیز کاربران اینترنت و ایجاد شبکه‌های گوناگون و پیچیده اجتماعی، ظهور سنجه‌های مرکزیت پیچیده‌تر را به وضوح می‌توان مشاهده کرد (بادر و هوگو^۹، ۲۰۰۳؛ لوزارس، لپزولدان، بولیبار، و مانتانیول^{۱۰}، ۲۰۱۵). بنا بر حجم بسیار زیاد این شبکه‌ها، یعنی تعداد بسیار زیاد اعضا

و روابط بین آن‌ها، لازم می‌باشد که سنجه‌ها به صورت عملی قابل محاسبه باشند. امروزه، سنجه‌هایی وجود دارند که بر اساس این که چه مقدار اطلاعات از گره در شبکه پخش می‌شود یا به کار گرفته می‌شوند (الطاف امین و همکاران^{۱۱}، ۲۰۰۳؛ کیتساک و همکاران^{۱۲}، ۲۰۱۰؛ گومز، فیگوریا، و یوسیبو^{۱۳}، ۲۰۱۳). همچنین شاخص‌هایی برای شبکه‌های اجتماعی به خصوصی ایجاد شدند مانند شبکه تویتر که بیش از هفتاد سنجه مرکزیت فقط از سال ۲۰۱۰ برای آن تعیین شده‌اند (ریکولم، گونزال، مولینرو و ساما^{۱۴}، ۲۰۱۸). دو تا از شناخته شده‌ترین سنجه‌ها، سنجه رتبه صفحه (پیج، برین، موتاوانی، و وینگورال^{۱۵}، ۱۹۹۹)، و مرکزیت کاتز (کاتز^{۱۶}، ۱۹۵۳) هستند. هر دوی این سنجه‌ها، مشتق شده از مرکزیت بردار ویژه^{۱۷} هستند (بوناریچ^{۱۸}، ۱۹۷۲). از آنجایی که ارتباط بین کاربران برای بسیار از کاربری‌ها و سامانه‌ها کاربرد دارد مانند بازاریابی و پیروسی (دامیونوگ و ریچاردسون^{۱۹}، ۲۰۰۱)، انتشار اطلاعات (گاتلر، و پاترینگانی^{۲۰}، ۲۰۰۴؛ گراهل، گاه، لیل نیل، و تامکینز^{۲۱}، ۲۰۰۴)، استراتژی‌های جست‌وجو (آدامیک، و آدر^{۲۲}، ۲۰۰۵)، پیشنهاددهی حرفه‌ای و تخصصی (سانگ، تسنگ، لین، و سان^{۲۳}، ۲۰۰۶)، دستگاه‌های مجتمع (ژانگ، ژوو، وانگ، و ژائو^{۲۴}، ۲۰۱۳)، مدیریت ارتباط مشتری اجتماعی تئوری نفوذپذیری (مورون و میکس^{۲۵}، ۲۰۱۵)، فراتر از این موارد سنجه‌های مرکزیت برای شناسایی فعال‌ترین،

- 1- Bavelas
- 2- Sun, & Tang
- 3- Seidman
- 4- Centrality
- 5- Betweenness
- 6- Closeness centrality
- 7- Bollobás
- 8- Freeman
- 9- Bader, & Hogue
- 10- Lozares, López-Roldán, Bolibar, & Muntanyola
- 11- Altaf-Ul-Amine, etal
- 12- Gómez, Figueira, & Eusébio
- 13- Riquelme, Gonzalez-Cantergiani, Molinero, & Serna
- 14- Page, Brin, Motwani, & Winograd
- 15- Katz
- 16- Eigen Vector
- 17- Bonacich
- 18- Domingos, & Richardson
- 19- Gaertler, & Patrignani
- 20- Gruhl, Guha, Liben-Nowell, & Tomkins
- 21- Adamic, L., & Adar, E. (2005)
- 22- Song, Tseng, Lin, & Sun
- 23- Zhang, Zhu, Wang, & Zhao
- 24- Morone & Makse

و سنجه‌های مرکزی سازی ارائه شده را برای ۴ شبکه اجتماعی در نظر گرفته شده است. دوتا از آن‌ها شبکه‌های بزرگی محسوب می‌شوند: شبکه هیگز (جهت دار) و شبکه آرکایو (غیر جهت دار). دو شبکه دیگر شبکه‌های کوچک ولی مشهوری هستند. شبکه میز ناهارخوری (جهت دار)، شبکه اجتماعی دلفین‌ها (غیر جهت دار). هر سنجه مرکزیت یک شاخص مرکزیت جداگانه‌ای را ارائه می‌دهد مگر برای شبکه اجتماعی دلفین‌ها، جایی که روش مقام آستانه خطی جدید^{۱۳}، رتبه صفحه و مرکزیت کاتز شبیه به هم عمل می‌کنند.

۲- مبانی نظری

در این بخش به معرفی سنجه‌های مرکزیت شناخته شده پرداخته شده است. همچنین بررسی شده است که هر کدام به چه صورت کار می‌کنند و چگونه سنجه‌های مرکزیت همبستگی دارند. در نهایت سنجه‌های مرکزی سازی را بیان شده است. شبکه اجتماعی مانند یک گراف می‌باشد که در آن V نماینده گره‌ها و E نماینده یال‌های ارتباطی خواهند بود. بعضی اوقات نیاز به یک گراف وزن دار است که در آن G گراف مورد نظر و W تابع وزنی است که به هر گره تخصیص داده می‌شود.

