

رویکردهای ارزیابی نظامهای بازیابی اطلاعات: پس‌زمینه و چشم‌انداز پیش‌رو

محمدرضا کیانی^۱

چکیده

این نوشتار با نگاهی گذرا به سیر تکاملی نظامهای بازیابی اطلاعات و ارزیابی آنها، انواع ارزیابی در این نظامها را از هم تفکیک و به طور مفصل‌تر دربارهٔ رویکردهای موجود در ارزیابی عملکرد نظامهای بازیابی، بحث کرده است. با نقد هر کدام، مدلی ترکیبی با رویکردی نظری - تصمیمی از «وانگ و فورگیونه» را معرفی می‌کند و در پایان نتایج و پیشنهادها را ارائه می‌دهد.

کلیدواژه‌ها: ارزیابی نظامها، بازیابی اطلاعات، رویکردهای ارزیابی.

مقدمه

نظامهای بازیابی اطلاعات^۲ به بازنمون، ذخیره‌سازی و بازیابی اطلاعات و دانش مرتبط با نیازها و سؤالهای خاص کاربران می‌پردازند. کاربر در قالب یک پرسش، با واژه‌هایی از زبان طبیعی نیاز خود را مطرح می‌کند. این پرسش به وسیلهٔ نظام و بر اساس یک الگوریتم مشخص با مدارک نمایه‌سازی شده مقایسه و مدارکی که بیشترین مشابهت را با پرسش کاربر دارند، عرضه می‌شوند. در نهایت، این کاربر است که میزان

۱. عضو هیئت علمی دانشگاه بیرجند و دانشجوی دکترای کتابداری و اطلاع‌رسانی دانشگاه فردوسی

مشهد (kiani.mreza@gmail.com).

2. Information Retrieval System (IRS).

ربط آنها را با نیاز خود مورد سنجش قرار می‌دهد(مندل^۱، ۲۰۰۸). بازیابی اطلاعات، فرایندی دارای ابهام است. این ابهام به مبهم بودن عناصر دخیل در آن یعنی زبان طبیعی، نیاز اطلاعاتی، ربط و ... باز می‌گردد.

پس زمینه تاریخی نظامهای بازیابی رایانه‌ای، به گسترش روشهای نمایه‌سازی خودکار متون در دهه ۶۰ میلادی باز می‌گردد؛ روشی که همچنان به صورت گسترده توسط ارائه‌دهندگان اطلاعات و همچنین در خدمات اینترنتی به کار می‌رود. در دهه ۷۰، مدل‌های بازیابی بولی و فضای برداری و بعدها مدل‌های بازیابی احتمالاتی گسترش یافتند. لکن پذیرش تجاری آنها به عنوان مدل‌های انطباق جزئی (در مقابل مدل‌های انطباق کامل اولیه) در موتورهای کاوش اینترنتی که نتایج را به صورت فهرستی رتبه‌بندی شده ارائه می‌دهند، تا دهه ۹۰ به تعویق افتاد.

امروزه نظامهای بازیابی اطلاعات نقش مهمی را در تمرکز اطلاعاتی^۲ و توسعه دانش بازی می‌کنند (وانگ و فورگیونه^۳، ۲۰۰۶) چنان‌که مدیریت دانش به منظور دسترس‌پذیر کردن حجم قابل توجه اطلاعات و مجموعه‌های گسترده علمی متنی، بدون این نظامها قابل تصور نیست. در زمان ما، استفاده از وب و موتورهای کاوش جزئی از زندگی روزمره تلقی می‌شود که با روشها، ابزار و رویکردهای نظامهای بازیابی اطلاعات عجین شده‌اند. در این راستا، همچون بسیاری از مفاهیم و فرایندهای دیگر، انسان دست به ارزیابی می‌زند و درباره کیفیت کار اظهار نظر می‌کند.

