

بررسی قابلیت به کارگیری سنج‌های مرکزیت به عنوان شاخص‌های ارتباط استنادی مدارک در بازیابی اطلاعات رابطه‌ای: مطالعه‌ی مقدماتی

جواد عباس‌پور (نویسنده‌ی مسئول)

دانشجوی دکتری علم اطلاعات و دانش‌شناسی دانشگاه تهران و بورسیه‌ی دانشگاه شیراز
javad.abbaspour@gmail.com

نادر نقشینه

استادیار گروه علم اطلاعات و دانش‌شناسی دانشگاه تهران
nnaghsh@ut.ac.ir

غلامرضا فدائی

استاد گروه علم اطلاعات و دانش‌شناسی دانشگاه تهران
ghfadaie@ut.ac.ir

فریده عصاره

استاد گروه علم اطلاعات و دانش‌شناسی دانشگاه شهیدچمران اهواز
osareh.f@gmail.com
تاریخ دریافت: ۱۳۹۱/۱۰/۰۹؛ تاریخ پذیرش: ۱۳۹۲/۰۱/۱۹

چکیده

هدف: هدف اصلی این پژوهش بررسی همبستگی سنج‌های مرکزیت (درجه‌ای، نزدیکی، بینابینی، و بردار ویژه) با سنج‌های اشتراک در مأخذ است که به‌عنوان یکی از سنج‌های استنادمحور برای برآورد میزان شباهت مدارک با یکدیگر در نظام‌های بازیابی اطلاعات رابطه‌ای شناخته شده است.

روش: این پژوهش مطالعه‌ی مقدماتی است که با رویکرد کمی و با روش تحلیل استنادی انجام شده است. جامعه‌ی این پژوهش را ۴۰ مقاله‌ی منتشر شده در سال ۲۰۱۱ و نمایه شده در پایگاه وب آو ساینس مربوط به چهار رشته‌ی علوم پایه/فنی - مهندسی (شامل فیزیک، شیمی، زیست‌شناسی و کامپیوتر) و چهار رشته‌ی علوم انسانی/اجتماعی (شامل اقتصاد، علوم تربیتی، روانشناسی و جامعه‌شناسی) به‌علاوه ۴۰۰۰ پیوند استنادی (۱۰۰ پیوند استنادی برای هر مقاله) که سایر مقاله‌ها با آنها برقرار کرده‌اند، تشکیل می‌دهند. داده‌های پژوهش با استفاده از نرم افزارهای Bib excel و UCINET گردآوری و تجزیه و تحلیل شدند.

یافته‌ها: یافته‌های پژوهش نشان داد که بین سنج‌های شباهت استنادی اشتراک در مأخذ و سنج‌های مرکزیت (درجه‌ای، بردار ویژه، نزدیکی و بینابینی) همبستگی مثبت در سطح آلفای ۰/۰۱ وجود دارد. بنابراین، سنج‌های مرکزیت ویژگی‌های لازم را برای اینکه مطالعات جامع‌تر و دقیق‌تری از حیث قابلیت به کارگیری به‌عنوان سنج‌های شباهت استناد محور روی آنها انجام شود، برخوردار هستند.

اصالت/ارزش: نتایج این پژوهش سنج‌های استنادی جدیدی را پیشنهاد می‌دهد که می‌توانند به‌عنوان سنج‌های مستقل و یا مکمل سنج‌های پیشین، در الگوریتم‌های بازیابی موتورهای جست‌وجو و پایگاه‌های اطلاعاتی به‌ویژه پایگاه‌های اطلاعاتی استنادی، برای سنجش دقیق‌تر شباهت مدارک با یکدیگر مورد استفاده قرار گیرند.

کلیدواژه‌ها: سنج‌های مرکزیت، اشتراک در مأخذ، سنج‌های شباهت، بازیابی اطلاعات رابطه‌ای، بازیابی

استناد محور

مقدمه

یکی از رویکردهایی که علیرغم سابقه نسبتاً طولانی آن در بازیابی اطلاعات، در طی سالیان اخیر بیشتر مورد توجه طراحان موتورهای جست‌وجو و پایگاه‌های اطلاعاتی قرار گرفته است «بازیابی رابطه‌ای»^۱ است. در سال‌های اخیر تعبیه گزینه‌هایی مانند Find similar و یا Related Documents و مواردی مانند آن به رابط کاربر موتورهای جست‌وجو و پایگاه‌های اطلاعاتی شناخته شده مانند گوگل اسکالر^۲، الزویر^۳، پاب مد^۴، به خوبی نشان دهنده گسترش و استقبال از این رویکرد است.

در این رویکرد، فرض بر آن است که هنگام جست‌وجو، به جای الزام کاربر به درج کلمات و کلیدواژه‌های مختلف و سپس اصلاح و دستکاری آنها، از منابع مرتبطی که کاربر پیش از انجام عملیات جست‌وجو در نظام بازیابی اطلاعات در اختیار دارد و یا در مراحل مختلف جست‌وجو می‌یابد - حتی اگر تعدادشان بسیار اندک باشد - به عنوان طعمه‌ای برای هدایت وی به منابع مشابه موجود در نظام و بسط دامنه منابع مرتبط بازیابی شده، بهره‌برداری شود.