معیارهای مرتبط با مرکزیت: یکی از پرکاربردترین سنجه‌ها در بین شاخص‌های مربوط، مرکزیت بردار ویژه هست (بانریچ و همکاران، ۱۹۷۲)، که مشخص می‌کند یک عضو در یک گروه مهم می‌باشد اگر همسایه‌های مهمی به آن لینک شده باشند یا اعضای خیلی زیادی با آن همسایه باشد. به بیان دیگر، یک ماتریس مجاورت A را در نظر بگیرید، عنصر a_{ij} از ماتریس A ، مقدار یک را دارد اگر عضو i گره j به عضو متصل شده باشد در غیر این صورت مقدار

محبوب‌ترین و تأثیرگذارترین کاربر در یک شبکه نیز استفاده می‌شوند (لی و همکاران، ۲۰۱۴). مدل انتشار تأثیر مدلی است که در آن اعضا بر روی یکدیگر توسط ارتباط بین خود با سایر اعضا در شبکه تأثیر می‌گذارند. گره‌ها تأثیر خود را در گراف اعمال می‌کنند. به این صورت که اگر مجموعه‌ای از اعضا به یک‌روند مطابق شوند، ممکن خواهد بود بر روی دیگر اعضا تأثیر بگذارند که آنان نیز به این روند تطبیق پیدا کنند. مطمئناً این یکی واضح‌ترین و شناخته‌ترین پدیده در تحلیل شبکه‌های اجتماعی است (یاسلی، و کلینبرگ، ۲۰۱۰). یکی از شناخته‌ترین مدل‌های عمومی برای انتشار نفوذ، مدل آستانه خطی است (وانگ، ما، وانگ، ایکسی، کوه، و چوانگ، ۲۰۲۱؛ کمپ، کلینبرگ، و تاردوس، ۲۰۰۵). مدل آستانه خطی بر اساس یک سری ایده از رفتار جمعی الگو گرفته (گرانووتر، ۱۹۷۸؛ اسچلینگ، ۲۰۰۶)، مدل آبشاری مستقل برای موضوع بازاریابی مطرح شده بود (گلدنبرگ، لیبای و مولر، ۲۰۰۱)، بسیاری از تلاش‌های تحقیقاتی به مطالعه مسئله پیشینه‌سازی تأثیر بر اساس مدل آستانه خطی و دیگر مدل‌ها اختصاص داده شده است (کندال، ۱۹۸۳؛ لیو و همکاران، ۲۰۱۶؛ فاستر، موت، پوترات، و روتنبرگ، ۲۰۰۱). در این مسئله تلاش برای پیدا کردن تعداد k عضو کلیدی برای پیشینه‌سازی انتشار تأثیر در میان تمام مجموعه‌ها با سایز یکسان است. بنابراین مسئله پیشینه‌سازی تأثیر بر اساس مدل آستانه خطی یک مسئله NP-hard خواهد بود (سارامکی، کیولا، اونلا، کاسکی، و کرتز، ۲۰۰۷؛ کورو متیوز، ۲۰۲۰؛ لانگیول، مییر، ۲۰۰۴) مطالعات در مورد سنجه‌های مرکزیت مشتق شده ناقص بوده و فقط مدل آبشاری مستقل را در نظر گرفته بودند این رتبه باندی‌ها در نظر گرفته و محاسبه شدند تا یک سری راه حل برای مسئله پیشینه‌سازی تأثیر ارائه شود. مرکزیت

- 1- Li et al
- 2- Easley, & Kleinberg
- 3- Wang, Ma, Wang, Xie, Koh, & Cheong
- 4- Kempe, Kleinberg, & Tardos
- 5- Granovetter
- 6- Schelling
- 7- Goldenberg, Libai, & Muller
- 8- Kendall
- 9- Foster, Muth, Potterat, & Rothenberg,
- 10- Saramäki, Kivelä, Onnela, Kaski, & Kertesz
- 11- Corr, & Matthews
- 12- Langville, & Meyer
- 13- LTR

مرکزیت مرتبط با نفوذ: در کنار تمام این سنجه‌هایی که بر اساس مرکزیت بردار ویژه بنا شده‌اند، یک سری سنجه‌های مرکزیت وجود دارند که بر اساس نفوذ تأثیر مبتنی بر مدل آبخاری مستقل بنا شده‌اند؛ بنابراین مدل، هر عضوی احتمال دارد که بر عضوی که هدف قرار گرفته تأثیرگذار باشد. برای انتشار این تأثیر، عضو باید فعال باشد و فقط یک شانس برای تأثیر بر عضوی دیگر دارد. وقتی یک عضو سعی در تأثیر بر عضوی دیگر را دارد این عضو فعال در نظر گرفته می‌شود و این روال برای این عضو تکرار می‌شود. تمام این فرآیند زمانی خاتمه می‌یابد که نه عضو فعالی وجود داشته باشد نه شانس برای تأثیرگذاری. به عنوان مثال، سنجه زیر را در نظر بگیرید (کمپ و همکاران، ۲۰۰۵)، که به آن رتبه‌بندی آبخاری^۳ مستقل گفته می‌شود.

$$ICR(u, p) = \frac{|F'(u, p)|}{\max \{|F'(v, p)| \mid v \in V(G)\}}$$

که در آن V مجموعه تمام اعضا در شبکه هست و $F'(u, p)$ فرآیند تأثیر نفوذ مبنی بر مدل آبخاری مستقل خواهد بود؛ که با عضوی با نشان u آغاز خواهد شد. این سنجه یک احتمال ثابت p برای تأثیر هر عضو در نظر می‌گیرد. بنابراین هر عضو به احتمال $p-1$ به صورت غیرفعال برای تأثیر از همسایه خود باقی مهم‌اند؛ و به احتمال $(p-1)$ به توان r از تعداد r همسایه که به آن متصل هستند غیرفعال باقی می‌ماند. در همه سنجه‌های مبتنی بر مدل آبخاری مستقل، مانند رتبه‌بندی آبخاری مستقل، نودهای فعال به احتمال انتشار بستگی دارند بنابراین یک سنجه می‌تواند رتبه‌بندی‌های متفاوتی برای هر اجرا داشته باشد. **مقایسه سنجه‌های مرکزیت:** به منظور مقایسه نتایج حاصل از دو سنجه مرکزیت، به طور معمول از همبستگی ایستا استفاده می‌شود. پرکاربردترین ضرایب برای ایجاد همبستگی در سنجه‌های مرکزیت، ضریب همبستگی رتبه‌ای اسپیرمن ρ و ضریب همبستگی رتبه‌ای کندال T