ارزیابی نظامهای بازیابی اطلاعات

زمانی که صحبت از ارزیابی می‌شود، منظور داوری در مورد کیفیت و کارآمدی است. به اعتقاد «ساراسویک»^۴ (۱۹۹۵) ارزیابی نفوذ عمده‌ای در تحقیقات، توسعه و

-
1. Mandl, T.
 2. Information focusing.
 3. Wang, Y.D., & Forgionne, G.
 4. Saracevic, T.

کاربردهای مربوط به بازیابی اطلاعات دارد. به نظر «بائزا - ریبیرو - نتو»^۱ (۱۳۸۵) در ارزیابی نظامهای بازیابی اطلاعات، به دو جنبه کلی تحلیل کارکردی^۲ و ارزیابی عملکرد^۳ نظام پرداخته می‌شود. در بخش تحلیل کارکردی نظام، موضوعاتی مانند آزمون تک تک کارکردها با معیارهای اثربخشی (هدفها)، هزینه - کارایی (هزینه‌های نهایی) و هزینه - سودمندی (سود ناشی از خدمات) و همچنین تحلیل خطا، مورد توجه قرار می‌گیرد.

عملکرد نظام از نظر فضا و زمان و معیارهای دیگری که مهم‌ترین آنها ربط^۴ است، ارزیابی می‌شود. در متون حوزه ما، زمانی که بحث ارزیابی نظامهای بازیابی اطلاعات مطرح می‌شود، بیشتر منظور مورد اخیر یعنی ارزیابی عملکرد بازیابی نظام از نظر ربط است.

هدف از توضیح این قسمت این بود که باید بین ارزیابی کل نظام بازیابی، ارزیابی عملکرد آن و ارزیابی بازیابی نظام از نظر ربط، تفاوت قایل شد؛ هرچند در بعضی از سنجش‌های ارائه شده، علاوه بر ربط، ارزیابی عملکرد به طور کلی، یعنی با توجه به جنبه‌های زمان و فضا، مورد توجه بوده است.

هدف از ارزیابی کل نظام، آمادگی برای پاسخگویی به منابع تأمین بودجه و توجیه عملکرد نظام و اثبات ارزشمندی آن، اطمینان از تداوم خدمات مناسب و اثربخشی (با مقایسه سطح مورد نظر عملکرد و سطح کنونی)، حل مشکلات نظام (بررسی نقاط قوت و ضعف) و به دست آوردن اطلاعات برای طرح‌ریزی عملکرد آتی و یا طراحی و پیاده‌سازی نظامهای جدید است (پائو، ۱۳۷۹).

در ارزیابی عملکرد یک نظام، هرچه زمان پاسخ کوتاه‌تر و فضای مورد استفاده [ی داده‌ها] کوچکتر باشد، نظام در وضعیت بهتری است که مربوط می‌شود به عملکرد

1. Baeza-Yates, Ricardo; Ribeiro-Neto, Berthier.

2. Functional analysis.

3. Performance.

4. Relevance.

ساختارهای نمایه‌سازی، تعامل با سیستم عامل، تأخیرهای کانالهای ارتباطی و پیامدهای لایه‌های نرم‌افزاری عمل‌کننده در پس زمینه (بائزا- بیتس و ریبرو- نتو، ۱۳۸۵). بنابراین، یک نظام خوب، نظامی است که نیازهای کاربر را برآورده سازد. برآوردن این نیازها، به عملکرد خوب نظام در ابعاد مختلف بستگی دارد. کیفیت پاسخها در ارتباط با نیاز اطلاعاتی، سرعت پاسخگویی و محیط رابط کاربر نظام، مهم‌ترین ابعاد هستند که در بین این سه، ارزیابی مورد اول از همه دشوارتر است.

