

عدد اردوش و شبکه تألیف‌های مشترک: نگاهی به کاربردی از نظریه گراف در علم‌سنجی

هادی صفری

چکیده. بیش از نیم‌قرن است که ریاضیدانان فاصله همکاری خود با پاول اردوش، ریاضیدان مجارستانی، در نوشتن مقالات علمی مشترک را عدد اردوش می‌نامند. عدد اردوش مثالی از فاصله گره‌ها در شبکه تألیف‌های مشترک ریاضیدانان است؛ یکی از شبکه‌هایی که در مطالعات علم‌سنجی برای بررسی همکاری‌های علمی میان پژوهشگران و رفتار آن‌ها به کار می‌رود. در این نوشته پس از نگاهی به تاریخچه عدد اردوش و جایگاه آن در بین ریاضیدانان، شبکه تألیف‌های مشترک را به صورت مثالی از کاربردهای نظریه گراف در حوزه علم‌سنجی بررسی می‌کنیم.

۱. مقدمه

پاول اردوش^۱ (۱۹۱۳-۱۹۹۶) ریاضیدان مشهور مجارستانی در طول عمر ۸۳ ساله خود مقالات بسیاری نوشت و با ریاضیدانان فراوانی همکاری کرد. گفته می‌شود تنها چهار ریاضیدان در طول تاریخ ریاضیات بیش از هفتصد مقاله به رشته تحریر درآورده‌اند و اردوش، با بیش از ۱۴۱۶ مقاله، در صدر این فهرست جای دارد [۸]. این تعداد مقاله فقط مقالاتی را در برمی‌گیرد که در بانک اطلاعاتی متمتیکال ریویوز^۲ نمایه شده‌اند. او در حوزه‌های مختلفی از ریاضیات و علوم وابسته پژوهش کرده است [۳]؛ هرچند که بیشترین مقالات او در حوزه‌های نظریه اعداد و ترکیبیات است [۶]. اردوش حداقل با ۵۰۹ محقق در نگارش مقالاتش همکاری کرده است و گفته می‌شود بیشترین تعداد همکاران را در بین ریاضیدانان چندصدساله اخیر داشته است [۸]. یکی از دلایل این موضوع را جابه‌جایی‌ها و مسافرت‌های متعدد اردوش می‌دانند [۶].

از اواسط قرن گذشته میلادی، ریاضیدانان - بیشتر به صورت نوعی سرگرمی - فاصله همکاری^۳ خود را با پاول اردوش اندازه گرفته‌اند و آن را عدد اردوش^۴ نامیده‌اند. عدد اردوش بر اساس شبکه تألیف‌های مشترک^۵ نویسندگان و معمولاً بر اساس اطلاعات بانک اطلاعاتی متمتیکال ریویوز محاسبه می‌شود. این شبکه‌ها از اوایل قرن اخیر میلادی مورد توجه پژوهشگران قرار گرفته‌اند و با روش‌های الهام گرفته‌شده از نظریه گراف و ابزارهای رایج تحلیل

عبارات و کلمات کلیدی: عدد اردوش، شبکه تألیف‌های مشترک، علم‌سنجی، گراف، ریاضیات کاربردی
دبیرتخصصی رابط: علیرضا عبدالمهی
تاریخ دریافت: ۱۴۰۰/۰۲/۰۴ تاریخ پذیرش: ۱۴۰۰/۰۲/۲۵
نوع مقاله: ترویجی

<http://dx.doi.org/10.22108/MSCI.2021.128339.1427>

¹Paul Erdős ²Math Reviews database ³collaborative distance ⁴Erdős number ⁵co-authorship network

شبکه‌های اجتماعی تحلیل شده‌اند. شبکه‌های تألیف‌های مشترک را می‌توان یکی از نمودهای همکاری‌های علمی دانست. این موضوع شبکه‌های مذکور را به ابزار مناسبی برای پژوهشگران حوزه علم‌سنجی^۶ تبدیل کرده است و از طریق آن می‌توانند رفتار پژوهشگران و ارتباطات آن‌ها را بررسی کنند. تحلیل شبکه‌های تألیف‌های مشترک، در کنار تحلیل شبکه‌های استنادی و انواع دیگری از شبکه‌ها، نمونه‌هایی از کاربردهای عملی نظریه گراف در حوزه‌های دیگر علوم به شمار می‌رود.

۲. عدد اردوش

اولین اشارات مکتوب به عدد اردوش احتمالاً به یادداشت گافمن با عنوان «خُب عدد اردوش شما چند است؟» [۵] در سال ۱۹۶۹ بازمی‌گردد. او تعریف ریاضی زیر را برای عدد اردوش ارائه کرد.

تعریف ۱.۲ ([۵]). فرض کنید A و B دو ریاضیدان باشند. همچنین فرض کنید دنباله A_i ها (برای $i \in \{0, 1, \dots, n\}$) زنجیره‌ای از ریاضیدان‌ها باشد که $A_0 = A$ و $A_n = B$ و به ازای هر $i \in \{0, 1, \dots, n-1\}$ ریاضیدان A_i حداقل یک مقاله مشترک با ریاضیدان A_{i+1} داشته باشد. در این صورت $\langle A_0, A_1, \dots, A_n \rangle$ یک زنجیره به طول n نامیده می‌شود که A را به B متصل می‌کند. عدد A ی ریاضیدان B که آن را با $v(A; B)$ نشان می‌دهیم برابر طول کوتاه‌ترین زنجیره‌ای است که A را به B متصل می‌کند. تعریف می‌کنیم $v(A; A) = 0$ و نیز اگر هیچ زنجیره‌ای A را به B متصل نکند، $v(A; B) = +\infty$.

اکنون دو قضیه بدیهی:

قضیه ۲.۲. برای هر دو ریاضیدان A و B ، $v(A; B) = v(B; A)$.