صفر را خواهد داشت. مرکزیت بردار ویژه از عضو u که با $EV(u)$ نشان داده می‌شود عبارت است از:

$$EV(u) = \frac{1}{\lambda} \sum_{v \in V(G)} (\alpha_{uv}) EV(v)$$

که در آن ثابت λ اندا، مقدار ویژه خوانده می‌شود. مرکزیت بردار ویژه زمانی نتایج قابل قبولی ارائه می‌دهد که گراف به شدت متصل باشد، مانند شبکه غیر جهت‌دار با کامپوننت^۲ های بسیار متصل. در شبکه‌های جهت‌دار واقعی، نودهایی به دست آورده می‌شود که مرکزیت بردار ویژه آن‌ها برابر صفر خواهد بود، بنابراین این سنجه ناکارآمد می‌شود. برای مثال، نودهایی که به کامپوننت به شدت متصل، نزدیک هستند ولی راهی برای اتصال به آن‌ها نیست. اگرچه، مرکزیت کاتز (کاتز، ۱۹۵۳) (براین کمبود بردار ویژه غلبه می‌کند ولی با استفاده از تخصیص دادن یک مقدار کوچک از مرکزیت برای نقش اعضای آزاد و بی‌اهمیت در شبکه مرکزیت کاتز یک عضو u که با $Katz(u)$ نمایش داده می‌شود عبارت است از:

$$KATZ(u) = \alpha \sum_{v \in V(G)} (\alpha_{vu}) KATZ(v) + \beta$$

که β یک ثابت که مستقل از شبکه است و α که فاکتور

محرك نامیده می‌شود، یک عدد بین صفر و $\frac{1}{\lambda \max}$ که اندای ماکس بزرگ‌ترین مقدار ویژه از A خواهد بود. لازم به ذکر است که وقتی α بیشترین مقدار و β برابر صفر باشد مقدار مرکزیت کاتز برابر مرکزیت بردار ویژه خواهد شد. علاوه بر این، یک سنجه مرکزیت شناخته شده دیگر نیز وجود دارد به نام رتبه صفحه (پیچ، ۱۹۹۹) که با $pr(u)$ نشان داده می‌شود.

$$PR(u) = (1 - \alpha) + \alpha \sum_{v \in V(G)} \frac{(\alpha_{vu}) PR(v)}{\delta + (v)}$$

1- Eigen Value
2- Component
3- ICR

$$C_i = \frac{2T(i)}{\delta(i)(\delta(i)-1)}$$

$\delta(i)$ درجه گره i هست. برای گراف های وزن دار، این مقادیر متفاوت خواهند بود. ولی کتابخانه NetworkX از رابطه زیر استفاده می کند.

$$C_i = \frac{1}{\delta(i)(\delta(i)-1)} \sum_{\substack{j,k \in V(G) \\ j \neq k}} (\hat{W}_{ij} \hat{W}_{ik} \hat{W}_{jk})^{1/3}$$

مدل آستانه خطی برای انتشار تأثیر

یک شبکه ی اجتماعی در قالب یک گراف قابل نمایش است که در آن گره ها اعضای شبکه و یال ها روابط فردی بین اعضا را نمایش می دهد. گاهی اوقات یال ها دارای وزن نیز هستند که این وزن بیان گر قوت هر رابطه است. یک ویژگی کمتر رایج، برچسب گره ها است که برای نمایش ویژگی های مختلف شبکه استفاده می شود. در این کار، برچسب ها بیان گر میزان مقاومت گره در تأثیرپذیری است، در حالی که وزن یال ها بیان گر قدرت تأثیر اعمال شده از یک عضو به عضو دیگر هست. یک شبکه ی اجتماعی وزن دار و برچسب دار، به صورت رسمی تر در قالب یک گراف تأثیر (مولینرو و همکاران، ۲۰۱۵) قابل تعریف است. در ادامه، در تمام موارد مدل آستانه ی خطی را برای تأثیر انتشار در نظر می گیریم.

تعریف ۱: یک گراف تأثیر یک چندتایی (G, w, f) است که در آن یک $G = (V, E)$ گراف جهت دار تشکیل شده از مجموعه ی افراد V و مجموعه ی یال های جهت دار E است. $E; w: E \rightarrow \mathbb{N}$ یک تابع وزن دهی است که یک وزن به هر یال تخصیص می دهد. $f: V \rightarrow \mathbb{N}$ یک تابع برچسب زنی است که میزان سهولت تأثیرپذیری هر عضو را اندازه گیری می کند. یک عضو $i \in V$ تنها در صورتی روی عضو $j \in V$ تأثیر می گذارد که $(i, j) \in E$.

توجه داشته باشید که می توان یک گراف بدون جهت را به صورت یک گراف جهت دار متقارن در نظر گرفت، بنابراین این معیار برای این نوع گراف نیز قابل استفاده است. با توجه به گراف (G, w, f) و یک مجموعه ی فعال ساز $X \subseteq V$ ، فرآیند تکراری فعال سازی زیر را در نظر

هستند. اگر w و γ دو لیست از n عضو باشند ما داریم (به و و، ۲۰۱۳).

$$\rho = 1 - \frac{6 \sum_{i=1}^n (x_i - y_i)^2}{n(n^2 - 1)} \text{ and } \tau = \frac{n_c - n_d}{0.5n(n-1)}$$

که x_i و y_i رتبه های اعضای i در لیست x و y هستند. علاوه بر این n_c تعداد زوج های موافق است و n_d تعداد زوج های ناموافق. مقدار هر دو ρ و τ در بازه -1 تا 1 هستند، به طوری که یک معنی این است که دو سنج برابر هستند و صفر یعنی این که دو تا کاملاً مستقل هستند و -1 به این معنا که دو خلاف یکدیگر هستند. محاسبه ضرایب همبستگی یک همگونی با مقدار احتمال خواهد داشت. به طور معمول، ما ۵۰٪ را به عنوان خط برش مشخص در نظر می گیریم بنابراین فرضیه در صورتی که احتمال کمتر از ۵۰٪ باشد رد خواهد شد.

