Journal of Business Intelligence Management Studies--
Vol. 9, No.35, Spring 2021
ims.atu.ac.ir
DOI: 10.22054/IMS.2020.50876.1698

eISSN:2538- ISSN:1735/112X Accepted: 11, 10, 2020 Received: 03, 20, 2020 Original Reserch / Review10,20,2020 / ...

Synergy Creation Of Users In Kousarnet Social Scientific Network Using Graph Based Clustering Methods

Zahra Shirani Department of Information and Computer Engineering, University Of Qom, Qom.

Amir Jalaly Bidgoly * Department of Information and Computer Engineering, University Of Qom, Qom.

Abstract

In recent years, the number of users of social networks has grown significantly. The big challenge for these networks' audience is How to communicate with the people present on these networks. Friend recommender systems try to fix this challenge by offering suggestions. In this study, data from the social and scientific network of Kousarent were used. In this research, using 10 types of relationships between users without considering the friendship relationships, network graph created, and then by using 3 algorithms Louvain, Kmeans and Hierarchical graph clustering was performed to identify communities. Clusters obtained from Louvain's clustering algorithm had higher percentages of matching with friendships. Then, weights were calculated by genetic algorithm for each of 10 relationships and by applying Louvain clustering algorithm on the network graph, the highest percentage of matching with the optimal weight of each of the 10 relationships was obtained. In this case, the resulting clusters are optimal clusters containing the most similar users. So other users in the same cluster can be suggested as friends. The weight of the edges between the individuals in the graph was also used to prioritize the bids. At the end, the friend proposed method was evaluated and the percentage of suggested friends matched with the individual's true friends was calculated.

Keywords: Recommendation Systems, Graph Clustering, Community Detection, Kousarnet Social Science Network.

* Corresponding Author: jalaly@qom.ac.ir
How to Cite: Shirani, Zahra & Jalaly Bidgoly, Amir,(1400). Synergy Creation Of Users In Kousarnet Social Scientific Network Using Graph Based Clustering Methods. Journal of Business Intelligence Management Studies, Vol. 9, No.35, Spring 2021.PP:187-216.

مطالعات مدیریت کسب‌وکار هوشمند
سال نهم، شماره ۳۵، بهار ۱۴۰۰، صص ۲۱۶- ۱۸۷
ims.atu.ac.ir
DOI: 10.22054/IMS.2020.50876.1698

مقاله پژوهشی
تاریخ دریافت: ۱۳۹۹/۰۱/۰۱ ؛ تاریخ بازنگری: ۱۳۹۹/۰۷/۲۹ تاریخ پذیرش: ۱۳۹۹/۰۸/۲۰
eISSN:2538-2527 ISSN:1735/112X

هم‌افزایی کاربران در شبکه علمی اجتماعی کوثرنت با استفاده از روش‌های خوشه‌بندی مبتنی بر گراف

زهرا شیرانی | دانشجوی کارشناسی ارشد رشته فناوری اطلاعات، دانشگاه قم، قم، ایران.

امیر جلالی بیدگلی * | استادیار مهندسی کامپیوتر، دانشگاه قم، قم، ایران.

چکیده

در سال‌های اخیر تعداد کاربران شبکه‌های اجتماعی رشد زیادی داشته‌اند. چالش بزرگ مخاطب این شبکه‌ها، نحوه برقراری ارتباط با افراد حاضر در این شبکه‌ها می‌باشد. سیستم‌های پیشنهاددهنده دوست با ارائه پیشنهاداتی سعی در رفع این چالش دارند. در این پژوهش از داده‌های شبکه علمی و اجتماعی کوثرنت استفاده شده است. در این تحقیق با استفاده از ۱۰ نوع رابطه بین کاربران و بدون در نظر گرفتن روابط دوستی،گراف شبکه ایجاد و سپس با استفاده از ۳ الگوریتم لووین[1] کی‌میانگین[2] و سلسله‌مراتبی[3] خوشه‌بندی گراف جهت تشخیص جوامع انجام گردید. خوشه‌های به دست آمده از الگوریتم خوشه‌بندی لووین دارای درصد مطابقت بالاتری با روابط دوستی بودند. سپس با استفاده از الگوریتم ژنتیک[4] برای هر یک از ۱۰ رابطه وزن‌های مختلفی در نظر گرفته شد و با اجرای الگوریتم خوشه‌بندی لووین بر روی گراف شبکه، بیشترین درصد مطابقت به همراه وزن بهینه هر یک از ۱۰ رابطه به دست آمد. در این حالت خوشه‌های حاصل، خوشه‌هایی بهینه حاوی کاربران با بیشترین شباهت هستند. بنابراین می‌توان سایر کاربرانی که در یک خوشه قرار گرفته‌اند به عنوان دوست به یکدیگر پیشنهاد داد. برای اولویت‌بندی پیشنهادات نیز از وزن یال‌های بین افراد در گراف استفاده شد. در پایان روش پیشنهاد دوست ارزیابی و درصد مطابقت دوستان پیشنهادی با دوستان واقعی فرد محاسبه گردید.

کلیدواژه‌ها: سیستم‌های‌پیشنهاددهنده، خوشه‌بندی‌گراف، تشخیص‌جامعه، شبکه‌علمی اجتماعی‌کوثرنت

* نویسنده مسئول: jalaly@qom.ac.ir

1. Louvain
2. Kmeans
3. Hierarchical
4. Genetic

مقدمه

در سال‌های اخیر تعداد کاربران شبکه‌های اجتماعی رشد چشم‌گیری داشته‌اند. حضور در این شبکه‌های اجتماعی افراد را در معرض ارتباط با حجم انبوهی از اطلاعات و افراد قرار داده است. چالش بزرگی که ذهن مخاطب این شبکه‌ها را به خود مشغول کرده است نحوه برقراری ارتباط با افراد حاضر در این شبکه‌ها می‌باشد. سیستم‌های پیشنهاددهنده دوست با ارائه پیشنهاداتی برای کاربر سعی در رفع این چالش دارند. در ابتدا در شبکه‌های اجتماعی عمومی ارائه پیشنهادات توسط سیستم‌های پیشنهاددهنده تنها با هدف دوستی و با استفاده از ویژگی‌های عمومی کاربر انجام می‌شد. به عنوان مثال فیس‌بوک به وسیله ویژگی به نام «افرادی که شاید شما بشناسید» افرادی را براساس رویکرد دوستِ دوست به کاربران معرفی می‌کند (چن و همکاران[1] ۲۰۰۹، ص ۲۰۱، سامر و آیاد[2] ۲۰۱۹).

با گذشت زمان و گذر از دوره شبکه‌های اجتماعی عمومی برای هر کاربرد خاص یک شبکه اجتماعی خاص با ویژگی‌های خاص به وجود آمد. به عنوان مثال شبکه اجتماعی گودریدز[3] که در آن هر فردی هرکتابی را که می‌خواند ثبت می‌نماید. در این سیستم پیشنهاد دوست براساس کتاب‌های مشابهی که خوانده شده انجام می‌شود. به عنوان مثالی دیگر افراد در شبکه اجتماعی لینکدین[4] پروفایلی از سوابق تحصیلی، سوابق کاری، تخصص‌ها و موقعیت شغلی فعلی خود قرار می‌دهند بنابراین پیشنهادات دوست در این سیستم براساس سوابق تحصیلی یا کاری مشترک و ... خواهد بود. سیستم‌های خاص منظوره دارای ویژگی‌های اختصاصی می‌باشند و اهداف خاصی را نیز دنبال می‌کنند. مثلا در شبکه‌های اجتماعی مبتنی بر مکان پیشنهاد دوست براساس موقعیت جغرافیایی کاربر خواهد بود (بائو و همکاران[5] ۲۰۱۵، ص ۵۲۵). در برخی از شبکه‌های اجتماعی، پیشنهادات براساس سبک زندگی کاربر (ونگ و همکاران[6] ۲۰۱۵، ص ۵۳۸) و یا میزان مطالعه کاربر در طول روز انجام می‌شود.

همچنین در برخی از شبکه‌های اجتماعی می‌توان با درنظرگرفتن علایق شخصی فرد

1. Chen et al.
2. Sammer& Ayad
3. Goodreads
4. Linkedin
5. Bao et al.
6. Wang et al.

پیشنهادات را به صورت شخصی سازی شده ارائه داد و برای بعضی از ویژگی‌های موردنظر کاربر وزن بیشتری درنظر گرفت (دینگ و همکاران[1] ۲۰۱۷، ص ۱۴۷۹). به عنوان مثال ممکن است فاصله جغرافیایی دوستان و یا رژیم غذایی و فعالیت بدنی برای برخی از کاربران مهم باشد.