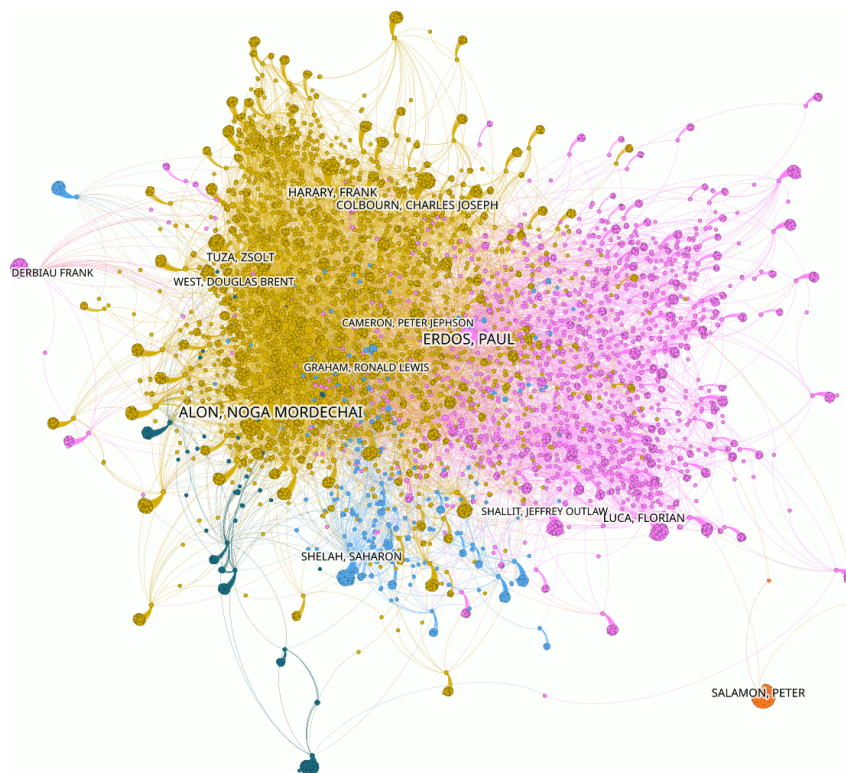
قضیه ۳.۲. برای هر سه ریاضیدان A ، B ، و C ، $v(A; B) + v(B; C) \geq v(A; C)$.

عدد اردوش در واقع حالت خاص تابع $v(\text{Erdős}; \cdot)$ با دامنه همه ریاضیدانان است. گافمن خاطرنشان می‌کند یک‌سال پیش از انتشار یادداشت مذکور اردوش را در لندن دیده است و او از اینکه ریاضیدانان عدد اردوش خود را محاسبه می‌کنند بی‌خبر بوده است [۵].

گراسمان در [۹، ۶] تحقیقات بیشتری درباره عدد اردوش انجام داده است و نیز پروژه عدد اردوش^۷ [۷] را در دانشگاه اوکلند نیوزلند تأسیس کرده است. او به همراه دکاسترو در [۳] عدد اردوش چندین دانشمند مشهور و برندگان برخی جوایز علمی بین‌المللی را محاسبه کرده است. در سال‌های اخیر تحقیقات بیشتری درباره عدد اردوش صورت گرفته است. در برخی پژوهش‌های جدیدتر تعریف گسترش‌یافته‌ای از عدد اردوش با عنوان عدد اردوش اعشاری ارائه شده است [۱۲].

شکل ۱ شبکه همکاری‌های ریاضیدانانی را نشان می‌دهد که عدد اردوش آن‌ها حداکثر ۲ است. همان‌طور که مشاهده می‌شود اردوش مستقیماً یا با یک واسطه با گروه‌های متفاوتی از دانشمندان همکاری کرده است. پژوهشگران حوزه گراف و علوم کامپیوتر (طلایی)، متخصصین نظریه اعداد (صورتی)، ریاضیدانان حوزه نظریه مجموعه‌ها (آبی)، محققین حوزه مهندسی کامپیوتر (سبز)، و گروه جداافتاده پیترو سلیمان و همکارانش گروه‌های عمده این پژوهشگران را تشکیل می‌دهند. نام پژوهشگرانی با بیش از ۱۹۰ همکار در شبکه (۱۰٪ اول پژوهشگران) نوشته شده است.

⁶scientometrics ⁷The Erdős Number Project



شکل ۱: شبکه همکاری‌های ریاضیدانانی با عدد اردوش حداکثر ۲ بر اساس داده‌های [۷]

۳. شبکه‌های تألیف‌های مشترک

فاصله با اردوش در زنجیره تألیف‌های مشترک یک حالت خاص از مفهومی است که در علم‌سنجی تحت عنوان شبکه‌های تألیف‌های مشترک مطالعه می‌شود. تألیف‌های مشترک را یکی از نمودهای ارتباطات علمی می‌دانند [۱۰] که خود یکی از محورهای نهاد اجتماعی علم و پیشرفت علمی محسوب می‌شود [۱۵].

بررسی شبکه‌های استنادات علمی سال‌هاست که در حوزه‌های علم‌سنجی و کتاب‌سنجی و پیش از آن‌ها در کتابداری مورد توجه بوده است. اما با وجود علاقه برخی پژوهشگران به بررسی شبکه‌های تألیف‌های مشترک، استفاده عملی از این شبکه‌ها در تحقیقات علم‌سنجی تا سال‌های آغازین قرن حاضر میلادی و رایج شدن نمایه‌ها و کتابخانه‌های دیجیتال ممکن نشد [۱۳]. پس از این زمان تحقیقات فراوانی درباره تألیف‌های مشترک و ویژگی‌های شبکه‌های تألیف‌های مشترک صورت گرفته است.