از جمله شاخص‌ترین زمینه‌های مطالعاتی که در حوزه بازیابی رابطه‌ای و موضوعات وابسته به آن انجام گرفته‌اند می‌توان به کارآیی یا عدم کارآیی گزینه‌های «مشابه‌یاب» یا «مدارک مرتبط» در رابط کاربر موتورهای جست‌وجو و پایگاه‌های اطلاعاتی (مانند لین و دیگران^۵، ۲۰۰۷؛ حریری^۶، ۲۰۰۸)؛ وجود یا عدم وجود رابطه محتوایی میان مدارک استنادکننده و استنادشونده (مانند هارتر و دیگران^۷، ۱۹۹۳؛ پیترز و دیگران^۸؛ شقاقی و حرری، ۱۳۸۹)؛ مقایسه انواع سنجه‌های شباهت مدارک با یکدیگر (مانند بیچلر و ایتون^۹، ۱۹۸۰؛ مک کین^{۱۰}، ۱۹۸۹؛ پائو و ورتن^{۱۱}، ۱۹۸۹؛ الگرن و کولیاندر^{۱۲}، ۲۰۰۹؛ کرمانی، نشاط و حرری، زودآیند)؛ امکان به‌کارگیری سنجه‌های شباهت پیشین در محیط‌های جدید مانند وب (مانند منتسر^{۱۳}، ۲۰۰۴؛ کوتو^{۱۴} و دیگران، ۲۰۰۶)؛ و در نهایت، ارائه سنجه‌های شباهت جدید به‌منظور شناسایی مدارک مرتبط‌تر و برآورد دقیق‌تری از درجه شباهت مدارک با هم (مانند وانجانتوک و کین^{۱۵}، ۲۰۰۴؛ یارنوبینگ^{۱۶}، ۲۰۰۵؛ جانسن، پرالات و ویلسون^{۱۷}، ۲۰۱۰) اشاره کرد.

1. Relational Retrieval
 4. Pubmed
 7. Harter et al.
 10. Mccain
 13. Menczer
 16. Jameving

2. Google Scholar
 5. Lin et al.
 8. Peters et al.
 11. Pao & Worthen
 14. Couto et al
 17. Janssen, Pralat and Wilson

3. Elsevier
 6. Hariri
 9. Bichteler & Eaton
 12. Ahlgren & Colliander
 15. Wanjantuk & Keane

سنجه‌های شباهت پیشنهادی که تاکنون توسط متخصصان و محققان ارائه شده‌اند، بر اساس آنچه مبنای محاسبه شباهت قرار می‌گیرد، به سه گروه اصلی دسته‌بندی می‌شوند: سنجه‌های متن‌محور^۱، سنجه‌های استنادمحور^۲ و سنجه‌های ترکیبی^۳ (بویاک و کلاوانس^۴، ۲۰۱۰). از میان انواع سنجه‌های پیشنهادی استنادمحور، به نظر می‌رسد سنجه‌های اشتراک در مأخذ^۵ و استنادهای مشترک × معکوس فراوانی مدارک^۶ بیش از بقیه در پایگاه‌های استنادی معتبر بین‌المللی برای سنجش شباهت مدارک با یکدیگر و رتبه‌بندی آنها به کار گرفته شده‌اند. برای مثال، در پایگاه‌های استنادی وب آو ساینس^۷ و اسکوپوس^۸ از سنجه اشتراک در مأخذ (الگرن و یارنوینگ، ۲۰۰۸) و در پایگاه استنادی سایت‌سیر^۹ از استنادهای مشترک × معکوس فراوانی مدارک برای پیشنهاد مدارک مشابه به کاربر، استفاده می‌شود (وانجاتوک و کین، ۲۰۰۴). با وجود برخی تفاوت‌های جزئی که بین این دو سنجه شباهت وجود دارد؛ اما، هر دوی آنها بر مبنای استنادهای مشترکی بنا نهاده شده‌اند که میان مدارک وجود دارد (وانجاتوک و کین، ۲۰۰۴).

از رویکردهای دیگری که برای بررسی میزان شباهت مدارک با یکدیگر پیشنهاد شده است، سنجش شباهت بر اساس «فاصله» مدارک با یکدیگر در شبکه استنادی است (اسمال^{۱۰}، ۱۹۹۹؛ جانسن، پرالات و ویلسون^{۱۱}، ۲۰۱۰؛ خسروی و فتاحی، ۱۳۸۹). در حین مطالعات، سنجه‌هایی با عنوان «سنجه‌های مرکزیت»^{۱۲} نظر برخی پژوهشگران را به خود جلب نمود که در یک فضای دو بعدی گراف گونه و بر اساس تعداد پیوندهای میان اشیاء و فاصله آنها از نقاط کانونی شبکه، مجموعه‌ای از سنجه‌ها را به وجود آورده‌اند. «سنجه‌های مرکزیت» انواع مختلفی دارند که مهمترین آنها «مرکزیت درجه‌ای»^{۱۳}، «مرکزیت نزدیکی»^{۱۴}، «مرکزیت بینایی»^{۱۵} و «مرکزیت بردار ویژه»^{۱۶} است.