با مطرح شدن بحث شبکه‌های اجتماعی، هر کدام از سازمان‌های بزرگ، شبکه اجتماعی مخصوص خود را به وجود آوردند. به عنوان مثال شبکه اجتماعی بی هایو[2] برای سازمان آی بی ام[3] (چن و همکاران[4] ۲۰۰۹، ص ۲۰۱) و یا شبکه اجتماعی کوثرنت برای طلاب خانم حوزه علمیه طراحی شده است و یا این که هر دانشگاهی شبکه اجتماعی مخصوص به خود را به وجود آورده است. در واقع یک شبکه اجتماعی سازمانی، شبکه‌ای است که برای افراد داخل همان سازمان طراحی شده است. شبکه‌های سازمانی شبکه‌هایی مختص مشتریان و کارمندان یک سازمان خاص هستند و معمولا این‌گونه شبکه‌ها به سایر سیستم‌های سازمان نیز متصل بوده و تمام اطلاعات سازمانی در اختیار جهت بهبود شبکه استفاده می‌شود. از این رو حجم اطلاعات در اختیار یک شبکه سازمانی بسیار بیشتر و دقیق‌تر از سایر شبکه‌های اجتماعی مطرح است. شبکه اجتماعی سازمانی برای اتصال به بقیه سیستم‌های سازمان به واسطی مانند گذرگاه خدمات سازمانی (ESB) [5] نیاز دارد.

گذرگاه خدمات سازمانی روشی برای اتصال شبکه به سایر سیستم‌های سازمان است. در واقع در این روش انواع مختلف برنامه‌های سازمانی به یک گذرگاه پیام متصل شده، و به برقراری ارتباط با یکدیگر می‌پردازند. این عملکرد باعث از بین رفتن وابستگی میان برنامه‌های مختلف می‌شود، و اجازه می‌دهد این برنامه‌ها بدون نیاز به داشتن اطلاع از دیگر برنامه‌های روی گذرگاه تنها با گذرگاه ارتباط برقرار کنند.

معرفی شبکه اجتماعی کوثرنت

شبکه اجتماعی کوثرنت یک شبکه سازمانی خاص منظوره برای طلاب خانم سراسر کشور می‌باشد و در حال حاضر در حدود صد هزار نفر عضو دارد. این شبکه به واسطه گذرگاه

1. Ding et al.
2. Beehive
3. IBM
4. Chen et al.
5. Enterprise Service Bus (ESB)

خدمات سازمانی به سیستم‌های سازمان متصل است. در این شبکه، فرایند تجمیع خدمات برای طلاب به‌منظور تسهیل در دسترسی به خدمات و افزایش بهره‌وری در استفاده‌ی بیش از بیست سیستم نرم‌افزاری مختلف حاصل می‌شود ازجمله: پذیرش (سنجش بدو ورود)، آموزش، آموزش مجازی (ارائه دروس و محتوا به‌صورت غیرحضوری)، بانک مقالات و پایان‌نامه‌ها، سیستم پژوهش، هم‌مباحثه‌ای، فراخوان‌های پژوهشی و المپیادهای علمی، سامانه وبلاگ (کوثربلاگ)، کتاب‌داری، گروه‌ها، انجمن‌های علمی اساتید و ... . در این شبکه اجتماعی هر فرد دارای یک نمایه است که شامل اطلاعات شخصی، اطلاعات سطح طلبه و آمار فعالیت در شبکه می‌باشد. صفحه اصلی سایت که کاربر پس از وارد شدن به سامانه مشاهده می‌کند رواق نام دارد. رواق شامل پست‌های کاربران شبکه می‌باشد و کاربران دیگر می‌توانند نظرات خود را برای پست دیگران بنویسند و یا با دادن احسنت پست دیگران را تایید کنند. همچنین در همین صفحه کاربر می‌تواند پست خود را ایجاد نماید و به نمایش بگذارد.

یکی از بخش‌های شبکه اجتماعی کوثرنت پیشنهاد هم‌بحث می‌باشد. با توجه به حجم وسیع اطلاعاتی که شبکه در مورد کاربر در اختیار ما می‌گذارد می‌توان هم‌بحث مورد نظر کاربر را پیشنهاد داد. ارائه پیشنهاد هم‌بحث می‌تواند با در نظر گرفتن اهداف گوناگون انجام شود و با توجه به همین اهداف تعریف‌های متعددی ارائه داد.

حجم اطلاعات در اختیار شبکه کوثرنت بسیار جامع‌تر و کامل‌تر از سایر شبکه‌های اجتماعی متداول است از این رو نیاز است سیستم پیشنهاددهنده هم‌بحث نیز به طور اختصاصی برای آن طراحی شود. هدف در پژوهش جاری یک روش پیشنهاد هم‌بحث یا دوست در شبکه‌های اجتماعی علمی و به طور خاص این سیستم است.

اهداف پژوهش

هدف اصلی این پژوهش ایجاد سیستم پیشنهاددهنده دوست برای شبکه علمی و اجتماعی کوثرنت می‌باشد. در این تحقیق تلاش شده است پیشنهاد هم‌بحث یا دوست با تشکیل گراف شبکه که در آن کاربران به عنوان گره‌های گراف و ارتباطات میان آن‌ها به صورت یال بین گره‌ها هستند انجام شود. سپس با در نظر گرفتن وزن بهینه برای هر یک از روابط موجود بین کاربران و سپس تقسیم‌بندی کاربران به خوشه‌ها یا جوامع، افراد شبیه به یکدیگر در یک خوشه قرار گیرند. سپس به دلیل وجود شباهت زیاد بین کاربران موجود

در یک خوشه می‌توان افرادی که داخل یک خوشه قرار گرفته‌اند و تا به حال رابطه دوستی با یکدیگر برقرار ننموده‌اند به عنوان دوست به یکدیگر پیشنهاد داد. روش مورد استفاده در این پژوهش برای یافتن جوامع مناسب الگوریتم لووین می‌باشد، که یکی از الگوریتم‌های خوشه‌بندی مبتنی بر گراف است.

این الگوریتم با استفاده از اطلاعات شبکه و با در نظر گرفتن اطلاعات کاربر، افراد را به جوامع مناسب تقسیم می‌نماید. همچنین وزن بهینه هریک از روابط موجود بین کاربران از طریق الگوریتم ژنتیک به دست می‌آید. به دلیل این‌که ممکن است افراد زیادی در یک خوشه قرار بگیرند و برای پیشنهاد دوست به کاربر افراد زیادی در لیست پیشنهادی قرار بگیرند، در این تحقیق از یک روش اولویت بندی بر اساس وزن یال‌های موجود بین کاربران در گراف شبکه کوثرنت استفاده شده است. بنابراین با بررسی ویژگی‌های موجود در شبکه مناسب‌ترین کاربران برای دوستی به یکدیگر پیشنهاد داده می‌شوند. سپس با توجه به روابط دوستی فعلی کاربران، دوستان پیشنهادی به کاربران مورد ارزیابی قرار گرفته و درصد مطابقت دوستان پیشنهادی با دوستان واقعی افراد به دست می‌آید و دقت روش محاسبه می‌گردد.

پیشینه پژوهش

مقدمه‌ای بر خوشه‌بندی

امروزه با حجم عظیمی از داده‌ها روبرو هستیم که برای استفاده از آن‌ها به ابزارهای کشف دانش نیاز داریم. داده‌کاوی یک توانایی پیشرفته، در تحلیل داده و کشف دانش می‌باشد. یکی از مهم‌ترین روش‌های داده‌کاوی، خوشه‌بندی است. خوشه‌بندی، گروه‌بندی داده ها است، به نحوی که داده‌ها در خوشه‌های یکسان بیشترین شباهت و در خوشه‌های متفاوت کمترین شباهت را با یکدیگر داشته باشند. در ذیل برخی از معروف‌ترین روش‌های خوشه‌بندی ارائه شده است (برخین و همکاران[1] ۲۰۰۶، ص ۲۵) .

1. Berkhin et al.

روش‌های خوشه‌بندی مبتنی بر تقسیم‌بندی

روش‌های تقسیم‌بندی، داده‌ها را بر اساس معیار تشابه به تعداد خاص خوشه‌ها تقسیم‌بندی می‌کنند. معروف‌ترین الگوریتم در این گروه الگوریتم کی‌میانگین می‌باشد (فیلیپن و همکاران[1] ۲۰۰۸، ص ۱۷۶) (لیکاس و همکاران[2] ۲۰۰۳، ص ۴۵۱).

روش‌های خوشه‌بندی سلسله مراتبی

الگوریتم‌های خوشه‌بندیِ سلسله مراتبی به دو دوسته متراکم‌شونده[3] و تقسیم‌شونده[4] تقسیم می‌شوند. در روش متراکم‌شونده که به آن روش پایین به بالا نیز گفته می‌شود، هر نمونه ابتدا یک خوشه است و در هر مرحله نمونه‌ها به هم می‌چسبند تا خوشه‌های بزرگ‌تر را تولید کنند و در نهایت همه‌ی نمونه‌ها با هم یک خوشه‌ی بزرگ را تشکیل دهند. ولی روش تقسیم‌شونده برعکس روند متراکم‌شونده می‌باشد (ژائو و همکاران[5] ۲۰۰۵، ص ۱۴۱).

روش‌های ارزیابی خوشه بندی

خوشه‌بندی یک روش بدون نظارت محسوب می‌شود و معمولا ارزیابی نتایج خوشه‌بندی براساس تابع فاصله و برچسب‌های تولید شده در تحلیل خوشه‌بندی انجام می‌شود. ارزیابی نتایج خوشه‌بندی می‌تواند از دو منظر درونی و بیرونی بررسی شود (آمیگو و همکاران[6] ۲۰۰۹، ص ۴۶۱).