تألیف‌های مشترک ابتدا در علوم اجتماعی باب شد و سپس به حوزه‌های علمی دیگر نیز تسری یافت [۱۶]. هرچند تمایل پژوهشگران حوزه‌های مختلف به تألیف‌های مشترک متفاوت است و در رشته‌های نظری مانند ریاضیات تمایل کمتری به تألیف‌های مشترک وجود دارد [۱۳]، از سال‌های دور نسبت تعداد مقالات با چند نویسنده به کل مقالات حوزه علوم ریاضی روندی صعودی داشته است [۹].

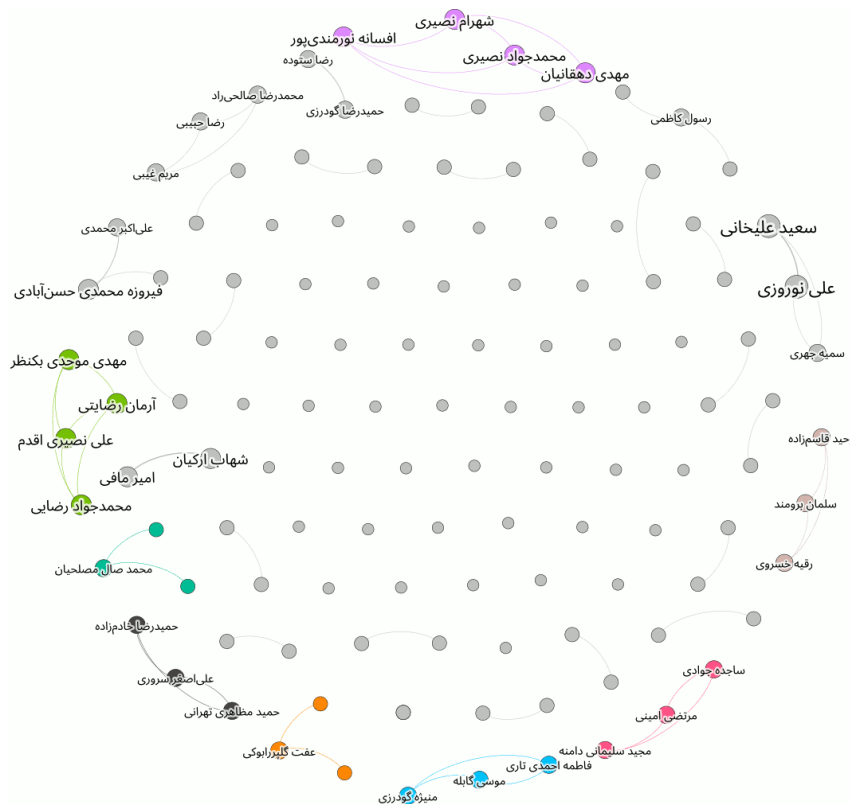
شبکه‌های تألیف‌های مشترک را می‌توان با گراف مدل‌سازی کرد. در چنین گرافی هر گره^۸ نشان‌دهنده یک نویسنده است و وجود یال^۹ بین دو گره به معنی آن است که نویسندگان متناظر آن‌ها حداقل در نگارش یک مقاله

^۸node ^۹edge

همکاری کرده‌اند [۱۳]. گاهی از گراف وزن دار برای مدل سازی استفاده می شود و وزن هر یال را برابر تعداد مقالات مشترک میان نویسندگان نظیر گره های دو سر یال در نظر می گیرند.

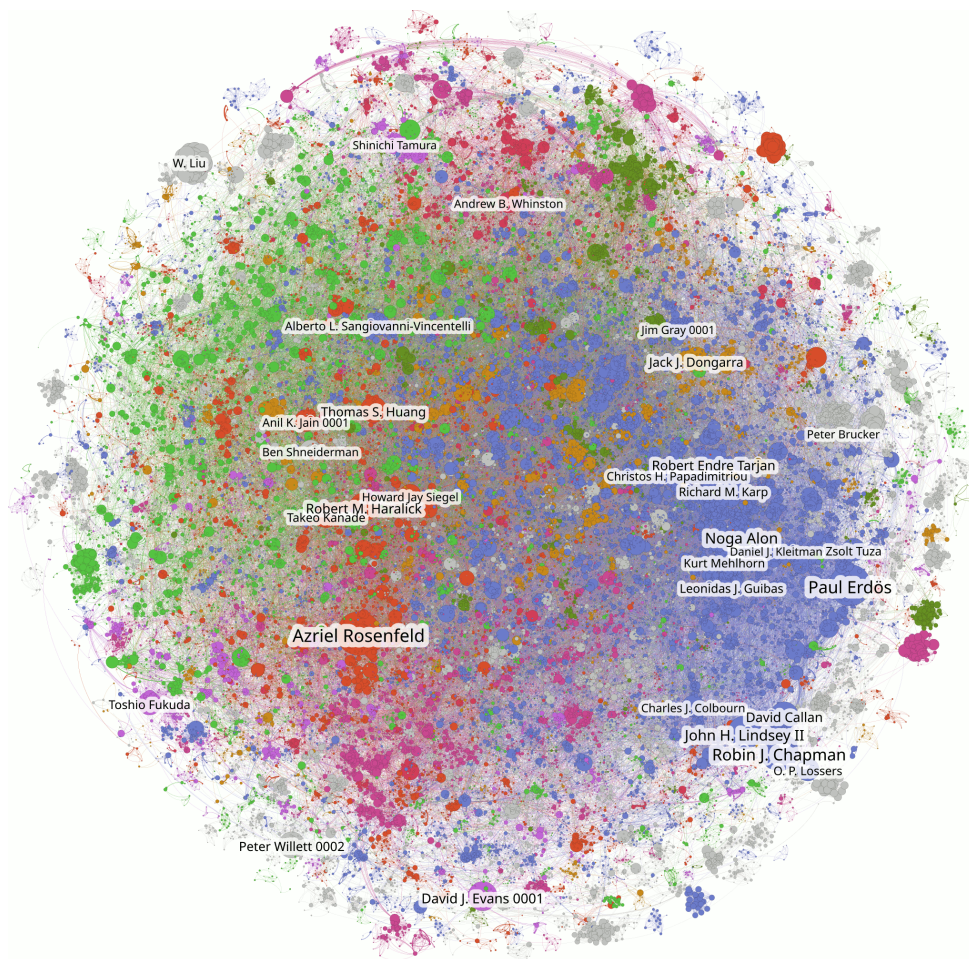
انواع گوناگونی از شبکه تألیف‌های مشترک در علم‌سنجی مورد بررسی قرار می‌گیرد. در برخی تحقیقات مانند [۱۷] شبکه تألیف‌های مشترک نویسندگان یک مجله خاص بررسی شده است. در تحقیقات دیگری به بررسی شبکه تألیف‌های مشترک پژوهشگران یک دانشگاه یا تألیف‌های مشترک بین‌المللی پژوهشگران کشوری خاص پرداخته شده است. به نظر می‌رسد با چنین تحقیقاتی نتوان دیدی جامع نسبت به وضعیت ارتباطات علمی و اهمیت پژوهشگران یک حوزه علمی به دست آورد و فقط بتوان به پرسش‌هایی محدود و مشخص پاسخ داد. در بسیاری از تحقیقات، تألیف‌های مشترک پژوهشگران یک حوزه علمی عام یا یک شاخه علمی خاص بررسی می‌شود و گاهی پژوهش به پژوهشگران یک کشور محدود می‌شود.