رویکردهای ارزیابی عملکرد بازیابی نظام (ربط) الف) رویکرد نظام مدار

ارزیابی‌ها اغلب در بازیابی اطلاعات از رویکرد نظام مدارانه^۱ استفاده می‌کنند که آن هم بیشتر بر تعیین الگوریتمهای جستجو با استفاده از سنج‌های آماری بویژه دو ضریب بسیار شناخته شده «بازیافت و دقت^۲» متمرکز است (سالتون، ۱۹۹۲). نمونه‌های این روش‌شناسی ارزیابی که به «پارادایم کرنفیلد^۳» معروف است (کلوردون، ۱۹۹۷) پروژه‌های کرنفیلد (کلوردون، میلز، و کین، ۱۹۶۶)، «اسمارت^۴» (سالتون، ۱۹۷۱، ۱۹۸۹)، «استیرز^۵» (بلرو میرون، ۱۹۹۰، ۱۹۸۵)، و «ترک^۶» (از سال ۱۹۹۲ توسط موسسه استاندارد و فناوری) است.

این مطالعات اگرچه به صورت طرح تجربی استاندارد انجام شده‌اند، چشم‌انداز آنها به ارزیابی، کامل نیست. دو اشکال اساسی به این رویکرد گرفته‌اند. نخست، از ربط موضوعی به عنوان مبنا یا تنها معیار ارزیابی بازیابی استفاده می‌شود. سنج‌های آماری عملکرد مانند بازیافت و دقت، به زمینه متن کار ندارند و مستقل از متن هستند و منعکس‌کننده فرایند تعاملی بازیابی نیستند و تنها نوعی داوری مقدماتی در مورد

1. System oriented approach.
2. Recall & Precision.
3. Cranfield paradigm.
4. Smart.
5. Stairs.
6. TREC(Text REtrieval Conference).

بازنمون دوجهی ربط (مربوط یا نامربوط) می دهند (بورلند، ۲۰۰۰، ۲۰۰۳). از طرفی، تنها برون داد را مورد توجه قرار می دهند و راهبردی یکجایه (که در ادامه توضیح داده خواهد شد) دارند و به فرایند تعاملی بازیابی توجهی ندارند. از سوی دیگر، مدارک را به مربوط و نامربوط تقسیم می کنند. هرچند جامعیت و مانعیت باز تعریف شده اند (ککالاینن و جارولین^۱، ۲۰۰۲)، همچنان این سنجه های ربط محور از انعکاس طبیعت پویا و چند بعدی ربط ناتوان است (ساراسویک، ۱۹۹۶).

دوم، بُعد کاربری از فرایند ارزیابی حذف شده است، در حالی که ربط مفهومی شناختی است که معنای آن به درک کاربران از اطلاعات بازیابی شده و موقعیتهای نیاز اطلاعاتی خودشان ارتباط عمده دارد (بورلاند، ۲۰۰۳). این رویکرد تنها به درجه ای که بازنمون پرسش (عبارت جستجو)^۲ با متن عینیت های اطلاعاتی بازیابی شده مطابقت دارد، می پردازد. به بیان دیگر، مطالعات ارزیابی در این حوزه، تعریف نیاز اطلاعاتی توسط سیستم را به جای تعریف های کاربران پذیرفته اند. کاربر در این رویکرد یک انتزاع^۳ است نه کاربر واقعی (مندل، ۲۰۰۸).

ب) رویکرد کاربر مدار

با تغییر مسیر پارادایم به سمت جنبه های رفتاری و شناختی در بازیابی اطلاعات، حجم عمده ای از مطالعات بر ارزیابی رضایت، عملکرد و استفاده کاربر نهایی از نظامهای بازیابی اطلاعات تمرکز کردند (الیس، ۱۹۹۶؛ رابرتسون و هنکاک بولیو، ۱۹۹۲؛ چمبر، ۱۹۹۴). مطالعات کاربر مدار^۴ به کاربران و سطوح استفاده آنها می پردازند. به دلیل دشواری ارزیابی جنبه های فردی جستجوها و ارزیابی قضاوت کاربر از سودمندی نتایج، کار در این رویکرد بسیار دشوارتر از رویکرد نظام مدار است. در این رویکرد، تمرکز بیشتر بر فرایند است تا خروجی ها.