در این پژوهش، که در واقع یک مطالعه مقدماتی برای پژوهش‌های آتی به شمار می‌رود، قصد داریم این ایده را در یک آزمون اولیه مورد سنجش قرار دهیم. بنابراین مسأله اصلی

1. Text- Based Measures

3. Hybrid Measures

5. Bibliographic Coupling

7. Web of Science

9. CiteCeer

11. Janssen, Pralat, and Wilson

13. Degree Centrality

15. Betweenness Centrality

2. Citation- Based Measures

4. Boyack and Klavans

6. Common Citation × Inverse Document Frequency

8. Scopus

10. Small

12. Centrality

14. Closeness Centrality

16. Eigenvector Centrality

پژوهش این است که آیا سنجه‌های مرکزیت از قابلیت لازم برای به‌کارگیری به‌عنوان سنجه‌های شباهت استنادی مدارک برخوردار هستند یا خیر؟ بدین منظور همبستگی میان سنجه‌های مرکزیت را با سنجه شباهت استنادی «اشتراک در مأخذ» که هم از بُعد نظری و هم از جهت به‌کارگیری در عمل، یکی از شناخته شده‌ترین و معتبرترین سنجه‌های شباهت به شمار می‌رود، مورد آزمون قرار خواهیم داد.

پژوهش حاضر دارای اصطلاحاتی تخصصی است که لازم به معرفی است. درباره اصطلاح «مرکزیت» باید گفت که به زعم متخصصان جامعه‌شناسی و علوم اجتماعی، اگر هر رأس (گره) از گراف را یک فرد و یا در اصطلاح جامعه‌شناسی یک «کنش‌گر» به شمار آوریم و یال‌ها (خطوط) را معادل انواع تعاملات اجتماعی میان آنها در نظر بگیریم، برخی از رئوس به واسطه قرار گرفتن در موقعیت‌های کلیدی و خاص، نسبت به سایر رئوس از اهمیت بیشتری برخوردارند و در نتیجه کنترل بیشتری بر کل شبکه دارند. جامعه‌شناسان رئوس یاد شده را رئوس دارای «مرکزیت» نامیده‌اند (واسرمن و فاست^۱، ۱۹۹۴، ص ۱۷۳). به‌طور مشابه در شبکه استنادی مقاله‌ها، اگر هر رأس (گره) را برابر با یک مقاله و یال‌های (خطوط) آن را معادل با پیوندهای استنادی آنها در نظر بگیریم، برخی از مقاله‌ها به دلیل قرار گرفتن در موقعیت‌های کلیدی شبکه استنادی، نسبت به سایر مقاله‌ها از اهمیت بیشتری برخوردارند. به این مقاله‌ها، در تحقیق حاضر، مقاله‌های دارای «مرکزیت» می‌گوییم.

درباره اصطلاح «مرکزیت درجه‌ای» باید گفت که در یک گراف، رئوسی که دارای بیشترین یال‌ها (خطوط) هستند از بالاترین مرکزیت درجه‌ای برخوردارند و به آنها رئوس دارای مرکزیت درجه‌ای بالا می‌گویند. بر این اساس، در شبکه‌های استنادی، برخی از مقاله‌ها دارای بیشترین استناد هستند که در این پژوهش به این موارد، مقاله‌های دارای مرکزیت درجه‌ای بالا می‌گوییم.

«مرکزیت بردار ویژه» بر این اصل بنا نهاده شده است که همه یال‌های گراف دارای ارزش‌های یکسان نیستند بلکه یال‌هایی که با رئوس تأثیرگذار شبکه پیوند دارند به نسبت سایر یال‌ها از ارزش بالاتری برخوردارند. بنابراین رئوس با مقادیر ویژه بالاتر، رئوسی هستند که با

1. Wasserman and Faust

رئوس تأثیرگذار پیوند برقرار کرده‌اند (بورگاتی^۱، ۲۰۰۵؛ نیومن^۲، ۲۰۰۸). در این پژوهش منظور از مرکزیت بردار ویژه مقاله‌ها در شبکه استنادی، مقاله‌هایی است که با مقاله‌های تأثیرگذار در شبکه استنادی، پیوند برقرار ساخته‌اند.

«مرکزیت نزدیکی» بر اساس مفهوم فاصله یا طول مسیر بنا نهاده شده است. در یک گراف، رئوسی که دارای حداقل فاصله (کمترین تعداد یال‌ها) با تمامی رئوس دیگر هستند، مرکزیت نزدیکی بالاتری دارند (دگن و فورس^۳، ۱۹۹۹، ص ۱۳۵). در این پژوهش منظور از مرکزیت نزدیکی در شبکه استنادی، مقاله‌هایی است که با حداقل فاصله (کمترین تعداد پیوندها) با سایر مقاله‌های شبکه استنادی، پیوند برقرار کرده‌اند.

درباره «مرکزیت بینایی» باید گفت که بر اساس نظریه گراف، رئوسی که بین تعداد زیادی از فواصل ژئودزیک یا کوتاه‌ترین فاصله‌های رئوس قرار گرفته‌اند؛ از مرکزیت بینایی بالاتری برخوردارند (دگن و فورس، ۱۹۹۹، ص ۱۳۶). بنابراین در این پژوهش، مقاله‌های دارای مرکزیت بینایی بالا مقاله‌هایی هستند که در شبکه استنادی بین تعداد زیادی از کوتاه‌ترین فاصله‌ها قرار گرفته‌اند.