شکل ۲ شبکه همکاری‌های نویسندگان در ۱۷ شماره اول مجله ریاضی و جامعه را نشان می‌دهد. در این شکل نام نویسندگانی با درجه وزن دار ۲ یا بیشتر نوشته شده است. ۵۵٪ از مقالات منتشرشده در این نشریه یک نویسنده داشته‌اند. همان‌طور که مشخص است شبکه تألیف‌های مشترک این نشریه یک، دو و یا حداکثر سه همکاری میان نویسندگان را نشان می‌دهد و نمی‌توان با کمک آن اطلاعات جامعی درباره نویسندگان و همکاری‌های علمی آن‌ها به دست آورد.



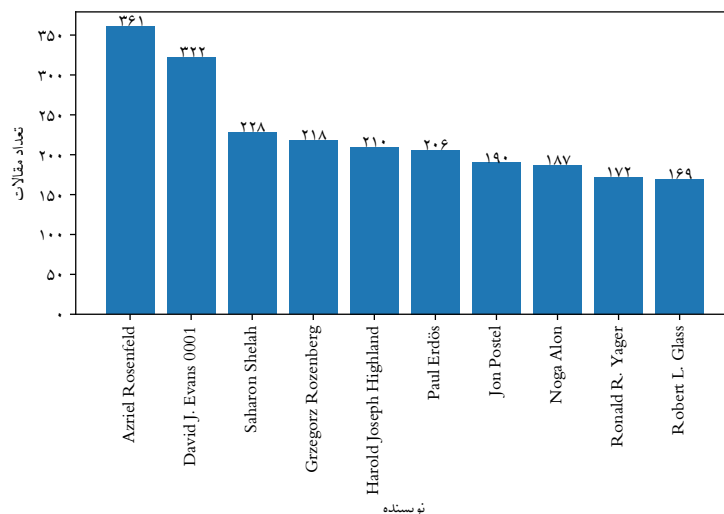
شکل ۲: شبکه همکاری‌های مجله ریاضی و جامعه از ابتدا تا پایان سال ۱۳۹۹

به عنوان یک نمونه دیگر، شکل ۳ شبکه همکاری‌های محققین علوم کامپیوتر را در نگارش مقالات منتشر شده در نشریات تا سال ۲۰۰۰ م. بر اساس اطلاعات بانک اطلاعاتی DBLP [۲] نشان می‌دهد. از اردوش بیش از ۲۰۰ مقاله با چنین مشخصاتی در این پایگاه داده یافت می‌شود؛ هرچند تا سال ۲۰۱۵ نیز (نزدیک به دو دهه پس از درگذشت او) مقالات جدیدی از اردوش منتشر شده است که به نظر می‌رسد ادامه فعالیت‌های مشترک او و همکارانش باشند. دانشمندانی که هیچ مقاله مشترکی با افراد دیگر نداشته‌اند و همچنین مقالاتی که نویسنده و یا سال انتشار یا عنوان آن‌ها نامشخص بوده است از تحلیل‌ها حذف شده‌اند.



شکل ۳: شبکه تألیف‌های مشترک محققین علوم کامپیوتر در مجلات منتشر شده تا سال ۲۰۰۰ بر اساس داده‌های [۲]

شبکه تألیف‌های مشترک دانشمندان علوم کامپیوتر ۲۴۳۷۵۵ گره (نویسنده) و ۴۶۸۲۷۵ یال (رابطه تألیف‌های مشترک) دارد که با بررسی ۳۴۰۶۹۹ مقاله به دست آمده‌اند. پس از حذف گره‌های منفرد، ۲۱۱۱۷۳ گره باقی می‌ماند. بیشترین مقالات متعلق به ایزریئل روزنفلد است (۳۶۱ مقاله)؛ پاول اردوش با ۲۰۶ مقاله در جایگاه ششم قرار دارد. شکل ۴ ده نویسنده برتر را بر اساس تعداد مقالاتشان نشان می‌دهد. در کل شبکه، هر نویسنده به طور میانگین ۲/۷۶ مقاله تألیف کرده است.