سنجه های مرکزی سازی: برای هر سنج مرکزیت، سنج مرکزی سازی فریمن نیاز به دانستن بیشترین مقدار ممکن جمع تمام مرکزیت ها را برای تمامی شبکه ها دارد. این مورد برای بسیاری از سنج های مرکزیت مانند مرکزیت درجه، بینابینی و نزدیکی در شبکه های غیر جهت دار آسان خواهد بود ولی در مورد دیگر سنج ها، این مورد سخت خواهد بود. علاوه بر این، ما می توانیم ضریب خوشه بندی متوسط را به عنوان سنج مرکزی سازی بیان کنیم (واتس و استروگاتز، ۱۹۹۸). این ضریب، متوسط تمامی ضرایب خوشه بندی محلی است ولی در شبکه های وزن دار تعریف دیگری دارد. این مورد با رابطه زیر معرفی می شود.

$$ACC = \frac{1}{n} \sum_{i=1}^n C_i$$

C_i ضریب خوشه بندی محلی برای عضو i خواهد بود. برای گراف های غیر وزن دار برای مثال وقتی که تمام یال های ارتباطی وزنی برابر یک داشته باشند C_i برابر خواهد بود با تعداد مثلث هایی که i در آن عضو هست در برابر تمام تعداد مثلث های ممکن

where $\text{neighbors}(i) = \{i \in V \mid (i, j) \in E \vee (j, i) \in E\}$

در تعريف ما G يك گراف جهت دار است اما همسايه هاي عامل i عواملی هستند كه به وسيله لبه در هر جهت به متصل شده اند. اين به ما اجازه می دهد كه فعال سازی اوليه را افزايش دهيم. مشاهده كنيد كه آن عوامل با زويه خروجی كوچك نمی توانند نفوذ خود را در شبكه انتشار دهند. بدین ترتيب معيار به دست آمده شبیه به درجهی مركزيت است. از اين رو، ما اجتماع $F_0(X) = \{i\}$ و همسايه هاي i را به عنوان فعال سازی اوليه در نظر می گيريم. هنگامی كه $F(X)$ می تواند در زمان مشخص پردازش محاسبه شود، $LTR(i)$ نیز قابل محاسبه در اين زمان است. علاوه بر اين، به عنوان اندازه گيري رتبه بندی آبخاري و به جای آن رتبه صفحه و مركزيت كاتز، مقام آستانه خطی جديد همیشه همگراست. اين نکته قابل توجه است كه ما هم چنين می توانيم معيارهاي ديگر مربوط به همسايه هاي عامل را در نظر بگيريم. هم چنين اين مجموعه ها می توانند با توجه به وزن ها به منظور تنظيم نقش عامل مركزی در گسترش نفوذ پالايه شوند. به عنوان مثال ما در فعال سازی اوليه می توانيم تنها آن دسته از همسايگان را كه برای w_{ij} مطابق معيار پايين قرار دارند را در نظر بگيريم. تحت تأثير دوم، نفوذ گسترش $F(\{i\})$ عامل i می تواند با اندازه گيري مقام آستانه خطی جديد همبستگي بيشتري داشته باشد. با اين حال اگر محدوديت هاي همسايگان خیلی زياد است، اندازه گيري ريسك ابتلا به شبیه درجهی مركزيت را به جريان می اندازد كه يك اقدام محلی است كه تنها بخش كوچكي از كل شبكه را در نظر می گيرد. ما چنين گزینه هایی را به عنوان كارهاي آينده ترك كرديم. اندازه گيري مقام آستانه خطی جديد می تواند به اين عنوان تفسير شود كه چه مقدار عامل i می تواند نفوذ خود را در يك شبكه گسترش دهد و چه مقدار سرمايه گذاري منابع خارجي سيستم رسمي بدون توجه به جهت هاي لبه قادر به متقاعد كردن همسايگان خود است. از نقطه نظر مثبت، اين سرمايه گزاري منابع می تواند ظرفيت عامل i را برای مدیریت مخاطبين خود در شبكه نشان دهد. از نقطه نظر منفی يا مشكوك، اين می تواند بيانگر توانايی او در رشوه باشد. به هر حال اين اندازه گيري وجود

بگيريد. فرض كنيد $F_t(X) \subseteq V$ مجموعه ای از گره ها باشد كه در تكرار t فعال می شود. در ابتدا در گام $t = 0$ ، تنها گرهی X فعال باشد، به اين معنا كه $F_0(X) = X$ در $t+1$ تكرار بعدي، گرهی $i \in V$ فعال می شود اگر و تنها اگر شرط زیر برقرار باشد:

$$\sum_{j \in F_t(X)} w_{ji} \geq f(i)$$

اين فرآيند زمانی كه هيچ فعال سازی ديگري رخ ندهد متوقف می شود. به عبارت ديگر، يك گرهی i زمانی فعال می شود كه مجموع وزن هاي گره هاي فعال متصل به آن بزرگ تر يا مساوی مقاومت آن در مقابل تأثير پذيری باشد. **تعريف ۲:** فرض كنيد (G, w, f) يك گراف تأثير باشد، انتشار تأثير X برابر است با:

$$F(X) = \bigcup_{t=0}^k F_t(X) = f_0(X) \cup \dots \cup F_k(X)$$

كه در آن، $k = \min \{t \in \mathbb{N} \mid F_t(X) = F_{t+1}(X)\} \leq n$ ، مقدار t در $F_t(X)$ سطح فعلی انتشار را نشان می دهد. توجه كنيد كه برای هر مجموعهی فعال ساز X ، انتشار تأثير $F(X)$ در زمان چند جمله ای قابل محاسبه است.