«اشتراک در مأخذ» یکی از سنجه‌های شباهت مدارک با یکدیگر است. در تعاریف آمده است که اگر دو مدرک در فهرست مأخذ خود به یک یا بیش از یک مدرک مشترک استناد کنند، آن دو مدرک با هم اشتراک در مأخذ خواهند داشت. هر چه دو مدرک تعداد منابع مشترک بیشتری داشته باشند، اشتراک در مأخذ قوی‌تر و در نتیجه احتمالاً شباهت بیشتری به هم خواهند داشت (گلنزل و زرون^۴، ۱۹۹۶). در این پژوهش به مقاله‌هایی که حداقل به یک یا بیش از یک مقاله مشترک استناد کرده باشند مقاله‌های دارای اشتراک در مأخذ گفته می‌شود.

پرسش‌های پژوهش

در این پژوهش در صدد پاسخ به دو پرسش اصلی زیر هستیم:

۱. آیا میان درجه اشتراک در مأخذ مقاله‌ها و سنجه‌های مرکزیت (درجه‌ای، بردار ویژه، نزدیکی و بینایی) همبستگی وجود دارد؟
۲. و در صورت مثبت بودن پاسخ پرسش بالا
۳. رتبه هر یک از سنجه‌های مرکزیت از نظر میزان همبستگی با اشتراک در مأخذ چگونه است؟

1. Borgatti

2. Newman

3. Degenne and Forse

4. Glanzel and Czerwon

روش پژوهش

این پژوهش، مطالعه‌ای است که با رویکرد کمی و با روش تحلیل استنادی انجام شده است. تحلیل استنادی یکی از فنون متداول کتاب‌سنجی است که در آن قواعد حاکم بر روابط میان مدرک استناد کننده (متن) و مدرک مورد استناد (سند) جست‌وجو و مطالعه می‌شود (حری، ۱۳۷۲).

جامعه این پژوهش را مقاله‌های منتشر شده در سال ۲۰۱۱ در پایگاه وب آو ساینس مربوط به چهار رشته علوم پایه/ فنی - مهندسی (شامل فیزیک، شیمی، زیست‌شناسی و کامپیوتر) و چهار رشته علوم انسانی/ اجتماعی (شامل اقتصاد، علوم تربیتی، روانشناسی و جامعه‌شناسی) تشکیل می‌دهند. از هر یک از این علوم به صورت تصادفی ۵ مقاله (جمعاً ۴۰ مقاله) به عنوان مقاله‌های اصلی انتخاب شدند. در مرحله بعد، شبکه استنادی هر مقاله در پایگاه وب آو ساینس استخراج گردید. سپس، با توجه به متغیر بودن تعداد پیوندهای استنادی هر مقاله و برای جلوگیری از تأثیرپذیری از تعداد پیوندهای استنادی، برای هر یک از مقاله‌های اصلی، به صورت تصادفی ۱۰۰ مقاله که دارای پیوند استنادی با مقاله منتخب بودند (جمعاً ۴۰۰۰ مقاله) انتخاب گردید. پس از این مرحله، با استفاده از نرم‌افزار Bib excel شبکه استنادی هر مقاله با پیوندهای استنادی آن، و خود پیوندهای استنادی با یکدیگر، ترسیم گردید.

داده‌های به دست آمده یک‌بار با استفاده از نرم‌افزار Bib excel برای محاسبه اشتراک در مآخذ هر مقاله موجود در شبکه استنادی، و بار دیگر با استفاده از نرم‌افزار UCINET برای محاسبه انواع مرکزیت (درجه‌ای، بردار ویژه، نزدیکی و بینایی) به کار گرفته شدند. در پایان، داده‌های حاصل از محاسبه اشتراک در مآخذ و سنجه‌های مرکزیت، با استفاده از نسخه ۱۶ نرم‌افزار تحلیل آماری SPSS مورد تجزیه و تحلیل قرار گرفتند. در این پژوهش از دو آزمون کولموگروف-اسمیرنف برای سنجش نرمال بودن داده‌ها و آزمون ضریب همبستگی اسپیرمن برای سنجش همبستگی میان متغیرها استفاده گردید.

یافته‌ها

برای پاسخ به پرسش‌های پژوهش، ابتدا پیش‌فرض‌های ورود به نوع آزمون مورد بررسی قرار گرفت. با توجه به اینکه همه متغیرهای پژوهش از نوع فاصله‌ای بودند و تعداد داده‌ها زیاد بود، با استفاده از آزمون کولموگروف-اسمیرنف، نرمال بودن توزیع داده‌ها بررسی شد (جدول ۱).