شکل ۴: نویسندگانی با بیشترین تعداد مقالات منتشرشده در مجلات تا سال ۲۰۰۰ بر اساس داده‌های [۲]

بیشتر مقالات مورد بررسی (۱۴۲۷۷۷ مقاله) یک نویسنده داشته‌اند و رتبه بعدی با ۱۱۶۹۰۲ مقاله به مقالاتی با دو نویسنده اختصاص داشته است؛ اما مقالات بسیاری نیز با بیش از ۲ نویسنده وجود دارد. میانگین تعداد نویسندگان مقالات برابر ۱/۹۷ بوده‌اند. **شکل ۵** روند افزایش سهم مقالاتی را که نتیجه تألیف‌های مشترک بوده‌اند نشان می‌دهد. از نیمه دوم سال‌های هشتاد میلادی بیش از نیمی از مقالات منتشرشده در هر سال نتیجه تألیف‌های مشترک بوده است. از سال ۱۹۹۴ بیشترین سهم مقالات منتشرشده در هر سال متعلق به مقالاتی با دو نویسنده بوده است.

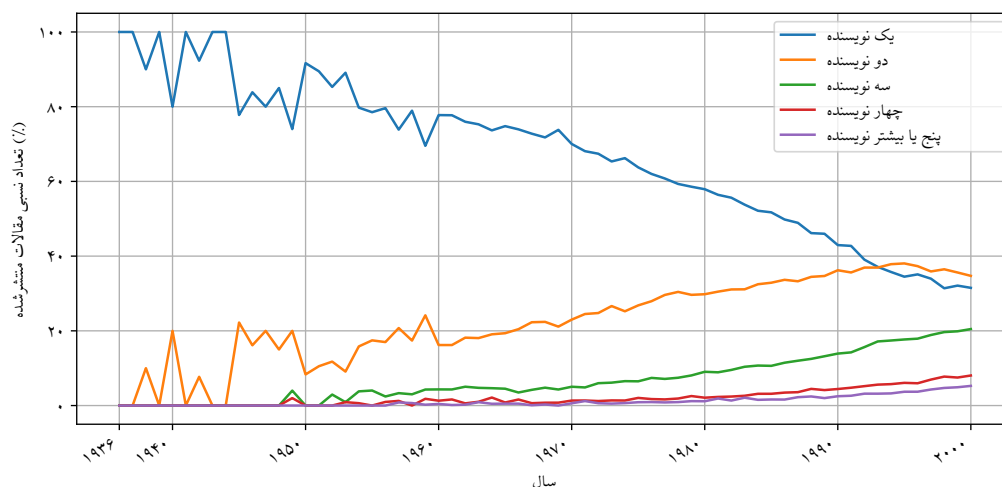
۴. نظریه گراف و شبکه‌های تألیف‌های مشترک

سال‌هاست که در ریاضیات کاربردی از نظریه گراف برای بررسی شبکه‌ها استفاده می‌شود. شبکه‌های تألیف‌های مشترک نیز از این قاعده مستثنی نیستند. الگوریتم‌های مختلفی برای کار با گراف‌ها و شبکه‌ها توسعه یافته است که محققین علم‌سنجی برای تحلیل شبکه‌های تألیف‌های مشترک از آن‌ها بهره می‌گیرند. در اثر حاضر برای چینه‌گره‌های شبکه در **شکل ۱** و **شکل ۳** از الگوریتم ForceAtlas2 [۱۱] و در **شکل ۲** از الگوریتم Fruchterman-Reingold [۴] استفاده شده است. همچنین، برای تشخیص اجتماعات در هر سه شکل از الگوریتم Louvain [۱] استفاده شده است.

بسیاری از سنجه‌های رایج در نظریه گراف نیز برای مقایسه شبکه‌های تألیف‌های مشترک به کار می‌روند. اندازه بزرگ‌ترین مؤلفه همبندی (مؤلفه غول‌آسا)^{۱۰}، قطر^{۱۱} و شعاع^{۱۲} گراف، ضریب خوشه‌بندی^{۱۳}، توزیع درجات، انواع سنجه‌های مرکزیت^{۱۴} و غیره برخی از این سنجه‌ها هستند. نیومن در [۱۳] بسیاری از این سنجه‌ها را برای شبکه تألیف‌های مشترک بانک اطلاعاتی متمتیکال ریویوز محاسبه کرده است.

تعریف برخی از این سنجه‌ها در ادامه بیان می‌شود. مجموعه گره‌های گراف G را با $V(G)$ و مجموعه یال‌های آن‌را با $E(G)$ نشان می‌دهیم.

¹⁰size of largest connected component (giant component) ¹¹diameter ¹²radius ¹³clustering coefficient ¹⁴centrality



شکل ۵: تعداد نویسندگان مقالات منتشرشده در مجلات تا سال ۲۰۰۰ در گذر زمان بر اساس داده‌های [۲]

تعریف ۱.۴ ([۱۴، ص. ۱۳۹]). اگر u و v دو گره گراف باشند، طول کوتاه‌ترین مسیر از u به v فاصله این دو گره نامیده می‌شود و با $d_{u,v}$ نمایش داده می‌شود.

به عبارت دیگر، فاصله u به v برابر حداقل تعداد یال‌هایی است که باید طی شود تا از u به v برسیم. اگر مسیری بین دو گره وجود نداشته باشد، فاصله میان آن‌ها برابر ∞ در نظر گرفته می‌شود.

تعریف ۲.۴ ([۱۴، ص. ۱۴۰]). قطر یک گراف برابر طولانی‌ترین فاصله متناهی میان گره‌های گراف است.

تعریف ۳.۴ ([۱۴، ص. ۱۹۹]). اگر گره u به گره v و گره v به گره w متصل باشد، مسیر uvw در گراف وجود خواهد داشت. اگر u مستقیماً به w نیز متصل باشد مسیر مذکور را یک مسیر بسته می‌نامیم. نسبت تعداد مسیرهای بسته به طول ۲ به تعداد کل مسیرهای به طول ۲ در یک گراف ضریب خوشه‌بندی نامیده می‌شود و با C نمایش داده می‌شود:

$$(۱) \quad C = \frac{\text{تعداد مسیرهای بسته به طول ۲}}{\text{تعداد مسیرها با طول ۲}}$$

ضریب خوشه‌بندی میانگین احتمال این رویداد است که دو همسایه یک گره خودشان مستقیماً بهم متصل باشند.