سنجه هاي جديد مركزيت و مركزی سازی

اقدامات جديد مركزيت و مركزی سازی در اين بخش اقدامات جديد نفوذ بر اساس گسترش نفوذ تحت مدل آستانه ی خطی بررسی می شود.

آستانه ی خطی مركزيت: مدل آستانه ی خطی به جای مدل آبخار مستقل به هيچ گونه شانسی بستگی ندارد بلکه به ظرفيت نفوذ هر عامل و مقاومت آن ها در برابر نفوذ وابسته است. ما يك اندازه گيري مركزی برای رتبه بندی کاربران شبكه به صورت زیر معرفی می كنيم.

تعريف ۳: اجازه دهيد (G, w, f) گراف نفوذ با $G = (V, E)$ و $|V| = n$ و a عضو V يك عامل باشد. رتبه بندی آستانه ی خطی a به صورت زیر مشخص می شود:

$$LTR(i) = \frac{|F(\{i\}) \cap \text{neighbors}(i)|}{n}$$

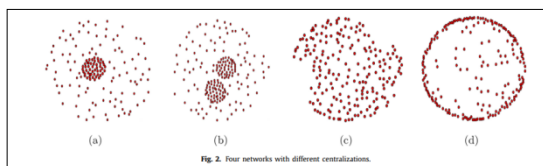
تعريف ۴: اجازه دهيد (G, w, f) گراف نفوذ با $G=(V, E)$ و $n=|V|$.
مرکزی سازی آستانه ی خطی به صورت زیر تعريف می شود:

$$LTC(G) = \frac{|F(\hat{C}(G))|}{n}$$

where $\hat{C}(G) = \{i \in V | i \text{ belongs to the main core of } G\}$

توجیه این اندازه گیری کاملاً طبیعی است. همانطور که بازیگران خارج از k -shell درجه ای کوچکتر از بازیگران داخلی دارند، اولین افراد بیشتر می توانند با دومین ها تحت تاثیر قرار بگیرند.

در مثالی دیگر، چنین اندازه گیری را به چهار شبکه اعمال شده است. شکل شماره (۱) نتایجی را که تقریباً تجسم تجزیه معمول k -shell را به تصویر می کشد، یعنی گره ها به گونه ای توزیع شده اند که یک گره دارای درجه بزرگتری است که نزدیک تر به مرکز است. توجه داشته باشید که گراف a در مرکز بیشتر مرکزی سازی شده است؛ بنابراین فعال سازی اولیه ی یک مجموعه ی بزرگ است که قادر است نفوذ خود را بر گره های دور منتشر کند. از این رو، اگر گره ها در مرکز توانایی نفوذ خود را بر گره های دور فراهم آورند، این می تواند نمونه خوبی از یک گراف با یک مرکز گرادین خطی بالا باشد. گراف (b) دو هسته بزرگ را به جای یکی نشان می دهد. اگرچه مرکزی سازی بالایی دارد، گره هایی از سطح هستند که فقط از یک هسته از دو هسته قابل دسترسی هستند؛ بنابراین اندازه گیری ما می تواند مقادیر پایین تری را نسبت به نمودار اول به دست آورد. نهایتاً به عنوان هسته اصلی هر دو گراف (c) و (d) پایین تر است، فعال سازی اولیه ممکن است برای گسترش نفوذ خود را به گره های محیطی کافی نباشد و در نتیجه مقادیر کمتری برای اندازه گیری مرکزی سازی به دست آید.



شکل (۱): نتایج تجسم تجزیه معمول k -shell

دو شبکه ی مختلف را نشان می دهد. یک شبکه ی رسمی روابط شناخته شده ی بین فردی و یک شبکه ی غیررسمی که عوامل می توانند در شبکه ی رسمی استفاده کنند. این واقعیت بسیار رایجی است که ما دائماً در معرض آن هستیم. به عنوان مثال، در شبکه تویتر، هزاران نفر از کاربران می توانند به روش های مختلفی برای انتشار یک رویداد تعامل داشته باشند. این اخبار بسته به ظرفیت نفوذ هر عامل مورد علاقه در اخبار، از طریق شبکه تحت پوشش قرار می گیرد. اگر چه کاربرانی که دارای بیشترین نفوذ هستند تأثیر بیشتری روی شبکه ایجاد می کنند، هر کاربر آزاد است اگر می خواهد در مورد رویدادی نظر دهد. در این شبکه ی رسمی کاربران نمی توانند به طور مستقیم با تصمیمات سایر کاربران مواجه شوند. با این وجود، سازمان دهندگان این رویداد می توانند از شبکه خارجی خود از مخاطبین (دوستی، رسانه ها و غیره) برای کمک به گسترش این رویداد استفاده کنند؛ بنابراین، مجموعه فعال سازی اولیه معمولاً توسط تنها یک بازیگر تشکیل نمی شود، بلکه توسط چندین مورد که به روشی دیگر که قابل مشاهده در شبکه ی رسمی نیست تشکیل می شود. توجه داشته باشید که افزایش اندازه مقام آستانه خطی جدید به اندازه گره های فعال سازی اولیه بستگی ندارد، همان طور که میزان تأثیر آن در عوامل می تواند باهم به عنوان یک ائتلاف عمل کند. در این اندازه گیری، یک عامل با سهم کمی در گسترش نفوذ ممکن است به دلیل همسایگانش رتبه خوبی داشته باشد. این در گسترش پدیده نفوذ و همچنین در بسیاری از اقدامات مرکزی است که رفتار کل شبکه را در نظر می گیرد. به عنوان مثال، اقدامات مبتنی بر محدوده ی خاصی از منابع مانن رتبه صفحه، جایی که یک عامل تنها می تواند به دلیل ارتباط مستقیم با بازیگران با نفوذ دیگر تأثیر بگذارد.