جدول ۱. نتایج آزمون کولموگروف-اسمیرنف روی توزیع داده‌های متغیرها

عوامل	آماره Z کولموگروف-اسمیرنف	سطح معنی داری	نتیجه
اشتراک در مآخذ	۰/۱۸۷	۰/۰۰۰	غیر نرمال
درجه	۰/۲۵۷	۰/۰۰۰	غیر نرمال
بینابینی	۰/۳۸۲	۰/۰۰۰	غیر نرمال
نزدیکی	۰/۲۴۸	۰/۰۰۰	غیر نرمال
بردار ویژه	۰/۲۶۹	۰/۰۰۰	غیر نرمال

نتایج آزمون کولموگروف-اسمیرنف نشان می‌دهد که در هیچ یک از انواع متغیرها، داده‌ها از توزیع نرمال برخوردار نیستند. بنابراین، برای سنجش همبستگی میان سنجه‌های مرکزیت با اشتراک در مآخذ بایستی از آزمون‌های ناپارامتری استفاده گردد. به همین دلیل، آزمون ضریب همبستگی اسپیرمن، به‌عنوان آزمون مطلوب مورد استفاده قرار گرفت. برای پاسخ به پرسش اول مبنی بر اینکه «آیا میان درجه اشتراک در مآخذ مقاله‌ها و سنجه‌های مرکزیت (درجه‌ای، بردار ویژه، نزدیکی و بینابینی) همبستگی وجود دارد یا خیر»، آزمون اسپیرمن انجام و در جدول زیر ارائه شده است. بر اساس یافته‌های حاصل از آزمون که در جدول ۲ ذکر شده‌اند، می‌توان دریافت که بین تمامی سنجه‌های مرکزیت با سنجه اشتراک در مآخذ همبستگی مثبت در سطح آلفای ۰/۰۱ وجود دارد.

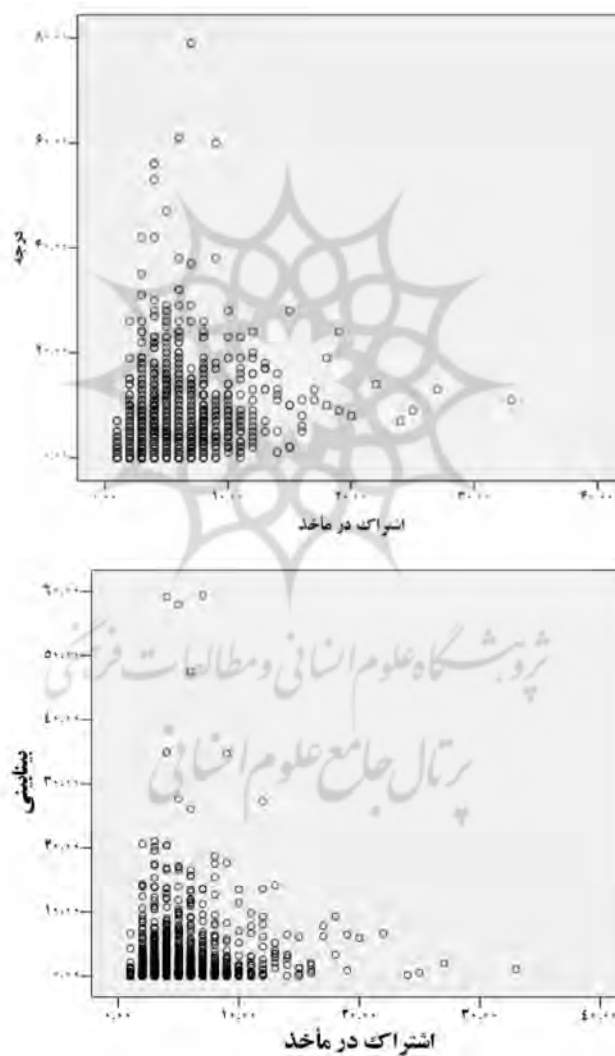
جدول ۲. ضریب همبستگی اسپیرمن بین اشتراک مآخذ و سنجه‌های مرکزیت

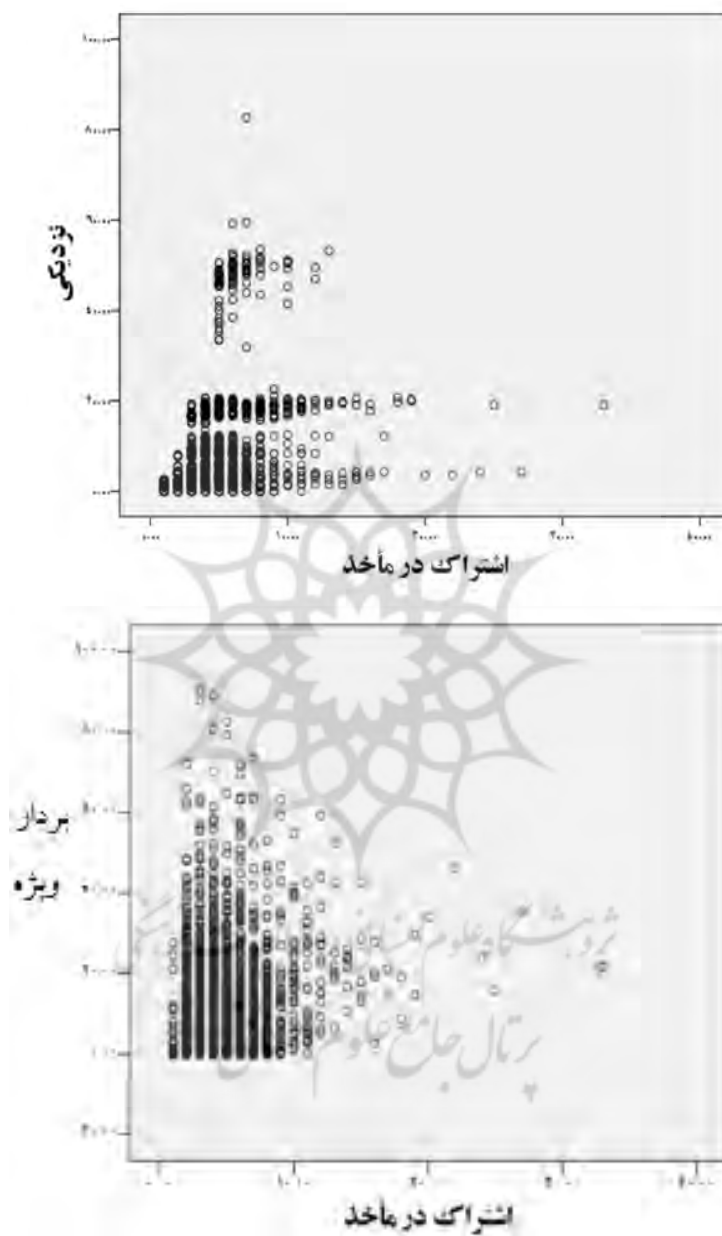
اشتراک در مآخذ	درجه	بینابینی	ویژه	نزدیکی
۱	۰/۴۳۵*	۰/۳۳۸*	۰/۳۹۱*	۰/۵۱۶*
۰/۴۳۵*	۱	۰/۸۵۱*	۰/۸۲۹*	۰/۸۰۲*
۰/۳۳۸*	۰/۸۵۱*	۱	۰/۶۶۹*	۰/۵۶۶*
۰/۳۹۱*	۰/۸۲۹*	۰/۶۶۹*	۱	۰/۶۷۸*
۰/۵۱۶*	۰/۸۰۲*	۰/۵۶۶*	۰/۶۷۸*	۱