مطابق مشاهدات نیومن در [۱۳] در شبکه تألیف‌های مشترک بانک اطلاعاتی متمتیکال ریویوز تا سال ۲۰۰۴، ۲۵۳۳۳۹ گره (نویسنده) وجود داشته است. هر نویسنده به طور میانگین ۳/۹ همکار داشته است. بزرگ‌ترین مؤلفه همبندی ۸۲٪ گراف را در بر می‌گرفته است. میانگین فاصله گره‌ها برابر ۷/۶ و بیشترین فاصله دو گره ۲۷ بوده است. ضریب خوشه‌بندی نیز برابر ۰/۱۵ بوده است. این‌ها بخشی از مشاهداتی بوده‌اند که باعث شده‌اند نیومن تألیف‌های مشترک را در علوم محض و مجرد کمتر از علوم کاربردی و آزمایشگاهی بداند.

در شبکه تألیف‌های مشترک محققین علوم کامپیوتر در مجلات منتشرشده تا سال ۲۰۰۰، پس از حذف گره‌های تنها، هر نویسنده به طور میانگین ۵/۲۱ همکار و ۶/۹۳ همکاری داشته است. ۲۱۱۲۹ مؤلفه همبندی دیده شده است. بزرگ‌ترین مؤلفه همبندی شامل ۱۴۴۶۰۹ گره بوده است که برابر ۶۸/۴۸٪ گره‌های گراف هرس شده و ۵۹/۳۳٪ گره‌های گراف اولیه است.

سنجه‌های مرکزیت گروهی دیگر از سنجه‌ها هستند که برای تعیین مهم‌ترین و تأثیرگذارترین نویسندگان در شبکه‌های تألیف‌های مشترک استفاده می‌شوند. از بین سنجه‌های مرکزیت، سه سنجه مرکزیت درجه^{۱۵}، مرکزیت بینابینی^{۱۶} و مرکزیت بردارویژه^{۱۷} بیشتر از سایر سنجه‌های مرکزیت در تحلیل شبکه‌های تألیف‌های مشترک استفاده می‌شوند.

تعریف ۴.۴ ([۱۴، ص. ۱۶۸]). مرکزیت درجه برابر درجه یک گره یا تعداد یال‌های متصل به آن گره است.

گاهی از صورت وزن‌دار مرکزیت درجه استفاده می‌شود که برابر مجموع وزن یال‌های متصل به یک گره است.

تعریف ۵.۴ ([۱۴، ص. ۱۸۶]). مرکزیت بینابینی گره i برابر تعداد کوتاه‌ترین مسیرهایی بین گره‌های گراف است که از i می‌گذرند و با x_i نمایش داده می‌شود. به‌طور دقیق‌تر، مقدار $n_{s,t}^i$ را چنان تعریف می‌کنیم که اگر گره i در حداقل یکی از کوتاه‌ترین مسیرهای s به t قرار داشته باشد برابر ۱ وگرنه برابر ۰ باشد. در این صورت

$$(۲) \quad x_i = \sum_{s,t \in V(G)} n_{s,t}^i.$$

توجه کنید که در این تعریف مسیرهای هر گره به خودش نیز محاسبه شده است و هر یال غیرجهت‌دار نیز دو بار شمرده شده است؛ هرچند برخی ترجیح می‌دهند از صورت‌های تغییریافته این سنجه مرکزیت استفاده کنند که جهت‌دار نیست یا مسیرهای هر گره به خودش را نمی‌شمارد.

محاسبه مرکزیت بردارویژه پیچیده‌تر است و مرحله‌به‌مرحله صورت می‌گیرد. مرکزیت بردارویژه بر اساس این تفکر طراحی شده است که مرکزیت یک نقطه در شبکه نه تنها به جایگاه خود آن گره در شبکه بستگی دارد، بلکه به میزان اهمیت همسایگان آن وابسته است.

اگر مرکزیت بردارویژه گره i در شبکه را در هر مرحله با x_i نشان دهیم، مرکزیت بردارویژه آن گره در مرحله بعد با x'_i نشان داده می‌شود و از رابطه زیر محاسبه می‌گردد

$$(۳) \quad x'_i = \sum_{j \in V(G)} A_{i,j} x_j,$$

که در آن $A_{i,j}$ یک درایه از ماتریس مجاورت گراف G (A) است که اگر بین دو گره i و j یالی وجود داشته باشد برابر ۱ و در غیر این صورت برابر ۰ خواهد بود. محاسبه را می‌توان با هر مقدار اولیه مثبت دلخواه یکسانی برای همه x_i ها آغاز کرد. نتایج مثبت بودن تمامی مقادیر اولیه در [۱۴، ص. ۱۷۰] مفصل‌تر بررسی شده است. اگر x یک بردار با درایه‌های x_i باشد، می‌توان محاسبات را به شکل ماتریسی نوشت:

$$(۴) \quad x' = Ax$$

اگر شماره مرحله را نیز به نمادگذاری x اضافه کنیم:

$$(۵) \quad x(t) = Ax(t-1)$$

¹⁵degree centrality ¹⁶betweenness centrality ¹⁷eigenvector centrality

که یعنی

$$(۶) \quad \mathbf{x}(t) = \mathbf{A}^t \mathbf{x}(0)$$

اگر $\mathbf{x}(0)$ را به شکل ترکیبی خطی از بردارویژه‌های $\mathbf{v}_i$ ماتریس مجاورت بنویسیم، با در نظر گرفتن مقادیر اولیه x_i ، خواهیم داشت

$$(۷) \quad \mathbf{x}(0) = \sum_i c_i \mathbf{v}_i$$

و

$$(۸) \quad \mathbf{x}(t) = \mathbf{A}^t \sum_i c_i \mathbf{v}_i = \sum_i c_i \kappa_i^t \mathbf{v}_i = \kappa_1^t \sum_i c_i \left[\frac{\kappa_i}{\kappa_1} \right]^t \mathbf{v}_i,$$

که در آن κ_i مقدارویژه‌های $\mathbf{A}$ و κ_1 بزرگ‌ترین آن‌هاست.