مرکزی سازی آستانه ی خطی: با استفاده از مدل آستانه ی خطی، ما یک اندازه گیری مرکزی جدید برای تعیین اینکه چه قدر کل شبکه مرکزی سازی شده است را تعریف می کنیم. هسته ی K یک گراف، بیشینه زیرگرافی است که در آن هر راس دارای کمینه درجه ی K است. K -shell زیرگرافی از گره های k -core است، اما نه در $(1+k)$ -core. هسته ی مرکزی، هسته با بیشترین درجه است.

۳- یافته‌های تحقیق

مثالی از یک گراف با آستانه تمرکز خطی بالا. گراف b دو هسته بزرگ را به جای یک هسته نشان می‌دهد. با این وجود، تمرکز زیادی در مرکز دارد، گره (عضو) هایی در سطح وجود دارند که می‌توانند تنها از طریق یک و یا دو هسته در دسترس باشند و بنابراین اندازه گیری ما مقدار کمتری را نسبت به گراف اول بازمی‌گرداند. در نهایت، هنگامی که هسته اصلی هر دو گراف c و d در سطح پایین تر هستند، انرژی فعالسازی اولیه احتمالاً برای تأثیر گذاشتن بر عضو های پیرامونی کافی نخواهند بود که این امر نتیجه اندازه گیری تمرکز در سطوح پایین تر انرژی است.

آزمایشات و نتایج: در این قسمت، نتایج معیار تأثیر مختلف را در چهار شبکه مختلف ارائه شده است:

- شبکه Higgs: شبکه ای بزرگ و جهت دار
- شبکه arXiv: شبکه ای بزرگ و فاقد جهت
- شبکه Dining-table: شبکه کوچک جهت دار
- شبکه Dolphins social network: شبکه کوچک فاقد جهت

ما چهار معیار برای مرکزیت در نظر می‌گیریم: رتبه صفحه، مرکزیت کاتر، رتبه بندی آشاری و مقام آستانه خطی جدید. همچنین LTC و ACC را به عنوان معیارهای متمرکز کردن در نظر گرفته شده است. دو مجموعه شبکه داده اول توسط Collection SNAP's Stanford Large Network Dataset ساخته شده‌اند و مورد سوم برای Pajek datasets در دسترس بوده و مورد آخر نیز برای Data Repository UCI Network. آزمایشات با زبان پایتون ۳ برنامه نویسی شده‌اند. برای کار با گراف‌ها، ما از NetworkX library استفاده شده است.

هر شبکه به عنوان یک گراف تأثیر نشان داده شده است (G, w, f) . برای هر واکنش دهنده $v \in V$ ، ما یک $f(i) = \lfloor \bar{w}/2 \rfloor + 1$ برچسب ایجاد شده است که، بنابراین یک واکنش دهنده تنها زمانی که به $\sum_{(j,i) \in E(G)} w_{ji}$ فعال سازی اولیه تعلق داشته باشد، فعال می‌شود و یا اعضا فعال بیش از نصف کل انرژی تأثیر وارده را به آن وارد می‌کنند.

شبکه Higgs

اولین مجموعه داده که همه ی توپیت های مرتبط با

آزمایش‌ها Higgs boson که بین اول تا هفتم جولای ری-توپیست شده بوده است را شامل می‌شود. این مجموعه نخست برای پروسه توسعه اطلاعات بر روی توپیتر استفاده می‌شده است. این مجموعه داده اجازه می‌دهد تا یک گراف تأثیر با $|V|=n=255491$ واکنش دهنده و $|E|=328132$ یال جهت دار تولید کنیم. یک یال جهت دار $E \ni (i, z)$ نشان دهنده یک واکنش دهنده i است که ری-توپیست می‌کند واکنش دهنده z را. بنابراین می‌توان گفت که i تغییر معینی را بر اعمال می‌کند. وزن یال w_{ij} نشان دهنده آن است که چند بار واکنش دهنده z واکنش دهنده i را ری-توپیست کرده است.

طبق تعریف، معیارهای رتبه بندی آشاری و مقام آستانه خطی جدید هر دو همگرا می‌شوند. برای محاسبه معیار رتبه بندی آشاری، ما از احتمال $p = 0.1$ استفاده می‌کنیم. برای محاسبه رتبه صفحه از عامل محرک $\alpha = 0.85$ که به صورت پیش فرض گذاشته شده استفاده می‌کنیم و الگوریتم همگرا شده قبل از ۱۰۰ تکرار نیز به صورت پیش فرض گذاشته می‌شوند. برای محاسبه مرکزیت کاتب نیز از مقادیر $\alpha = 0.1$ و $\beta = 0.1$ استفاده می‌کنیم که آن‌ها نیز به صورت پیش فرض انتخاب شده‌اند. هر چند که به جای رتبه صفحه، با مقادیر ترانس استاندارد و حتی یک میلیون تکرار، الگوریتم کاتب واگرا شده است؛ بنابراین برای آنکه همگرایی بتواند مقدار مناسب و عملی از منابع زمان و حافظه استفاده کند، ما تصمیم گرفتیم تا ترانس را از استاندارد قابل توجه ۶ رقم به تنها ۲ رقم کاهش دهیم. قبل از مقایسه نتایج معیارهای تأثیر مختلف، اجازه بدهید تا بر روی معیار مقام آستانه خطی جدید متمرکز شویم. جدول شماره (۱) و شکل شماره (۲) نشان می‌دهند که چه تعداد واکنش دهنده به سطوح گسترش برابر ۰، ۱ و ۲ و... رسیده‌اند. توجه کنید که تقریباً ۹۰ درصد واکنش دهنده‌ها به چهارمین سطح گسترش نمی‌رسند. هر چند که منحنی شکل گرفته توسط این دو متغیر کاهش شدید یکنواختی ندارد. در حقیقت، رایج ترین واکنش دهنده‌ها آن‌هایی هستند که دقیقه به دو سطح گسترش رسیده‌اند و هیچ واکنش دهنده‌ای وجود ندارد که دقیقه یازده سطح گسترش را طی کرده باشد. علاوه بر آن، بیشترین سطح گسترش واکنش دهنده‌ها که با معیار مقام آستانه خطی