مقدمه‌ای بر خوشه‌بندی گراف

گراف‌ها به عنوان ساختارهای ریاضی که روابط اشیا با هم را در سطح انتزاع مناسبی نشان می‌دهند به طور گسترده در مدل‌سازی مسائل مختلف مورد استفاده قرار گرفته‌اند. در خوشه‌بندی، داده‌ها را در گروه‌هایی تقسیم می‌کنیم که داده‌های مربوط به هر گروه

---

1. Filippone et al.
2. Likas et al.
3. Agglomerative
4. Partitioning
5. Zhao et al.
6 Amigó et al.

ویژگی‌های نزدیک به هم داشته باشند و داده‌های موجود در دو گروه مختلف، ویژگی‌هایی با اختلاف زیاد داشته باشند. روش‌های خوشه‌بندی گراف به سه دسته کلی زیر تقسیم می‌شوند.

۱- روش‌های خوشه‌بندی گراف مبتنی بر پیمانگی(به عنوان مثال روش لووین، روش گیروان و نیومن[1] و ...)

۲- روش‌های خوشه‌بندی مبتنی بر فاصله (به عنوان مثال روش کی‌میانگین، روش K-Medoids و ...)

۳- روش‌های خوشه‌بندی سلسله مراتبی

توضیح الگوریتم لووین

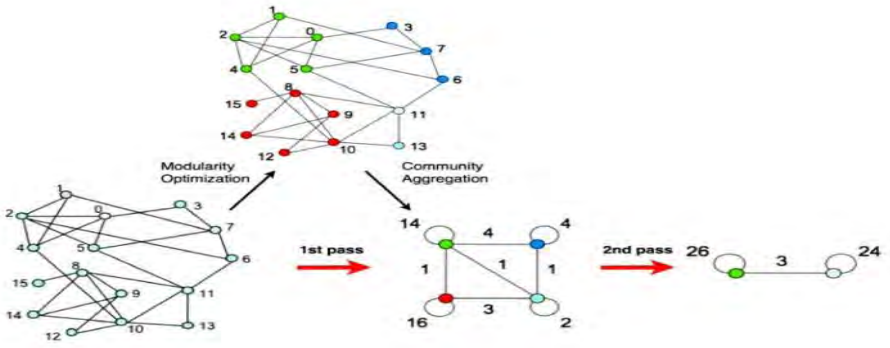
روش لووین (بلاندل و همکاران[2] ۲۰۰۸، ص ۱۰۰۰۸) یک الگوریتم ساده و حریصانه برای یافتن انجمن‌ها در شبکه‌های بزرگ است. در این روش محدودیت سایز گراف ورودی به محدودیت حافظه ما بستگی دارد، نه به محدودیت زمانی. به همین دلیل، این الگوریتم به راحتی بر گراف‌هایی با صدها میلیون راس و به صورت توزیع شده قابل اعمال است. این الگوریتم با استفاده از روش‌های حریصانه بر روی ماژولاریتی، این روش را بهینه کرده و به خوشه‌بندی رأس‌های گراف می‌پردازد. این بهینه‌سازی در ۲ مرحله انجام می‌شود:

مرحله اول: بابهینه‌سازی محلی، این الگوریتم به دنبال گروه‌های کوچک می‌گردد. فاز اول در یک نقطه بهینه محلی متوقف می‌گردد. نقطه‌ای که ماژولاریتی بیشتر با تغییر انجمن هیچ رأسی به دست نمی‌آید.

پژوهشگاه علوم انسانی و مطالعات فرهنگی
پرتال جامع علوم انسانی

1. Girvan & Newman
2. Blondel et al.

شکل ۱. نمایش چگونگی اجرای الگوریتم Louvain



مرحله دوم: سپس با ادغام گروه‌های کوچک که توانایی ایجاد گروه‌های بزرگ‌تر را دارند خوشه‌بندی را ادامه می‌دهد. این مراحل تکرار می‌شود تا تغییری در خوشه‌ها بدست نیاید و به مقدار بیشینه ماژولاریتی برسیم.

شکل بالا نمایش تصویری مراحل الگوریتم Louvain می‌باشد. این الگوریتم برای گراف‌های خلوت، مرتبه خطی دارد. دلیل سرعت بالای این روش، ساده بودن محاسبه، تفاوت در میزان پیمانگی و کاهش سایز گراف بعد از تعداد کمی تکرار است (بلاندل و همکاران، ۲۰۰۸، ص ۱۰۰۰۸).

مقدمه‌ای بر سیستم‌های پیشنهاددهنده

یک سیستم هوشمند که بتواند در میان حجم انبوهی از اطلاعات گوناگون، مناسب‌ترین و مورد پسندترین‌ها را بنا به شرایط و ویژگی‌های خاص هر کاربر به او پیشنهاد کند، سیستم‌های پیشنهاددهنده نام دارد. هدف اصلی سیستم‌های پیشنهاددهنده دوست، کمک به برقراری رابطه جدید و ایجاد یال‌های جدید بین کاربران شبکه است.

در تحقیقات اخیر در مورد شبکه‌های بزرگ و پیچیده، شبکه‌های اجتماعی از لحاظ ساختاری مورد تجزیه و تحلیل قرار گرفته‌اند. در این تحلیل‌ها شبکه به صورت یک گراف کلی در نظر گرفته می‌شود، گره‌ها نمایانگر افراد شبکه و یال‌های آن نشان‌دهنده تعامل، همکاری و یا نفوذ بین اشخاص است. شبکه‌های اجتماعی بسیار پویا هستند؛ آن‌ها از طریق اضافه شدن لبه‌های جدید به سرعت تغییر می‌کنند، که نشان‌دهنده ظهور تعاملات جدید در

ساختار اجتماعی پایه است. شناسایی مکانیسم‌هایی که شبکه‌های اجتماعی تکامل می‌یابند، یک سوال اساسی است که هنوز کاملا درک نشده و انگیزه کار ما را شکل می‌دهد (لیبن-نوول و کلینبرگ[1] ۲۰۰۷، ص ۱۰۱۹) (سامر و همکاران، ۲۰۱۹، ص ۴۲۲).

انواع الگوریتم‌های مورد استفاده در سیستم‌های پیشنهاددهنده

سیستم‌های پیشنهاددهنده دوست، شبکه اجتماعی عمومی را به صورت یک گراف کلی در نظر می‌گیرند و از الگوریتم‌های مختلفی برای پیشنهاد افراد جدید استفاده می‌کنند. مهم‌ترین این روش‌ها در ذیل تشریح می‌گردد.

الگوریتم تطابق محتوا

الگوریتم تطابق محتوا کاربرانی را که دارای محتوای مشابه هستند معرفی می‌نماید. این مدل، ابتدا با استفاده از اطلاعات متنی مربوط به کاربر، مجموعه‌ای از لغات ایجاد می‌کند که به عبارتی معرف کاربر مورد نظر است. برخی از این لغات عبارتند از: کلمات موجود در صفحه‌ی مشخصات کاربر، عنوان شغلی، شهر محل زندگی و کار کاربر و به طور کلی هر گونه محتوای متنی که کاربر به اشتراک گذاشته باشد (یو و همکاران[2] ۲۰۱۵، ص ۱۰۲). سپس سیستم پیشنهاددهنده کاربرانی که کلمات کلیدی مشترک زیادی با یکدیگر دارند را به یکدیگر پیشنهاد می‌دهد (چن و همکاران، ۲۰۰۹، ص ۲۰۱).

در یک مطالعه، سیستم توصیه دوست مبتنی بر هم‌بستگی و انسجام بررسی شده است، که کاربران با محتوای مشابه را در خوشه‌های خاص طبقه‌بندی می‌کند و افرادی که در یک خوشه قرار گرفتند را به یکدیگر پیشنهاد می‌دهد(حمید و همکاران[3] ۲۰۱۴، ص ۱۷۶).

همچنین مطالعه‌ای در مورد دانشجویان مهندسی انجام شده است تا عوامل مهم اجتماعی را که در نظر کاربران در هنگام ارسال یا پذیرش درخواست‌های دوستی مورد توجه قرار می‌گیرد، شناسایی کنند. براساس پاسخ‌های جمع‌آوری شده، رگرسیون لجستیک و مدل‌های شبکه‌های عصبی مصنوعی برای پیش‌بینی ویژگی‌های درخواست‌های

1. Liben-Nowell & Kleinberg
2. Yu et al.
3. Hamid et al.

کاربران استفاده شده‌اند (سوار ناکار و همکاران[1] ۲۰۱۷، ص ۲۸۷). در مطالعه دیگری امکان استفاده از نظرات افراد محبوب جامعه (رهبران عقیده) بررسی شده و نشان داده شده است که می‌توان بر اساس این اطلاعات دقت سیستم پیشنهاددهنده را بهبود داد (ونگ[2] و همکاران، ۲۰۲۰).