* همبستگی در سطح $\alpha = 0.01$ معنی‌دار است.

با توجه به مثبت بودن پاسخ پرسش اول، پرسش دوم مبنی بر اینکه «رتبه هریک از سنجه‌های مرکزیت از نظر میزان همبستگی با سنجه اشتراک در مآخذ چگونه است»، بررسی

گردید. از نظر میزان همبستگی سنجه‌های مرکزیت با سنجه اشتراک در مأخذ، یافته‌ها نشان می‌دهد که مرکزیت نزدیکی با $I=0/516$ بیشترین همبستگی را داراست و مرکزیت درجه‌ای با $I=0/435$ و بردار ویژه با $I=0/391$ در مراتب بعدی قرار دارند. همچنین مرکزیت بینابینی با $I=0/338$ کمترین همبستگی را داراست. شکل ۱ نمودار توزیع پراکنندگی سنجه اشتراک در مأخذ با سنجه‌های درجه، بینابینی، بردار ویژه و نزدیکی را نشان می‌دهد:





شکل ۱. نمودار توزیع پراکندگی سنجه اشتراک در مآخذ با سنجه‌های مرکزیت

نتیجه‌گیری

پیش از این، در پژوهش‌های علوم کتابداری و اطلاع‌رسانی سنجه‌های مرکزیت در بسترهای گوناگون مورد آزمون قرار گرفته‌اند. برای مثال، چن^۱ (۲۰۰۶) از سنجه مرکزیت برای شناسایی نقاطی محوری که در طی زمان تغییرات پارادایمی در متون علمی ایجاد می‌کنند استفاده کرد. لیدسدورف^۲ (۲۰۰۷) دریافت که سنجه مرکزیت بینایی بیان‌کننده ارتباطات بین رشته‌ای مجلات نمایه شده در پایگاه گزارش‌های استنادی مجله^۳ است. لیو و دیگران^۴ (۲۰۰۵) سه سنجه مرکزیت (درج‌های، نزدیکی و بینایی) را برای شبکه هم‌تألیفی حوزه کتابخانه‌های دیجیتال به کار گرفتند و به این نتیجه رسیدند که از میان سنجه‌های مرکزیت به کار گرفته شده، سنجه مرکزیت بینایی بهتر از سایر سنجه‌ها برای رتبه‌بندی مؤلفان در شبکه تألیفی عمل می‌کند. علاوه بر این، یان و دینگ^۵ (۲۰۰۹) در پژوهش خود به این نتیجه دست یافتند که در شبکه‌های هم‌تألیفی، سنجه‌های مرکزیت با تعداد استنادها همبستگی قوی (در سطح آلفای ۰/۰۱) دارند.

در این پژوهش، قابلیت سنجه‌های مرکزیت از زاویه‌ای دیگر، یعنی سنجش شباهت مدارک با هم و رتبه‌بندی آنها در بازیابی رابطه‌ای مطمئن نظر قرار گرفت. یافته‌های این پژوهش، که در یک مقیاس کوچک انجام شد، به طور کلی نشان داد که سنجه‌های مرکزیت علاوه بر کاربردهای پیشنهادی در پژوهش‌های پیشین، مانند شناخت تغییرات پارادایمی در متون علمی، شناسایی ارتباطات میان رشته‌ای و رتبه‌بندی مؤلفان، از امکان لازم برای اینکه به‌عنوان یک سنجه شباهت مدارک مورد استفاده قرار گیرند، برخوردار هستند. با توجه به اینکه ضریب همبستگی به‌دست آمده برای تمامی سنجه‌های مرکزیت در سطح آلفای ۰/۰۱ معنی‌دار است، می‌توان نتیجه گرفت که نه تنها میان سنجه‌های مرکزیت و اشتراک در مأخذ ارتباط مستقیم وجود دارد؛ بلکه، ارتباط میان آنها یک ارتباط قوی است. به عبارت دیگر، سنجه‌های مرکزیت قادرند سنجه اشتراک در مأخذ را پیش‌بینی نمایند.