مطابق قضیه پرون-فروبنیوس^{۱۸}، حداقل یک مقدارویژه نامنفی برای ماتریس مجاورت وجود دارد. با توجه به آن که κ_1 بزرگ‌ترین مقدارویژه این ماتریس است، می‌توان نتیجه گرفت $\kappa_1 \geq 0$. طبق یکی دیگر از نتایج قضیه مذکور، در گراف‌های همبند تنها یک بردارویژه با مقدارویژه‌ای برابر κ_1 وجود دارد. در گراف‌های ناهمبند معمولاً مرکزیت بردارویژه را برای هر مؤلفه همبندی جداگانه محاسبه می‌کنند. جزئیات بیشتر قضیه‌های ریاضی مورد استفاده در [۱۴، ص. ۳۴۵] بیان شده است.

چون در گراف‌های همبند برای $i \neq 1$ می‌توان گفت $\frac{\kappa_i}{\kappa_1} < 1$ ، پس با بزرگ شدن t مقدار همه جمله‌های مجموع جز جمله اول به ۰ میل می‌کنند؛ بنابراین

$$(۹) \quad \lim_{t \rightarrow \infty} \mathbf{x}(t) = \lim_{t \rightarrow \infty} \kappa_1^t \sum_i c_i \left[\frac{\kappa_i}{\kappa_1} \right]^t \mathbf{v}_i = \lim_{t \rightarrow \infty} \kappa_1^t c_1 \left[\frac{\kappa_1}{\kappa_1} \right]^t \mathbf{v}_1 = \lim_{t \rightarrow \infty} c_1 \kappa_1^t \mathbf{v}_1.$$

پس

$$(۱۰) \quad \mathbf{x} = \lim_{t \rightarrow \infty} \mathbf{x}(t) = \lim_{t \rightarrow \infty} c_1 \kappa_1^t \mathbf{v}_1.$$

بنابراین، مقدار حدی بردار مرکزیت‌ها ضریبی از بردارویژه پیشرو^{۱۹} ماتریس مجاورت است؛ یعنی در معادله زیر صدق می‌کند

$$(۱۱) \quad \mathbf{A}\mathbf{x} = \kappa_1 \mathbf{x}.$$

در نتیجه، یک مجموعه از مقادیر معتبر برای مرکزیت بردارویژه که در معادله ۳ صدق کند همان بردارویژه پیشرو ماتریس مجاورت است. جالب توجه است که مرکزیت گره‌ها در مقام مقایسه معنی‌دار است و مجموعه مقادیر معتبر یکتایی برای مرکزیت بردارویژه وجود ندارد.

اگر تمام مقادیر بردارویژه اولیه مثبت باشند و گراف نیز همبند باشد، مطابق قضیه پرون-فروبنیوس، همه درایه‌های بردارویژه پیشرو ماتریس مجاورت نامنفی است [۱۴، ص. ۱۷۰].

تعریف ۶.۴ ([۱۴، ص. ۱۷۰]). اگر $\mathbf{A}$ ماتریس مجاورت گراف همبند G و κ_1 بزرگ‌ترین مقدارویژه آن باشد، مرکزیت بردارویژه گره i برابر درایه x_i از یک بردار $\mathbf{x}$ است که در معادله ۱۱ صدق می‌کند

برای محاسبه عملی مرکزیت بردارویژه معمولاً از روش توانی استفاده می‌شود. در این روش که بر پایه معادله ۶ توسعه یافته است، با شروع از یک بردار آغازین دلخواه از مرکزیت گره‌ها (با شرایطی که ذکر شد)، هربار آن‌را در ماتریس مجاورت ضرب می‌کنند. با توجه به آنچه گفته شد، ماتریس به ضریبی از بردارویژه پیشرو ماتریس مجاورت میل خواهد کرد. این کار را آنقدر ادامه می‌دهند که به دقت مورد نظر برسند. همچنین برای پرهیز از کار با اعداد

¹⁸Perron-Frobenius ¹⁹leading eigenvector

بزرگ، معمولاً در هر مرحله همه اعداد ماتریس حاصل را بر عددی یکسان تقسیم می‌کنند. جزئیات بیشتر محاسبات ریاضی این روش در [۱۴، ص. ۳۴۵] در دسترس است.

مرکزیت بردارویژه محدودیت‌هایی دارد و مثلاً برای گراف‌های جهت‌دار بدون دور قابل استفاده نیست. با تغییراتی در این سنجۀ مرکزیت، سنجه‌های مرکزیت دیگری مانند کتر^{۲۰} و رتبه^{۲۱} صفحه طراحی شده است که بررسی آن‌ها از حوزه این مقاله خارج است. گاهی با تقسیم همه مرکزیت‌ها بر یک عدد خاص (معمولاً بزرگ‌ترین مرکزیت) صورت نرمال‌شده هر یک از سنجه‌های مرکزیت را محاسبه می‌کنند.

در شبکه‌تألیف‌های مشترک محققین علوم کامپیوتر در مجلات منتشرشده تا سال ۲۰۰۰، پس از حذف گره‌های تنها، برترین نویسندگان بر اساس مرکزیت درجه به‌ترتیب روزنفلد (۱۸۸)، اردوش (۱۷۸)، چپمن (۱۶۵)، آلون (۱۴۶)، و لیندزی (۱۴۳) هستند. بیشترین مرکزیت بر اساس مرکزیت بردارویژه نرمال‌شده نیز به لیو (۱)، میلر (۰/۹۸)، و منگ (۰/۹۸) اختصاص دارد که در یک مقاله بین‌المللی در سال ۱۹۹۹ با هم همکاری داشته‌اند.