الگوریتم محتوا به علاوه ارتباط

در این روش معیار دیگری به نام ارتباط اجتماعی قابل قبول بین کاربران، علاوه بر کلمات کلیدی در معرفی کاربران مورد استفاده قرار می‌گیرد و افرادی را که کلمات کلیدی مشابهی در لیست لغاتشان دارند و بین آن‌ها حداقل یک رابطه اجتماعی قابل قبول وجود دارد، به یکدیگر معرفی می‌کند (چن و همکاران، ۲۰۰۹، ص ۲۰۱). در پژوهش‌های جدیدتر از روش‌های مانند یادگیری عمیق برای تحلیل محتوی در کنار تحلیل‌های اجتماعی مانند تحلیل روابط برای تخمین میزان قابل اعتماد بودن کاربران استفاده شده است (خرد و جلالی[3] ۲۰۲۰)

الگوریتم دوستِ دوست

الگوریتم دوستِ دوست یکی از معروف‌ترین الگوریتم‌های پیشنهاددهنده دوست مبتنی بر ساختار شبکه می‌باشد. این الگوریتم بر این اصل استوار است که اگر برخی از دوستان X، با کاربر Y ارتباط دارند، به احتمال زیاد نیز شخص X خواهان ارتباط با کاربر Y می‌باشد. ایده اصلی این الگوریتم این است که یک کاربر با احتمال بیشتر دوستانِ دوستانش را می‌شناسد، تا هر شخص تصادفی دیگر(چن و همکاران، ۲۰۰۹، ص ۲۰۱). شبکه اجتماعی فیس‌بوک از این الگوریتم استفاده نموده است. (ژنگ و وو[4] ۲۰۱۷، ص ۱).

در یکی از مراجع، یک سیستم پیشنهاددهنده دوست مبتنی بر ساختار گراف شبکه و الگوریتم ژنتیک معرفی شده است. در این روش برای پیشنهاد دوست به یک کاربر، زیرگراف شامل آن کاربر و ارتباط‌های تا ۳ درجه فاصله از آن کاربر مورد بررسی قرار

1. Swarnakar et al.
2. Weng et al.
3. Kherad & Jalaly
4. Zheng & Wu

می‌گیرد. در این روش پس از تشکیل زیرگراف مربوط به هر کاربر، تعدادی ارتباط (یال) جدید برای آن کاربر با استفاده از الگوریتم ژنتیک به‌دست می‌آید. این الگوریتم شامل دو فاز است: ۱- فاز فیلتر کردن گراف کلی شبکه و تشکیل زیرگراف مربوط به هر کاربر،و ۲- پیشنهاد تعدادی از افرادی که در این زیرگراف حضور دارند، به کاربر موردنظر، با استفاده از الگوریتم ژنتیک. در این روش از سه معیار ارزیابی برای پیشنهاد دادن نود j به نود مرکزی زیرگراف (نود i) استفاده می‌گردد. اولین معیار بیانگر تعداد دوست‌های مشترک کاربر i و کاربر j می‌باشد. دومین معیار تعداد ارتباط‌های مشترک بین دوستان کاربر i و کاربر j است. سومین معیار، اجتماع دوستان مشترک کاربر i و j می‌باشد. سپس هر یک از این معیارها در یک ضریب وزنی ضرب شده و ارزیابی نهایی برای پیشنهاد کاربر j به کاربر i از مجموع حاصل‌ضرب وزنی این سه معیار حاصل می‌گردد، که مقادیر این ضرایب وزنی توسط الگوریتم ژنتیک بهینه می‌شود (سیلوا و همکاران[1] ۲۰۱۰، ص ۱).

الگوریتم سونار

الگوریتم سونار اطلاعات روابط اجتماعی را از منابع داده جمع‌آوری می‌کند و در صورتی که دو کاربر در منبع داده‌ای با یکدیگر تعامل داشته باشند، آن‌ها را به یکدیگر معرفی می‌کند (چن و همکاران، ۲۰۰۹، ص ۲۰۱).

روش‌های فیلترینگ مشارکتی

روش‌های فیلترینگ مشارکتی بر مبنای جمع‌آوری و تحلیل یک حجم زیاد داده مربوط به رفتار، فعالیت یا ترجیحات کاربر می‌باشد و بر مبنای شباهت کاربر با سایر کاربرها پیش‌بینی می‌کنند که کاربر چه چیزی را دوست خواهد داشت. یکی از رایج‌ترین انواع فیلترینگ مشارکتی، فیلترینگ مشارکتی آیتم – آیتم می‌باشد (فردی که آیتم X را می‌خرد آیتم Y را هم خواهد خرید) (وان و همکاران[2] ۲۰۱۳، ص ۱۰۴۵).

در حال حاضر برای هر کاربرد خاص، یک شبکه اجتماعی خاص با اهداف و ویژگی‌های خاص پدید می‌آید. بنابراین روش‌های پیشنهاد دوست در این شبکه‌ها نیز از ویژگی‌های خاص جهت پیشنهاد دوست استفاده می‌نمایند. به عنوان مثال می‌توان به

1. Silva et al.
2. Wan et al.

پیشنهاد کاربران با سبک زندگی مشابه اشاره نمود. این روش کتاب دوست[1]نام دارد (دو و همکاران[2] ۲۰۱۴، ص ۴۵۷). در این روش با استفاده از اطلاعات سنسورهای گوشی کاربران، افرادی که دارای سبک زندگی مشابه هستند به یکدیگر توصیه می‌شوند (ونگ و همکاران، ۲۰۱۵، ص ۵۳۸).

پیشنهاد دوست در شبکه‌های اجتماعی خاص منظوره

برای پیشنهاد دوست در شبکه‌های اجتماعی خاص منظوره علاوه بر استفاده از ویژگی‌های خاص می‌توانیم علاقه‌مندی‌های فرد را نیز در نظر گرفته و روش‌های پیشنهاد دوست را با توجه به علایق کاربر و به صورت شخصی سازی شده ارائه دهیم (ژائو و همکاران[3] ۲۰۱۷، ص ۳۲۱). پژوهش‌های مختلفی در ارتباط با موضوع پیشنهاد دوست خاص منظوره انجام شده است. لی و همکاران (لی[4]و همکاران، ۲۰۲۰) در پژوهشی به بررسی پیشنهاد دوست در شبکه اجتماعی به صورت تفکیک شده برای دایره‌های خاص دوستی[5]پرداخته است. زو و همکاران[6]به مطالعه پیشنهاد دوست در شبکه‌های دانشگاهی پرداخته است. در مطالعه‌ای مسئله پیشنهاد دوست به وسیله مدل جدید شبکه‌های عمیق عصبی بیزی (BayDNN)[7]با ترکیب رتبه‌بندی شخصی بررسی شده است. در BayDNN، از یک شبکه عصبی یک بعدی برای استخراج ویژگی‌ها از ساختار اطلاعات شبکه استفاده شده، و سپس از یک یادگیری رتبه‌بندی شده شخصی بیزی برای تطبیق ترجیحات کاربران براساس ویژگی‌های استخراج شده استفاده شده است. (چن و همکاران، ۲۰۰۹، ص ۲۰۱) (سومان و همکاران، ۲۰۲۰، ص ۱۲۵ سامون[8]و همکاران، ۲۰۲۰).

شبکه‌های اجتماعی مکان محور نوع دیگری از شبکه‌های اجتماعی خاص منظوره هستند. در این شبکه‌ها، روش فیلترینگ مشارکتی، ترجیحات کاربر را از سابقه رفتارها (از

1. Friendbook
2. Du et al.
3. Zhao et al.
4. Li
5. Friends circles
6. Xu et al.
7. Bayesian deep neural networks (BayDNN)
8. Suman

جمله سابقه موقعیت مکانی) برمی‌دارد و با توجه به آن کاربران را به یکدیگر توصیه می‌کند (چو و همکاران[1] ۲۰۱۳، ص ۳۶۵) (بائو و همکاران، ۲۰۱۵، ص ۵۲۵).

مروری بر پژوهش با موضوع سیستم پیشنهاددهنده هم‌پژوهشی در سازمان کوثرنت

پژوهشی(جلالی، ۱۳۹۳) در سال ۱۳۹۳ با موضوع ارائه سیستم پیشنهاددهنده هم‌پژوهشی در این زمینه انجام گرفته است. این سیستم پیشنهاددهنده از دو منبع داده محتوایی و اطلاعات کاربر استفاده می‌کند. منابع محتوایی مربوط به عناوین پایان‌نامه‌هایی می‌باشد که افراد درگذشته انتخاب کرده‌اند و سیستم پیشنهاددهنده محتوایی نیز بر پایه همین عناوین بنا شده‌است. از سوی دیگر اطلاعاتی در پایگاه‌های داده دیگر واحد ارتباط با فناوری اطلاعات حوزه وجود دارند که از آن‌ها در راستای تشکیل پروفایل کاربران، استفاده شده است. این منابع شامل اطلاعات مدارس طلاب، اطلاعات بخش آموزش مجازی، وبلاگ‌ها، المپیاد، ثبت نام و دوره‌های تحصیلی می‌باشند. هدف از طراحی سیستم پیشنهاددهنده مبتنی بر محتوا که به طور مشخص از داده‌های مربوط به عناوین پایان‌نامه‌ها بنا نهاده می‌شود، معرفی افراد دارای زمینه پژوهشی مشترک مربوط به پایان نامه به یکدیگر می‌باشد اما در مواقعی ممکن است کاربر هنوز به مرحله مشخص کردن موضوع پایان‌نامه‌اش نرسیده باشد در این شرایط یک سیستم جایگزین به کار خواهد آمد که بر اساس سایر مشخصات و ویژگی‌های افراد بتواند پیشنهاداتی را به کاربر ارائه کند. در این مرحله با استفاده از الگوریتم‌های مختلف خوشه‌بندی، طبقه‌بندی و محاسبه شباهت، کاربرانی که دارای بیشترین میزان تشابه با کاربر بوده‌اند پیشنهاد داده می‌شوند.