از طرفی، یافته‌های این پژوهش حاکی از آن است که از میان انواع سنجه‌های مرکزیت، مرکزیت نزدیکی به دلیل اینکه بیشترین همبستگی (۰/۵۱۶) را با سنجه اشتراک در مأخذ

1. Chen
3. Journal Citation Reports (JCR)
5. Yan and Ding

2. Leydesdorff
4. Liu et al.
6. Citation Counts

برخوردار است، به احتمال فراوان از بالاترین امکان برای شناسایی مدارک مرتبط برخوردار است. از این حیث، سنجه‌های مرکزیت درجه‌ای و بردار ویژه در رتبه‌های بعدی قرار دارند و سنجه بینابینی در پایین‌ترین سطح قرار دارد. این نتایج نشان می‌دهد که بر خلاف نتایج پژوهش‌های پیشین که مرکزیت بینابینی را بهترین گزینه از میان انواع سنجه‌های مرکزیت برای تبیین ارتباطات میان رشته‌ای (لیدسدورف، ۲۰۰۷) و رتبه‌بندی مؤلفان (لیو و دیگران، ۲۰۰۵) پیشنهاد داده‌اند، برای سنجش ارتباط محتوایی مدارک با یکدیگر، سنجه «مرکزیت نزدیکی» عملکرد به مراتب بهتری از خود به نمایش می‌گذارد.

لازم به ذکر است که اظهار نظر دقیق‌تر در رابطه با نتایج این پژوهش منوط به انجام پژوهش‌های جامع‌تر است که از زوایای گوناگون این موضوع را مورد کنکاش قرار دهند. بنابراین، در پژوهش‌های آتی قصد داریم تا ضمن انجام پژوهشی مشابه در سطحی گسترده‌تر، قابلیت تعمیم نتایج به علوم مختلف را مورد بررسی قرار دهیم. همچنین سنجه‌های مرکزیت را که از جمله سنجه‌های استنادمحور محسوب می‌شوند، با یک سنجه معتبر از نوع متن‌محور، محک بزیم. پژوهشگران انتظار دارند - در صورت تکرار یافته‌های این پژوهش در پژوهش‌های آتی - بتوان از سنجه‌های مرکزیت در الگوریتم‌های بازیابی پایگاه‌های اطلاعاتی، به ویژه پایگاه‌های استنادی، و موتورهای جست‌وجو برای تشخیص بهینه مدارک مرتبط و رتبه‌بندی دقیق‌تر آنها در نتایج بازیابی شده برای کاربر، بهره گرفت.

منابع

- حری، عباس (۱۳۹۱). تحلیل استنادی. در *مروری بر اطلاعات و اطلاع‌رسانی*. چاپ دوم. تهران: نهاد کتابخانه‌های عمومی کشور.
- خسروی، عبدالرسول و فتاحی، رحمت‌ا. (۱۳۸۹). تحلیل ساختار و الگوریتم ذخیره و بازیابی اطلاعات در پایگاه‌های استنادی. *پژوهشنامه علوم و فن آوری اطلاعات*. ۲۶ (۹)، ۱۹۹-۲۲۲.
- شقاقی، مهدی و حری، عباس (۱۳۸۹). شباهت‌های موضوعی و شباهت‌های استنادی: مطالعه‌ای در پایان‌نامه‌های حوزه کتابداری و اطلاع‌رسانی. *کتابداری و اطلاع‌رسانی*. ۱۳ (۲)، ۱۶۷-۱۹۰.
- کرمانی، آناهیتا؛ نشاط، نرگس و حری، عباس (زودآیند). مقایسه عملکرد اشتراک در مآخذ و اشتراک واژگان عناوین استنادها در خوشه‌بندی فازی پروانه‌های ثبت اختراع. *پژوهشنامه پردازش و مدیریت اطلاعات*. بازیابی شده در تاریخ ۲۴ فروردین ۱۳۹۲ از: <http://Jipm.irandoc.ac.ir>