۵. جمع‌بندی

در این پژوهش ابتدا به معرفی عدد اردوش و تاریخچه آن پرداخته شد. عدد اردوش حالتی خاص از یک سنجۀ گرافی شبکه‌تألیف‌های مشترک ریاضیدانان است؛ یکی از شبکه‌هایی که در سال‌های اخیر در مطالعات علم‌سنجی مورد توجه قرار گرفته‌اند. سپس شبکه‌های تألیف‌های مشترک معرفی شدند و درباره اهمیت و انواع آن‌ها توضیحاتی ارائه شد. همچنین، به‌عنوان نمونه، شبکه‌های تألیف‌های مشترک نویسندگان مجله حاضر و نیز شبکه‌تألیف‌های مشترک محققین علوم کامپیوتر ترسیم و بررسی گردید. در بخش پایانی نیز برخی ابزارهای گرافی مورد استفاده در تحلیل شبکه‌های تألیف‌های مشترک معرفی شد.

شبکه‌های تألیف‌های مشترک و تحلیل آن‌ها، در کنار تمایل سابقه‌دار ریاضیدانان معاصر به محاسبۀ عدد اردوش خودشان، یکی از نمونه‌های ارتباط دوطرفه ریاضیات کاربردی و علوم میان‌رشته‌ای نوین برآمده از علوم انسانی است.

تشکر و قدردانی

پژوهش‌ها درباره شبکه‌های تألیف مشترک به خصوص در علوم ریاضی وامدار فعالیت‌های جری گراسمان و همکارانش در دانشگاه اوکلند است. نویسنده لازم می‌داند از آن‌ها قدردانی نماید. همچنین، نویسنده مراتب قدردانی خود را از داوران محترم که با پیشنهادهای خود به تکمیل و واضح‌تر شدن این مقاله کمک‌های شایانی کردند اعلام می‌نماید.

مراجع

- [1] V. D. Blondel, J.-L. Guillaume, R. Lambiotte and E. Lefebvre, Fast unfolding of communities in large networks, *J. Stat. Mech.*, (2008). doi.org/10.1088/1742-5468/2008/10/P10008.
- [2] DBLP Team, DBLP: computer science bibliography, available at <https://dblp.org/>, 2021.
- [3] R. De Castro and J. W. Grossman, Famous trails to Paul Erdős, *Math. Intelligencer*, **21** (1999) 51–53.
- [4] T. M. J. Fruchterman and E. M. Reingold, Graph drawing by force-directed placement, *Softw. Pract. Exper.*, **21** (1991) 1129–1164.
- [5] C. Goffman, And what is your Erdős number?, *Amer. Math. Monthly*, **76** (1969) 791–791.
- [6] J. W. Grossman, Paul Erdős: The master of collaboration, *Algorithms and Combinatorics*, **14** (1997) 467–475.

²⁰Katz ²¹PageRank

- [7] J. W. Grossman, The Erdős number project, available at <https://oakland.edu/en>, 2021.
 - [8] J. W. Grossman, Facts about Erdős numbers and the collaboration graph, The Erdős number project, available at <https://oakland.edu/enp>, 2021.
 - [9] J. W. Grossman and P. D. F. Ion, On a portion of the well-known collaboration graph, *Congr. Numer.*, **108** (1995) 129-132.
 - [10] G. Harirchi, G. Melin and Sh. Etemad, An exploratory study of the feature of Iranian co-authorships in biology, chemistry and physics, *Scientometrics*, **72** (2007) 11-24.
 - [11] M. Jacomy, T. Venturini, S. Heymann and M. Bastian, ForceAtlas2, a continuous graph layout algorithm for handy network visualization designed for the Gephi software, *PLoS ONE*, **9** (2014) 12 pp. doi:10.1371/journal.pone.009867
 - [12] V. Ströele, R. Crivano, G. Zimbrão, J. M. Souza, F. Campos, J. M. N. David and R. Braga, Rational Erdős number and maximum flow as measurement models for scientific social network analysis, *J. Brazilian Computer Soc.*, **24** (2018) 17 pp.
 - [13] M. E. Newman, Coauthorship networks and patterns of scientific collaboration, *Proc. Natl. Acad. Sci. USA*, **101** (2004) 5200-5205.
 - [14] M. E. J. Newman, *Networks: An Introduction*, OUP, Oxford, 2010.
- [۱۵] م. ا. ریاحی، مجلات علمی: مجرای ارتباط دانشمندان، رهیافت، ۵ (۱۳۷۴) ۱۰ - ۱۹.
- [۱۶] ن. حریری و م. نیکزاد، شبکه‌های تألیف‌های مشترک در مقالات ایرانی رشته‌های کتابداری و اطلاع‌رسانی، روان‌شناسی، مدیریت و اقتصاد در پایگاه ISI بین سال‌های ۲۰۰۰ تا ۲۰۰۹، علوم و فناوری اطلاعات، ۲۶ (۱۳۹۰) ۸۲۵ - ۸۴۴.
- [۱۷] ف. سهیلی و ف. عصاره، بررسی تراکم و اندازه شبکه اجتماعی موجود در شبکه هم‌نویسندگی مجلات علم اطلاعات، پردازش و هدایت اطلاعات، ۲۹ (۱۳۹۲) ۳۵۱ - ۳۷۱.

هادی صفری

گروه مهندسی نرم‌افزار، دانشکده مهندسی برق و کامپیوتر، پردیس دانشکده‌های فنی دانشگاه تهران، تهران
hadi.safari@ut.ac.ir

هادی صفری مدرک کارشناسی مهندسی کامپیوتر خود را در گرایش نرم‌افزار در سال ۱۳۹۸ از دانشگاه تهران اخذ کرد و هم‌اکنون دانشجوی کارشناسی ارشد مهندسی نرم‌افزار دانشگاه تهران است. علم شبکه، تحلیل شبکه‌های اجتماعی، علم‌سنجی و آزمون مبتنی بر مدل نرم‌افزارهای کامپیوتری از علایق پژوهشی او هستند.