در پژوهش انجام شده معرفی کاربران شبیه به هم صرفا در راستای موضوع پایان‌نامه انجام گرفته و با توجه به موضوع پژوهش مرتبط با پایان‌نامه انتخابی به ایجاد ارتباط بین افراد می‌پردازد. در این تحقیق از اطلاعات شبکه استفاده نشده است. همین علت‌ها سبب شد این پژوهش خروجی قابل قبولی نداشته و توسط سازمان استفاده نشود.

---

1. Chu et al.

آزمایش‌های انجام شده و تحلیل‌ها

در سال‌های اخیر تعداد کاربران شبکه‌های اجتماعی رشد چشم‌گیری را داشته‌اند. حضور در این شبکه‌های اجتماعی افراد را در معرض ارتباط با حجم انبوهی از اطلاعات و افراد قرار داده است. چالش بزرگی که ذهن مخاطب این شبکه‌ها را به خود مشغول کرده است نحوه برقراری ارتباط با افراد حاضر در این شبکه‌ها می‌باشد. سیستم‌های پیشنهاددهنده دوست با ارائه پیشنهاداتی برای کاربر سعی در رفع این چالش دارند.

توصیف پایگاه داده مورد استفاده

در این پژوهش از داده‌های کاربران شبکه علمی و اجتماعی کوثرنت استفاده شده است. اطلاعات ۲۰۸۷ نفر از کاربران این شبکه موجود می‌باشد که شامل موارد زیر می‌باشد:

استان، شهر، مدرسه، سطح (سطح ۲، سطح ۳، سطح ۴)، بخش (حضوری یا نیمه حضوری)، اساتید ترم جاری، درس‌های ترم جاری، کلاس‌های ترم جاری، پست‌هایی که مورد پسند کاربر قرار گرفته است و گروه‌های عضو

تعداد روابط کاربران شبکه در هر یک از ۱۰ مورد بالا به شرح شکل (۱) می‌باشد.

شکل ۱. توزیع هر یک از ۱۰ رابطه موجود بین کاربران

همچنین روابط دوستی فعلی این ۲۰۸۷ کاربر که شامل ۷۰۴۷ رابطه دوستی می‌باشد موجود است که فقط در مرحله ارزیابی از این روابط دوستی استفاده می‌گردد. در حال حاضر تعداد این روابط دوستی در شبکه کوثرنت بسیار کم می‌باشد و افراد با تعداد نفرات

کمی ارتباط دوستی برقرار نموده‌اند. مطابق با نمودار پراکندگی ذیل به عنوان مثال از ۲۰۸۷ کاربر ۲۹۶ کاربر فقط با ۱ نفر ارتباط دوستی دارند و ۱۵۶ کاربر فقط با ۲ نفر ارتباط دوستی دارند. یعنی حدود یک چهارم از کاربران تنها با ۱ یا ۲ نفر از سایر کاربران رابطه دوستی دارند. بقیه کاربران نیز با تعداد افراد کمی ارتباط دارند. مطابق شکل (۲) با افزایش تعداد افراد، تعداد دوستان فرد کاهش می‌یابد.

شکل ۲. پراکندگی تعداد افراد مرتبط با یکدیگر

تشکیل گراف شبکه علمی و اجتماعی کوثرنت

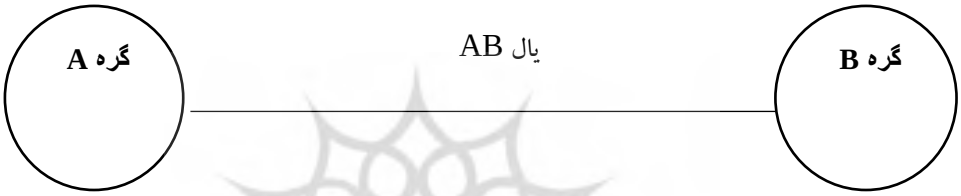
با استفاده از اطلاعات و روابط ۲۰۸۷ کاربر موجود و بدون در نظر گرفتن روابط دوستی گراف شبکه را ایجاد می‌نماییم. کاربران گره‌های گراف می‌باشند و ۱۰ رابطه بین کاربران که در بالا ذکر شد به صورت یال وزن‌دار به گراف اضافه می‌شوند. رابطه‌های افراد ذاتا دو طرفه می‌باشد و یال‌ها در گراف بدون جهت در نظر گرفته شده‌اند. در ابتدا وزن هریک از ۱۰ رابطه موجود بین کاربران برابر با ۱ در نظر گرفته شده است. به عنوان مثال اگر A و B دو گره از گراف باشند و بین این دو گره هر یک از ۱۰ رابطه برقرار باشد شکل (۳) را خواهیم داشت.

همانطور که در شکل (۳) مشاهده می‌کنید بعد از افزودن گره‌ها به گراف شبکه برای اضافه نمودن یال‌ها فایل اطلاعات هر یک از ۱۰ رابطه بررسی می‌شود و در صورت وجود رابطه یک یال بین دوگره رسم می‌شود و به ازای هر رابطه وزن ۱ به وزن یال اضافه

می‌گردد. حتی با وجود چندین رابطه بین دو گره فقط ۱ یال بین دو گره رسم می‌شود و با وجود هر رابطه وزن یال یک واحد افزایش می‌یابد. به عنوان مثال اگر بین دو گره c و d، ۳ رابطه هم‌استانی، هم‌شهری و هم مدرسه‌ای برقرار باشد بین این دو گره یک یال با وزن ۳ ایجاد خواهد شد. مشخصات گراف شبکه نیز به شرح جدول (۱) می‌باشد.

شکل ۳. نمایش دو گره از گراف و یال وزن‌دار بین دو گره

$$W = W_1 + W_2 + W_3 + \ldots + W_{10}$$



W = وزن یال ، W_1 = وزن هم‌استانی= ۱ ، W_2 = وزن هم‌شهری = ۱ ، . . . ،

W_{10} =وزن هم‌گروهی= ۱

جدول۱. مشخصات گراف شبکه

تعداد گره	۲۰۸۷
تعداد یال	۲۱۷۶۶۶۱
میانگین درجه گره‌های گراف	۲۰۸۵/۹۲۳۳
چگالی گراف	۰/۹۹۹۹۶

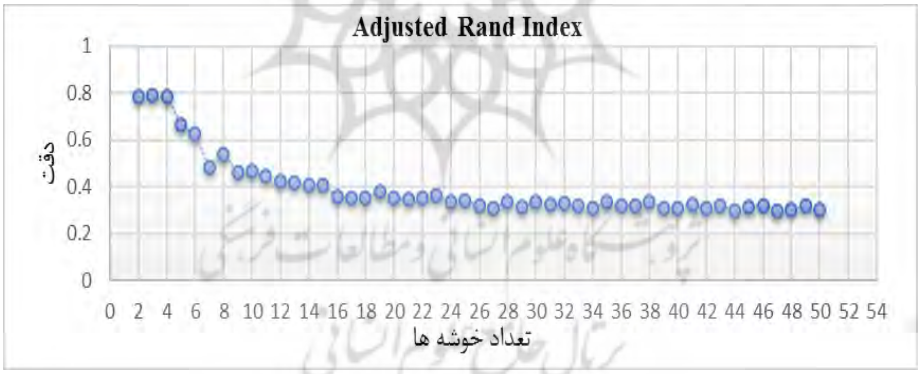
تشخیص جوامع

برای یافتن جوامع یا خوشه‌های منسجم گراف که افراد داخل خوشه بیشترین شباهت به یکدیگر و افراد در خوشه‌های متفاوت کمترین شباهت را داشته باشند از الگوریتم‌های لووین، کی‌میانگین و سلسله‌مراتبی استفاده شد.

تشخیص جامعه با الگوریتم کی‌میانگین

با استفاده از الگوریتم کی‌میانگین گراف را خوشه‌بندی می‌نماییم. قبل از خوشه‌بندی با این الگوریتم تعداد خوشه بهینه را با استفاده از معیار Adjusted Rand Index (ARI) به دست آورده‌ایم. مطابق با شکل (۴) با استفاده از معیار ارزیابی Adjusted Rand Index (ARI) تعداد بهینه خوشه‌ها ۳ به‌دست می‌آید.