References

- Ahlgren, P. & Colliander, C. (2009). *Document Similarity Approaches and Science Mapping: Experimental Comparison of Five Approaches*. *Journal of Informetrics*, 3, 49-63.
- Ahlgren, P. & Jarneving, B. (2008). *Bibliographic Coupling, Common Abstract Stems and Clustering: A Comparison of Two Document – Document Similarity Approaches in the Context of Science Mapping*. *Scientometrics*, 76 (2), 273-290.
- Bichteler, J. & Eaton, E. A. (1980). *The Combined Use of Bibliographic Coupling and Cocitation for Document Retrieval*. *Journal of the American Society for Information Science*, 31 (4), 278-282.
- Borgatti, P. (2007). Centrality and Network Flow. *Social Networks*, 27, 55–71.
- Boyack, K. W. & Klavans, R. (2010). *Co-Citation Analysis, Bibliographic Coupling , and Direct Citation: Which Citation Approach Represents the Research Front Most Accurately?* *Journal of the American Society for Information Science and Technology*, 61 (12), 2389-2304.
- Chen, C. M. (2006). *Cite Space II: Detecting and Visualizing Emerging Trends and Transient Patterns in Scientific Literature*. *Journal of the American Society for Information Science and Technology*, 57 (3), 359–377.
- Couto, T.; Cristo, M.; Andre, M. G.; Calado, P.; Ziviani, N.; Moura, E. et al. (2006). *A Comparative Study of Citations and Links in Document Classification*. In *Proceedings of the 6th ACM/IEEE-CS joint conference on Digital Libraries*, 75-84.
- Degenne, A. & Forse, M. (1999). *Introducing Social Networks*. London: Sage Publication.
- Glanzel, W. & Czerwon, H. J. (1996). *A New Methodological Approach to Bibliographic Coupling and its Application to the National, Regional and Institutional Level*. *Scientometrics*, 37 (2), 195-221.
- Hariri, H. (2008). *An Investigation of the Effectiveness of the "Similar Pages" Feature of Google*. *Online Information Review*, 32 (3), 370 – 378.
- Harter, S. P. & Nisonger, T. E. (1993). *Semantic Relationships between Cited and Citing Articles in Library and Information Science Journals*. *Journal of the American Society for Information Science*, 44 (9), 543-552.
- Janssen, J.; Pralat, P. & Wilson, R. (2010). *Estimating Node Similarity from Co-Citation in a Spatial Graph Model*. *Proceedings of the 2010 ACM Symposium on Applied Computing, March 22-26, 2010, Sierre, Switzerland*.
- Jarneving, B. (2005). *The Outbreak of SARS Mirrored by Bibliometric Mapping: Combining Bibliographic Coupling with the Complete Link Cluster Method*. *Library and Information Science Research*, 11, 37-63.
- Leydesdorff, L. (2007). *Betweenness Nentrality as an Indicator of the Interdisciplinarity of Scientific Journals*. *Journal of the American Society for Information Science and Technology*, 58 (9), 1303–1319.
- Lin J.; DiCuccio M.; Grigoryan V. & Wilbur W. J. (2007). *Exploring the Effectiveness of Related Article Search in PubMed*. In *Technical Report. LAMP-TR-145/CS-TR-4877/UMIACS-TR-2007-36/HCIL-2007-10 University of Maryland, College Park, Maryland*.

- Liu, X.; Bollen, J.; Nelson, M. L. & Sompel, H. V. (2005). *Co-Authorship Networks in the Digital Library Research Community*. Information Processing and Management, 41, 1462-1480.
- Mccain, W. (1989). *Descriptor and Citation Retrieval in the Medical Behavioral Sciences Literature: Retrieval Overlaps and Novelty Distribution*. Journal of the American Society for Information Science, 40 (2), 110-114.
- Menczer, F. (2004). *Combining Link and Content Analysis to Estimate Semantic Similarity*. Proceedings of the 13th international World Wide Web conference on Alternate track papers & posters. May 19-21, 2004, New York, NY, USA : 452-453.
- Newman, M. (2008). The Mathematics of Networks. In: Blume L., editor, The New Palgrave Encyclopedia of Economics. 2nd edition. Basingstoke: Palgrave Macmillan.
- Pao, M. L. & Worthen, B. (1989). *Retrieval Effectiveness by Semantic and Citation Searching*. Journal of the American Society for Information Science, 40 (4), 226-235.
- Peters, H. P. F.; Braam, F. R. & Raan, A. F. J. V. (1995). *Cognitive Resemblance and Citation Relations in Chemical Engineering Publications*. Journal of the American Society for Information Science, 46 (1), 9-21.
- Small, H. (1999). *Visualizing Science by Citation Mapping*. Journal of the American Society for Information Science, 50 (9), 799-813.
- Wanjanatuk, P. & Keane, J. A. (2004). *Finding Related Documents via Communities in the Citation Graph*. International Symposium on Communications and Information Technology, IEEE International Symposium on 26-29 Oct., 445-450.
- Wasserman, S. & K. Faust (1994). *Social Network Analysis: Methods and Applications*. Cambridge: Cambridge University Press.
- Yan, E. & Ding, Y. (2009). *Applying Centrality Measures to Impact Analysis: A Co-Authorship Network Analysis*. Journal of the American Society for Information Science, 60 (10), 2107-2118.
- Zhang, J. & Korfhage, R. (1999). *A Distance and Angle Similarity Measure Method*. Journal of the American Society for Information Science. 50 (9), 772-778.

به این مقاله این‌گونه استناد کنید:

عباس‌پور، جواد؛ نقشینه، نادر؛ فدائی، غلامرضا و عصاره، فریده (۱۳۹۱). بررسی قابلیت به‌کارگیری سنجه‌های مرکزیت به عنوان شاخص‌های ارتباط استنادی مدارک در بازیابی اطلاعات رابطه‌ای: مطالعه‌ی مقدماتی. *تحقیقات اطلاع‌رسانی و کتابخانه‌های عمومی*، ۱۹ (۱)، ۱۴۹-۱۶۱.