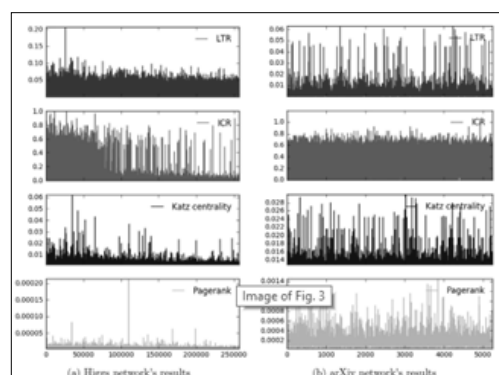
۴- نتيجه‌گيري

اين مقاله بر روي مقايسه سنجه‌هاي مركزيت كه مشتق شده از مدل آستانه خطي بود متمرکز شده است. در شبكه‌هاي اجتماعي مختلف بررسي شد. رتبه‌بندي آستانه خطي را معرفي شده است: يك سنجه مركزيت جديد مبتني بر مدل آستانه خطي كه براي هر عضو، به عنوان تعداد اعضاي در نظر گرفته مي‌شد كه مي‌توانست انتشار نفوذ صورت دهد، ارائه شد.

اين سنجه را با سه سنجه مركزيت ديگر (مركزيت كانز، رتبه صفحه و رتبه بندي آبشاري مستقل) بر اساس نفوذ و تأثير در ۴ شبكه واقعي دو شبكه بزرگ (يكی جهت دار، دیگری بي جهت) و دو شبكه كوچك (يكی جهت دار، دیگری بي جهت) مقايسه شده است. شبكه بزرگ تر ۲۵۵۴۹۱ عضو و ۳۲۸۱۳۲ ارتباط بين اعضا را دارا بود. براي شبكه بزرگ، حتي براي شبكه كوچك جهت دار، در ادامه نشان داده شد كه همبستگي مابين سنجه‌ها، پايين است كه به اين معني است كه همه آن‌ها حوزه نفوذ مختلفی را فراهم مي‌كنند و در نهايت، نتايج گوناگوني را بروز مي‌دهند. به طور كلي، يافته‌ها نشان مي‌دهد كه مقدار حقيقي از نتايج همبستگي در ضريب اسپيرمن بالاتر از ضريب كندال است. علاوه بر اين، رتبه بندي آستانه خطي با رتبه بندي مستقل آبشار، بيشترين ميزان انحراف را حاصل مي‌دهد. براي مقايسه اين سنجه، رتبه بندي آستانه خطي مقادير مختلف بيشترى را ايجاد مي‌كند. اين نشان مي‌دهد كه رتبه بندي آستانه خطي، بسيار پركاربردتر براي رتبه بندي اعضا در يك شبكه خواهد بود. در كنار سنجه مركزيت، در اين مقاله يك سنجه مركزي سازي به نام مركزي سازي آستانه خطي را معرفي شده است؛ كه به تعداد اعضايي كه مي‌توانستند توسط گره‌هايي كه به هسته اصلي شبكه تعلق داشتند تحت تأثير قرار بگيرند تطابق داشت. اثبات شد كه اين سنجه براي شبكه‌هاي با تعداد زياد اعضا بسيار مفيد واقع خواهد شد. همچنين اين سنجه با ضريب خوشه بندي ميانگين نيز مقايسه شد و نتيجه گرفته شد كه هر كدام از اين سنجه‌ها، مركزي سازي متفاوتي را ايجاد مي‌كنند و براي استخراج اطلاعات متفاوتي از شبكه کاربرد دارند.

جديد مرتبط مي‌شوند ($\tau = 0.77$ and $\rho = 0.89$). اين دين معناسست كه براي اين شبكه، گره‌ها (اعضا) مي‌توانند امتياز مقام آستانه خطي جديد بالايي را با رسيدن به سطح گسترش پايين داشته باشند.

جدول (۱): تعداد واكنش دهنده به سطوح گسترش	
Max t	#actors
۰	۴۲۱۲۹
۱	۶۳۱۸۹
۲	۷۳۲۶۰
۳	۵۱۱۱۸
۴	۱۱۵۱۶
۵	۱۴۸۶۱
۶	۲۱۱
۷	۱۷۱
۸	۲۲
۹	۸
۱۰	۴
۱۲	۲



شكل (۲): تعداد واكنش دهنده به سطوح گسترش

پيشنهادات و كارهاى آينده

براى كارهاى آتى، مى‌توان از مركزى سازى فريمن براى برخى از سنجه‌هاى مركزيت مبتنى بر حوزه انتشار به منظور

مقايسه هر چه بيشتر سنجه‌هاى مركزيت با روش مركزى سازى آستانه خطى استفاده كرد.

منابع

- 5-Adamic, L., & Adar, E. (2005). How to search a social network. *Social networks*, 27(3), 187-203.
- 6-Altaf-Ul-Amine, M., Nishikata, K., Korna, T., Miyasato, T., Shinbo, Y., Arifuzzaman, M., ... & Kanaya, S. (2003). Prediction of protein functions based on k-cores of protein-protein interaction networks and amino acid sequences. *Genome Informatics*, 14, 498-499.
- 7-Bader, G. D., & Hogue, C. W. (2003). An automated method for finding molecular complexes in large protein interaction networks. *BMC bioinformatics*, 4(1), 1-27.
- 8-Bavelas, A. (1948). A mathematical model for group structures. *Human organization*, 7(3), 16-30.
- 9-Bollobás, B. (1984). The evolution of random graphs. *Transactions of the American Mathematical Society*, 286(1), 257-274.
- 10-Bonacich, P. (1972). Factoring and weighting approaches to status scores and clique