شکل ۴. دقت معیار ارزیابی Adjusted Rand Index الگوریتم خوشه بندی کی‌میانگین

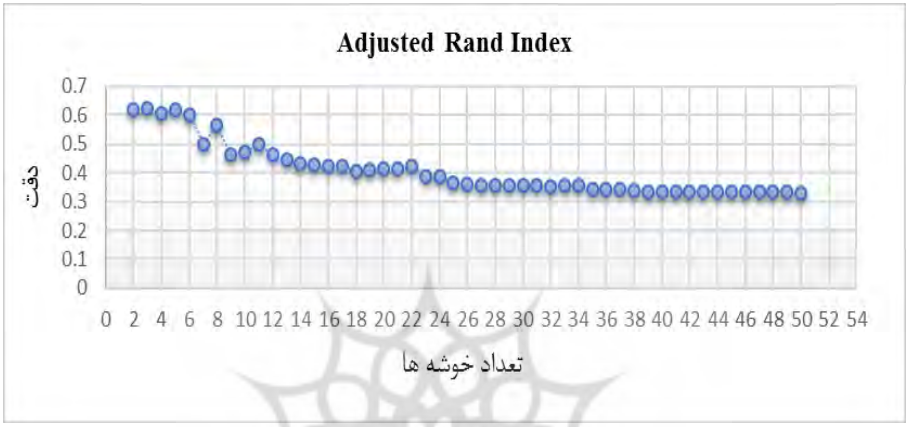


تشخیص جامعه با الگوریتم سلسله مراتبی

با استفاده از الگوریتم سلسله مراتبی گراف را خوشه‌بندی می‌نماییم. قبل از خوشه‌بندی با این الگوریتم تعداد خوشه بهینه را با استفاده از معیار Adjusted Rand Index (ARI) به

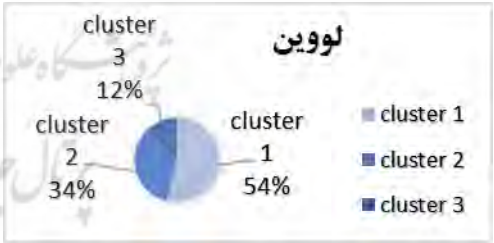
دست آورده‌ایم. مطابق با شکل (۵) با استفاده از معیار ارزیابی Adjusted Rand Index (ARI) تعداد بهینه خوشه‌ها ۳ به‌دست می‌آید.

شکل ۵. دقت معیار ارزیابی Adjusted Rand Index الگوریتم خوشه بندی سلسله‌مراتبی

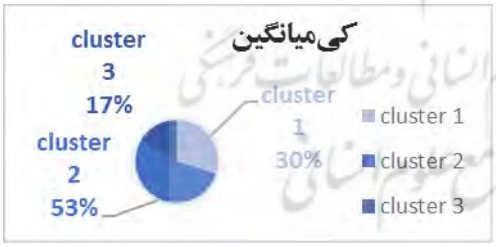


تشخیص جامعه با الگوریتم لووین

با استفاده از الگوریتم لووین گراف را خوشه‌بندی می‌نماییم. این الگوریتم احتیاجی به دانستن تعداد خوشه‌ها قبل از خوشه‌بندی ندارد. توزیع کاربران مطابق با هر سه الگوریتم در شکل‌های زیر قابل مشاهده است.



شکل ۶. توزیع کاربران در خوشه‌ها طبق الگوریتم خوشه‌بندی لووین



شکل ۷. توزیع کاربران در خوشه‌ها طبق الگوریتم خوشه‌بندی کی‌میانگین

شکل ۸. توزیع کاربران در خوشه‌ها طبق الگوریتم خوشه‌بندی سلسله‌مراتبی

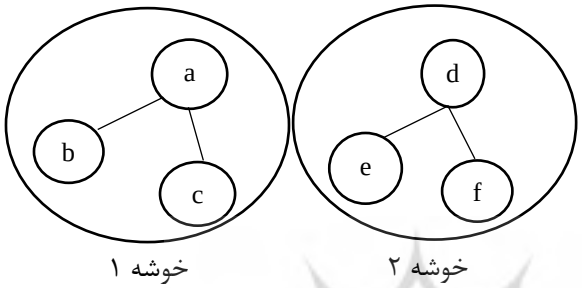
مقایسه نتایج حاصل از الگوریتم‌های لووین، کی‌میانگین و سلسله‌مراتبی

در این تحقیق با استفاده از ۱۰ نوع رابطه بین کاربران و بدون در نظر گرفتن روابط دوستی بین کاربران، گراف شبکه ایجاد شده و با استفاده از ۳ الگوریتم لووین، کی‌میانگین و سلسله‌مراتبی خوشه‌بندی گراف انجام شده است. حال برای مقایسه نتایج هر یک از روش‌های خوشه‌بندی، از اطلاعات روابط دوستی بین کاربران استفاده می‌شود. به این صورت که پس از خوشه‌بندی گراف با هر یک از روش‌های مذکور اگر دو کاربری که دارای رابطه دوستی با یکدیگر هستند، در یک خوشه نیز واقع شده بودند یک امتیاز مثبت در نظر گرفته می‌شود. به همین ترتیب برای هر یک از روابط دوستی بین دو کاربر خوشه‌ها بررسی می‌شوند و در صورت وجود آن دو کاربر در یک خوشه، یک امتیاز مثبت در نظر گرفته شده و درصد مطابقت داده‌های خوشه‌بندی شده با روابط دوستی به دست می‌آید. خوشه‌بندی مناسب است که درصد بیشتری از روابط دوستی را در خوشه‌های خود جای دهد.

مطابق با شکل (۹) به عنوان مثال اگر گراف به این صورت خوشه‌بندی شده باشد، به دلیل اینکه گره‌های a و b در یک خوشه واقع شده‌اند و همچنین با یکدیگر ارتباط دوستی نیز دارند بنابراین یک امتیاز مثبت در نظر گرفته می‌شود. به همین ترتیب برای سایر روابط دوستی نیز خوشه‌ها بررسی شده و درصد مطابقت وجود روابط دوستی در خوشه‌ها به دست می‌آید. مطابق جدول (۲) از نتایج حاصل از خوشه‌بندی متوجه می‌شویم که روابط

خوشه‌های به دست آمده از الگوریتم خوشه‌بندی لووین دارای درصد مطابقت بالاتری با روابط دوستی هستند.

شکل ۹. مثال خوشه‌بندی گره‌های گراف و وجود یا عدم وجود روابط دوستی در خوشه‌ها



روابط دوستی		وجود در یک خوشه
b	a	✓
d	c	×

جدول۲. مقایسه نتایج حاصل از الگوریتم‌های لووین، کی‌میانگین و سلسله‌مراتبی

	درصد مطابقت	ماژولاریتی	تعداد خوشه	توزیع داده‌ها در خوشه اول	توزیع داده‌ها در خوشه دوم	توزیع داده‌ها در خوشه سوم
لووین	۰/۷۶	۰/۳۷	۳	٪ ۵۴	٪ ۳۴	**٪ ۱۲**
کی‌میانگین	۰/۷۲	۰/۳۴	۳	٪ ۳۰	٪ ۵۳	**٪ ۱۷**
سلسله‌مراتبی	۰/۶۸	۰/۳۱	۳	٪ ۱۹	٪ ۵۳	**٪ ۲۸**

اجرای الگوریتم ژنتیک

در مرحله قبل برای هر یک از روابط وزنی برابر با ۱ درنظر گرفته شد و گراف شبکه ایجاد و خوشه‌بندی نیز انجام گردید. به‌وسیله بررسی نتایج متوجه شدیم که روابط خوشه‌های به دست آمده از الگوریتم خوشه‌بندی لووین دارای درصد مطابقت بالاتری با روابط دوستی هستند. در این مرحله با استفاده از الگوریتم ژنتیک برای هر یک از ۱۰ رابطه وزن‌های مختلفی در نظر گرفته شده و گراف شبکه با وزن یال‌های مختلف ایجاد شده است. سپس الگوریتم خوشه‌بندی لووین با در نظر گرفتن وزن جدید یال‌ها، بر روی گراف شبکه اجرا

شده و درصد مطابقت روابط خوشه‌های به دست آمده با روابط دوستی حساب می‌شود. این چرخه ادامه می‌یابد تا زمانی که درصد مطابقت بالاتری به دست آید. شکل (۱۰) چرخه اجرای الگوریتم ژنتیک را نشان می‌دهد.

بعد از اینکه الگوریتم ژنتیک ۳۰۰ مرتبه اجرا گردید، دیگر افزایشی در میزان درصد مطابقت مشاهده نگردید. با اجرای الگوریتم ژنتیک و در نظر گرفتن وزن مختلف برای هر یک از ۱۰ رابطه و خوشه‌بندی داده‌ها با الگوریتم لووین بیشترین درصد مطابقت مطابق جدول (۳) و وزن بهینه هر یک از ۱۰ رابطه به شرح جدول(۴) می‌باشد.

شکل۱۰. چرخه اجرای الگوریتم ژنتیک و خوشه‌بندی داده‌ها و به دست آوردن بیشترین درصد مطابقت و وزن بهینه هر یک از ۱۰ رابطه

جدول۳. نمایش میزان درصد مطابقت قبل از اجرای الگوریتم ژنتیک با وزن برابر با ۱ برای همه روابط و بعد از اجرای الگوریتم ژنتیک و به دست آوردن وزن بهینه

۰/۸۷۹	**درصد مطابقت حاصل از اجرای الگوریتم خوشه بندی لووین بعد از اجرای الگوریتم ژنتیک و به دست آوردن وزن بهینه روابط**
۰/۷۶۴	درصد مطابقت حاصل از اجرای الگوریتم خوشه بندی لووین با در نظر گرفتن وزن برابر با ۱ برای همه روابط
٪ ۱۵/۰۵	میزان بهبود درصد مطابقت

جدول۴. نمایش وزن بهینه روابط

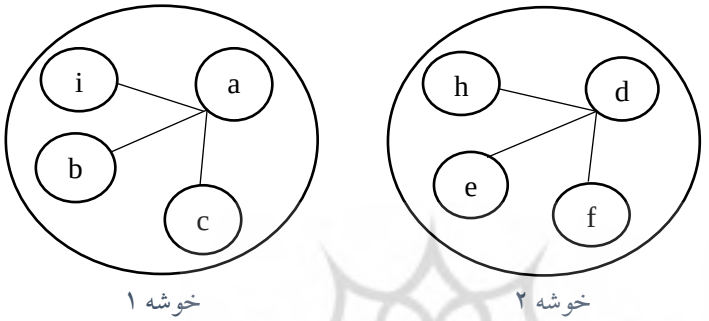
روابط	هم‌شهری	هم کلاسی	هم‌مدرسه‌ای	هم استانی	هم گروهی	پسند پست مشترک	هم‌درسی	هم استادی	هم‌سطحی	هم‌بخشی
وزن بهینه	۱۰	۱۰	۹/۵	۹	۸	۶/۷۵	۶/۵	۵/۲۵	۴/۵	/۷۵ ۳

پس از اجرای الگوریتم ژنتیک و خوشه بندی گراف متوجه شدیم که برای هر رابطه چه وزنی در نظر گرفته شده تا درصد مطابقت الگوریتم خوشه‌بندی به بیشترین مقدار خود برسد و کاربرانی که در یک خوشه قرار می گیرند بیشترین مطابقت را با کاربرانی که رابطه دوستی دارند داشته باشند. در این حالت خوشه‌های حاصل از خوشه‌بندی گراف، خوشه‌های بهینه، منسجم و با همبستگی بالا می‌باشند که حاوی کاربران با بیشترین شباهت به یکدیگر هستند. بنابراین می‌توان نتیجه گرفت سایر کاربرانی نیز که در یک خوشه قرار گرفته‌اند و با یکدیگر رابطه دوستی ندارند، می‌توان به عنوان دوست به یکدیگر پیشنهاد داد.

مطابق با شکل زیر به عنوان مثال گراف شبکه با استفاده از ۱۰ رابطه ایجاد شده و پس از اجرای الگوریتم خوشه‌بندی، ۲ خوشه زیر حاصل شده است. در خوشه ۱ گره a با ۳ گره b، c و i در ارتباط است. در روابط دوستی گره a فقط با گره b رابطه دوستی دارد. بنابراین چون گره‌های c و i با گره a در یک خوشه قرار گرفته‌اند و گره a با این دو گره

رابطه دوستی ندارد، می‌توان برای پیشنهاد دوست به گره a، گره‌های c و i را پیشنهاد داد. همچنین در خوشه ۲ نیز برای گره d می‌توان گره‌های h و f را به عنوان دوست پیشنهاد داد.

شکل ۱۱. نمایش مثالی از پیشنهاد دوست از بین کاربرانی که در یک خوشه قرار گرفته‌اند و با یکدیگر رابطه دوستی ندارند



روابط دوستی	
a	b
d	e

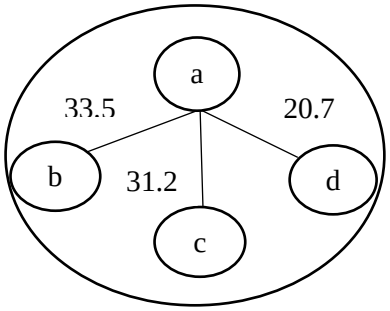
اولویت‌بندی پیشنهاد دوست به کاربران با درنظر گرفتن وزن روابط

پس از خوشه‌بندی گراف، ممکن است تعداد کاربرانی که در یک خوشه قرار می‌گیرند بسیار زیاد بوده و به دلیل این‌که کاربرانی که هم‌خوشه هستند به عنوان دوست به یکدیگر پیشنهاد داده می‌شوند، تعداد کاربران پیشنهادی بسیار زیاد شوند. به همین دلیل استفاده از یک روش اولویت‌بندی بین کاربران پیشنهادی ضروری می‌باشد.

روش اولویت‌بندی که در این تحقیق مورد بررسی قرار گرفته است استفاده از وزن روابط یا همان وزن یال‌های گراف می‌باشد. به این صورت که برای یک گره وزن تمام یال‌های متصل به گره بررسی می‌شود و به تعداد افراد پیشنهادی مورد نظر یال‌هایی با بیشترین وزن انتخاب می‌شوند و گره‌هایی که در انتهای یال‌های منتخب قرار گرفته‌اند به ترتیب به عنوان دوست به گره مورد نظر پیشنهاد داده می‌شوند.

به عنوان مثال اگر شکل زیر یکی از خوشه‌های حاصل از خوشه‌بندی باشد و گره a با سه گره b، c و d در رابطه باشد و روابط بین این گره‌ها و وزن بهینه آن‌ها نیز به شرح شکل زیر باشد. برای پیشنهاد دوست به گره a وزن هر یک از یال‌های متصل به گره a بررسی شده و به ترتیب گره b با اولویت اول، گره c با اولویت دوم و گره d با اولویت سوم پیشنهاد داده می‌شوند.

شکل۱۲. مثالی از یک خوشه گراف



جدول ۵. اولویت بندی پیشنهادات مطابق با وزن یال‌ها

		۱۰ نوع رابطه بین کاربران و وزن بهینه هر یک از روابط										
اولویت پیشنهاد	وزن یال	هم‌بخشی (۳/۷۵)	هم‌سطحی (۴/۵)	هم‌استادی (۵/۲۵)	هم‌درسی (۶/۵)	پسند پست مشترک (۶/۷۵)	هم‌گروهی (۸)	هم‌استانی (۹)	هم‌مدرسه‌ای (۹/۵)	هم‌کلاسی (۱۰)	هم‌شهری (۱۰)	روابط گراف
۱	۳۳/۵				۲							a—b
۲	/۲۵ ۳۱			۲	۲				۲			a—c
۳	/۷۵ ۲۰			۲	۲			۲				a—d

ارزیابی روش پیشنهاد دوست

برای ارزیابی روش پیشنهاد دوست از طریق اولویت‌بندی وزن یال‌های گراف، دوستان پیشنهادی با دوستان واقعی که در روابط دوستی موجود هستند مقایسه شده و درصد

مطابقت دوستان پیشنهادی با دوستان واقعی به دست می‌آید. درصد مطابقت به ترتیب برای ۱۰، ۲۰ و ۳۰ نفر دوست پیشنهادی در جدول (۵) مشخص است.

جدول ۵. درصد مطابقت دوستان پیشنهادی با دوستان واقعی

تعداد دوستان پیشنهادی	درصد مطابقت دوستان پیشنهادی با دوستان واقعی
۱۰	٪ **۶۹**
۲۰	٪ **۷۰**
۳۰	٪ **۷۲**

نتیجه‌گیری و ارائه پیشنهادها

هدف اصلی سیستم‌های پیشنهاددهنده دوست در شبکه‌های اجتماعی، کمک به کاربر برای برقراری رابطه جدید می‌باشد. به عبارت دیگر هدف سیستم‌های پیشنهاددهنده‌ی دوست، پیشنهاد ایجاد یال‌های جدید بین کاربران شبکه است. امروزه استفاده از انواع سیستم‌های پیشنهاددهنده در شبکه‌های اجتماعی رونق بسیاری پیدا کرده است. در این پژوهش از اطلاعات و داده‌های شبکه علمی و اجتماعی کوثرنت استفاده شده است. شبکه علمی اجتماعی کوثرنت یک شبکه سازمانی برای طلاب خانم سراسر کشور می‌باشد. با استفاده از ۱۰ نوع رابطه بین کاربران و بدون در نظر گرفتن روابط دوستی بین کاربران گراف شبکه ایجاد گردید و از داده‌های مربوط به روابط دوستی بین کاربران در قسمت ارزیابی و محاسبه درصد مطابقت استفاده شد. سپس با استفاده از ۳ الگوریتم لووین، کی‌میانگین و سلسله‌مراتبی خوشه‌بندی گراف جهت تشخیص جوامع با میزان شباهت زیاد و میزان اختلاف کم انجام گردید. برای مقایسه نتایج هر یک از روش‌های خوشه‌بندی، از اطلاعات روابط دوستی بین کاربران استفاده گردید. نتایج حاصل از ۳ روش خوشه‌بندی نشان داد که روابط خوشه‌های به دست آمده از الگوریتم خوشه‌بندی لووین دارای درصد مطابقت بالاتری با روابط دوستی هستند. در مرحله بعد با استفاده از الگوریتم ژنتیک برای هر یک از ۱۰ رابطه وزن‌های مختلفی در نظر گرفته شد و گراف شبکه با وزن یال‌های مختلف