- identification. *Journal of mathematical sociology*, 2(1), 113-120.
- 11-Corr, P. J., & Matthews, G. E. (2020). *The Cambridge handbook of personality psychology*. Cambridge University Press.
- 12-Domingos, P., & Richardson, M. (2001, August). Mining the network value of customers. In *Proceedings of the seventh ACM SIGKDD international conference on Knowledge discovery and data mining* (pp. 57-66).
- 13-Easley, D., & Kleinberg, J. (2010). *Networks, crowds, and markets: Reasoning about a highly connected world*. Cambridge university press.
- 14-Foster, K. C., Muth, S. Q., Potterat, J. J., & Rothenberg, R. B. (2001). A faster katz status score algorithm. *Computational & Mathematical Organization Theory*, 7, 275-285.
- 15-Freeman, L. C. (2002). Centrality in social networks: Conceptual clarification. *Social network: critical concepts in sociology. Londres: Routledge*, 1, 238-263.
- 16-Gaertler, M., & Patrignani, M. (2004, March). Dynamic analysis of the autonomous system graph. In *IPS 2004, International Workshop on Inter-domain Performance and Simulation, Budapest, Hungary* (pp. 13-24).
- 17-Goldenberg, J., Libai, B., & Muller, E. (2001). Using complex systems analysis to advance marketing theory development: Modeling heterogeneity effects on new product growth through stochastic cellular automata. *Academy of Marketing Science Review*, 9(3), 1-18.
- 18-Gómez, D., Figueira, J. R., & Eusébio, A. (2013). Modeling centrality measures in social network analysis using bi-criteria network flow optimization problems. *European Journal of Operational Research*, 226(2), 354-365.
- 19-Granovetter, M. (1978). Threshold models of collective behavior. *American journal of sociology*, 83(6), 1420-1443.
- 20-Gruhl, D., Guha, R., Liben-Nowell, D., & Tomkins, A. (2004, May). Information diffusion through blogspace. In *Proceedings of the 13th international conference on World Wide Web* (pp. 491-501).
- 21-Katz, L. (1953). A new status index derived from sociometric analysis. *Psychometrika*, 18(1), 39-43.
- 22-Kempe, D., Kleinberg, J. M., & Tardos, É. (2005, July). Influential nodes in a diffusion model for social networks. In *ICALP* (Vol. 5, pp. 1127-1138).
- 23-Kendall, M. G. (1938). A new measure of rank correlation. *Biometrika*, 30(1/2), 81-93.
- 24-Kitsak, M., Gallos, L. K., Havlin, S., Liljeros, F., Muchnik, L., Stanley, H. E., & Makse, H. A. (2010). Identification of influential spreaders in complex networks. *Nature physics*, 6(11), 888-893.
- 25-Langville, A. N., & Meyer, C. D. (2004). Deeper inside pagerank. *Internet Mathematics*, 1(3), 335-380.
- 26-Li, J., Peng, W., Li, T., Sun, T., Li, Q., & Xu, J. (2014). Social network user influence sense-making and dynamics prediction. *Expert Systems with Applications*, 41(11), 5115-5124.
- 27-Liu, W., Yue, K., Wu, H., Li, J., Liu, D., & Tang, D. (2016). Containment of competitive influence spread in social networks. *Knowledge-Based Systems*, 109, 266-275.
- 28-Lozares, C., López-Roldán, P., Bolibar, M., & Muntanyola, D. (2015). The structure of global centrality measures. *International Journal of Social Research Methodology*, 18(2), 209-226.
- 29-Morone, F., & Makse, H. A. (2015). Influence maximization in complex networks through optimal percolation. *Nature*, 524(7563), 65-68.
- 30-Page, L., Brin, S., Motwani, R., & Winograd, T. (1999). *The PageRank citation ranking: Bringing order to the web*. Stanford InfoLab.
- 31-Riquelme, F., Gonzalez-Cantergiani, P., Molinero, X., & Serna, M. (2018). Centrality measure in social networks based on linear threshold model. *Knowledge-Based Systems*, 140, 92-102.
- 32-Saramäki, J., Kivelä, M., Onnela, J. P., Kaski, K., & Kertesz, J. (2007). Generalizations of the clustering coefficient to weighted complex networks. *Physical Review E*, 75(2), 027105.
- 33-Schelling, T. C. (2006). *Micromotives and macrobehavior*. WW Norton & Company.
- 34-Seidman, S. B. (1983). Network structure and minimum degree. *Social networks*, 5(3), 269-287.
- 35-Song, X., Tseng, B. L., Lin, C. Y., & Sun, M. T. (2006, August). Personalized recommendation driven

- by information flow. In *Proceedings of the 29th annual international ACM SIGIR conference on Research and development in information retrieval* (pp. 509-516).
- 36-Sun, J., & Tang, J. (2011). A survey of models and algorithms for social influence analysis. *Social network data analytics*, 177-214.
- 37-Wang, L., Ma, L., Wang, C., Xie, N. G., Koh, J. M., & Cheong, K. H. (2021). Identifying influential spreaders in social networks through discrete moth-flame optimization. *IEEE Transactions on Evolutionary Computation*, 25(6), 1091-1102.
- 38-Watts, D. J., & Strogatz, S. H. (1998). Collective dynamics of 'small-world' networks. *nature*, 393(6684), 440-442.
- 39-Ye, S., & Wu, F. (2013). Measuring message propagation and social influence on Twitter. com. *International Journal of Communication Networks and Distributed Systems*, 11(1), 59-76.
- 40-Zhang, X., Zhu, J., Wang, Q., & Zhao, H. (2013). Identifying influential nodes in complex networks with community structure. *Knowledge-Based Systems*, 42, 74-84.

©Authors, Published by Journal of Intelligent Knowledge Exploration and Processing. This is an open-access paper distributed under the CC BY (license <http://creativecommons.org/licenses/by/4.0/>).