ایجاد گردید. سپس الگوریتم خوشه‌بندی لووین با در نظر گرفتن وزن جدید یال‌ها، بر روی گراف شبکه اجرا شد و درصد مطابقت روابط خوشه‌های به دست آمده با روابط دوستی محاسبه گردید. مراحل الگوریتم ژنتیک تا زمانی که دیگر بهبودی در میزان درصد مطابقت مشاهده نگردید ادامه یافت. سپس بیشترین درصد مطابقت به همراه وزن بهینه هر یک از ۱۰ رابطه به دست آمد و درصد مطابقت روابط موجود در خوشه‌ها با روابط دوستی نسبت به حالت قبلی یعنی بدون در نظر گرفتن وزن‌های مختلف بهبود یافت. در این حالت خوشه‌های حاصل از خوشه‌بندی گراف، خوشه‌های بهینه، منسجم و با همبستگی بالا می‌باشند که حاوی کاربران با بیشترین شباهت به یکدیگر هستند. بنابراین می‌توان نتیجه گرفت سایر کاربرانی نیز که در یک خوشه قرار گرفته‌اند و با یکدیگر رابطه دوستی ندارند، می‌توان به عنوان دوست به یکدیگر پیشنهاد داد. برای اولویت‌بندی پیشنهادات نیز وزن یال‌های بین افراد در گراف در نظر گرفته شد به این صورت که برای یک گره وزن تمام یال‌های متصل به گره بررسی گردید و به تعداد افراد پیشنهادی مورد نظر یال‌هایی با بیشترین وزن انتخاب شدند و گره‌هایی که در انتهای یال‌های منتخب قرار گرفتند به ترتیب به عنوان دوست به گره مورد نظر پیشنهاد داده شدند. در پایان نیز روش پیشنهاد دوست مورد ارزیابی قرار گرفت و درصد مطابقت دوستان پیشنهادی با دوستان واقعی فرد محاسبه گردید.

ارائه پیشنهادها

در این تحقیق برای تعیین درصد مطابقت روابط خوشه‌های به دست آمده از روابط دوستی استفاده گردید و وزن بهینه هر یک از روابط براساس آن تعریف گردید. حال اگر هر رابطه دیگری مانند رابطه همکار علمی بودن و کار کردن بر روی مقالات مشترک به جای روابط دوستی قرار گیرد آن‌گاه می‌توان وزن بهینه هر یک از روابط را برای آن رابطه به دست آورد و میزان تاثیرگذاری هر یک از ۱۰ رابطه را بر روی رابطه در نظر گرفته شده متوجه شد. همچنین می‌توان به جای استفاده از ۱۰ رابطه مورد استفاده در این پروژه از روابط دیگر نیز استفاده نمود و تاثیر آن‌ها را نیز بررسی نمود. همچنین به جای اولویت بندی افراد پیشنهادی بر اساس وزن یال می‌توان از معیارهای دیگر مانند معیارهای مرکزیت در گراف و یا اشتراک معیارهای مرکزیت با وزن یال‌ها استفاده نمود.

ORCID

Zahra Shirani http://orcid.org/0000-0001-8972-4408

Amir Jalaly Bidgoly http://orcid.org/0000-0002-8574-3537

منابع

جلالی، امید. (۱۳۹۳). سیستم پیشنهاددهنده هم‌پژوهشی (مورد مطالعه: شبکه اجتماعی کوثرنت). پایان نامه کارشناسی ارشد رشته مهندسی فناوری اطلاعات گرایش تجارت الکترونیک، دانشکده فنی و مهندسی، دانشگاه قم.

Amigó, E., Gonzalo, J., Artiles, J., & Verdejo, F. (2009). A comparison of extrinsic clustering evaluation metrics based on formal constraints. Information retrieval, 12(4), 461-486.

Bao, J., Zheng, Y., Wilkie, D., & Mokbel, M. (2015). Recommendations in location-based social networks: a survey. GeoInformatica, 19(3), 525-565.

Bradnova, V., Chernyavsky, M. M., Just, L., Haiduc, M., Kharlamov, S. P., Kovalenko, A. D., & Zarubin, P. L. (2003). Nuclear Clustering Quest in Relativistic Multifragmentation. Paper presented at the Few-Body Problems in Physics '02, Vienna: Springer Vienna, Vol. 14, 241-244.

Chen, J., Geyer, W., Dugan, C., Muller, M., & Guy, I. (2009). Make new friends, but keep the old: recommending people on social networking sites. Paper presented at the Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Boston, MA, USA, 201-210.

Chu, C.-H., Wu, W.-C., Wang, C.-C., Chen, T.-S., & Chen, J.-J. (2013). Friend Recommendation for Location-Based Mobile Social Networks. Paper presented at the 2013 Seventh International Conference on Innovative Mobile and Internet Services in Ubiquitous Computing, 365-370.

Ding, D., Zhang, M., Li, S.-Y., Tang, J., Chen, X., & Zhou, Z.-H. (2017). BayDNN : Friend Recommendation with Bayesian Personalized Ranking Deep Neural Network. Paper presented at the Proceedings of the 2017 ACM on Conference on Information and Knowledge Management - CIKM '17, 1479-1488.

Du, Z., Hu, L., Fu, X., & Liu, Y. (2014). Scalable and Explainable Friend Recommendation in Campus Social Network System. Paper presented at the Frontier and Future Development of Information Technology in Medicine and Education, Dordrecht: Springer Netherlands, Vol. 269, 457-466.

Filippone, M., Camastra, F., Masulli, F., & Rovetta, S. (2008). A survey of kernel and spectral methods for clustering. Pattern Recognition, 41(1), 176-190.

Hamid, M. N., Naser, M. A., Hasan, M. K., & Mahmud, H. (2014). A cohesion-based friend-recommendation system. Social Network Analysis and Mining, 4(1), 176-186.

Kherad, M., & Bidgoly, A. J. (2020). Recommendation system using a deep learning and graph analysis approach. arXiv preprint arXiv:2004.08100.

Li, S., Song, X., Lu, H., Zeng, L., Shi, M., & Liu, F. (2020). Friend recommendation for cross marketing in online brand community based on intelligent attention allocation link prediction algorithm. Expert Systems with Applications, 139, 112839.

Liben-Nowell, D., & Kleinberg, J. (2007). The link-prediction problem for social networks. Journal of the American Society for Information Science and Technology, 58(7), 1019-1031.

Likas, A., Vlassis, N., & J. Verbeek, J. (2003). The global k-means clustering algorithm. Pattern Recognition, 36(2), 451-461.

Sammer A, Q., & Ayad R, A. (2019). Survey of User to User Recommendation System in Online Social Networks. Engineering and Technology Journal, 37(10), 422-428.

Sander, J., Ester, M., Kriegel, H.-P., & Xu, X. (1998). Density-Based Clustering in Spatial Databases: The Algorithm GDBSCAN and Its Applications. Data Mining and Knowledge Discovery, 2(2), 169-194.

Suman Venkata, S., Yuvraj Singh, C., Swaraj, K.,& Tripathy, B. (2020). Recommendation System Using Community Identification. International Conference on Innovative Computing and Communications, Vol. 1087, 125-132.

Swarnakar, P., Kumar, A., & Tyagi, H. (2017). Network dynamics in friend recommendation: a study of Indian engineering students. Int. J. Information Technology and Management, 16(3), 287-300.

Wan, S., Lan, Y ,.Guo, J., Fan, C., & Cheng, X. (2013). Informational friend recommendation in social media. Paper presented at the Proceedings of the 36th international ACM SIGIR conference on Research and development in information retrieval, Dublin, Ireland, 1045-1048.

Wang, Z ,.Liao, J., Cao, Q., Qi, H., & Wang, Z. (2015). Friendbook: A Semantic-Based Friend Recommendation System for Social Networks. IEEE Transactions on Mobile Computing, 14(3), 538-551.

Weng, L., & Zhang, Q. (2020). A social recommendation method based on opinion leaders. Multimedia Tools and Applications, 1-16.

Xu, Y., Zhou, D., & Ma, J. (2019). Scholar-friend recommendation in online academic communities: An approach based on heterogeneous network. Decision Support Systems, 119, 1-13.

Yu, Z., Wang, C., Bu, J., Wang, X., Wu, Y & ,Chen, C. (2015). Friend recommendation with content spread enhancement in social networks. Information Sciences, 309, 102-118.
Zhao, X., Ma, Z., & Zhang, Z. (2017). A novel recommendation system in location-based social networks using distributed ELM. Memetic Computing, 10(3), 321-331.
Zheng, H., & Wu, J. (2017). Friend Recommendation in Online Social Networks: Perspective of Social Influence Maximization. Paper presented at the in Computer Communication and Networks (ICCCN), 2017 26th International Conference on IEEE, 1-9.

پژوهشگاه علوم انسانی و مطالعات فرهنگی
پرتال جامع علوم انسانی

استناد به این مقاله: شیرانی، زهرا. (۱۴۰۰). هم‌افزایی کاربران در شبکه علمی اجتماعی کوثرنت با استفاده از روش‌های خوشه‌بندی مبتنی بر گراف. *مطالعات مدیریت کسب‌وکار هوشمند*، سال ۱۴۰۰(شماره۳۵)، ص ۱۸۷- ص ۲۱۶.

Journal of Business Intelligence Management Studies is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.