1. Kekalainen, J., & Jarvelin, K.

2. Query.

3. Abstraction.

4. User oriented.

دسته‌بندی از نظر راهبرد جستجو

شاید مناسب باشد قبل از تقسیم مطالعات مربوط به ارزیابی عملکرد به نظام مدار و کاربر مدار، آنها را از نظر توجه به ماهیت درخواست پرسش، دسته‌بندی کنیم. در این تقسیم‌بندی، راهبرد جستجو مورد توجه قرار می‌گیرد. دو نوع راهبرد جستجوی ذکر شده، عبارتند از: راهبرد یکجایه^۱ و راهبرد تعاملی^۲.

در نوع اول، کاربر پرسش خود را ارائه و پاسخی را دریافت می‌کند. اما در راهبرد تعاملی، کاربر نیاز اطلاعاتی خود را از طریق مجموعه‌ای از مراحل تعامل با نظام، مشخص می‌کند.

در راهبرد یکجایه، وجه غالب کیفیت مجموعه پاسخ ارائه شده است، که خود می‌تواند نظام مدار یا کاربرمدار باشد. مطالعاتی که با این راهبرد صورت گرفته، ابتدا در محیط آزمایشگاهی (مانند ترک) و سپس در محیط واقعی انجام شدند. معیارهای ارزیابی^۳ راهبرد یکجایه در رویکرد نظام‌مدار عبارتند از: جامعیت و مانعیت، سنجه ارزیابی E، و میانگین توافقی F. این معیارها در رویکرد کاربرمدار عبارتند از: دستاورد جامعیت^۴، جامعیت نسبی^۵، نسبت تازگی^۶ و نسبت پوشش^۷.

در راهبرد تعاملی، بیشتر از کیفیت مجموعه پاسخ ارائه شده، به فرایند تعامل و اجزای آن مانند تلاش کاربر، ویژگیهای طراحی محیط رابط^۸، راهنماییهای ارائه شده از سوی نظام و طول مدت تعامل توجه می‌شود. بنابراین، باید بین راهبرد تعاملی با رویکرد کاربرمدار یکجایه، تفاوت قایل شد.

منظور از تعامل، ارتباط بین اطلاع‌جویان انسانی و نظامهای بازیابی اطلاعات است.

1. Batch mode.

2. Interactive.

۳. برای توضیح بیشتر به کتاب «قلمروهای نو در بازیابی اطلاعات» مراجعه شود.

4. Recall effort.

5. Relative recall.

6. Novelty ratio.

7. Coverage ratio.

8. User interface.

وقتی کاربران به یک نظام بازیابی اطلاعات وارد می‌شوند، اغلب از نحوه دستیابی خود به اطلاعات، درک مبهمی دارند. بنابراین، تعامل میان سیستم و کاربر در مورد درک و بیان نیازهای اطلاعاتی، تدوین پرسشها، انتخاب از میان مآخذ موجود، درک نتایج کاوش و پیگیری روند پیشرفت کاوش کاربران، می‌تواند انجام شود. رفتار اطلاع‌یابی^۱ و بازخورد ربط^۲ در این بخش با نظامهای بازیابی پیوند می‌خورند. بر اساس این نوع نگاه، ارزیابی با آزمایش اطلاع‌جویی واقعی کاربر نهایی و فرایندهای بازیابی آغاز می‌شود (جارولین و اینگورسن^۳، ۲۰۰۴). مدل اطلاع‌یابی گزینشی (میوه چینی)^۴ بیستس و سایر مدل‌های تشریح و بازنمایی اعمال و فعالیت‌های شناختی کاربر در خلال فرایند بازیابی اطلاعات (به عنوان مثال بلکین، ۱۹۸۰؛ اینگورسن، ۱۹۹۶؛ کولثا، ۱۹۹۱) از این دسته‌اند. در ارزیابی نظامهای تعاملی، به طراحی محیط رابط، بازیافت و دقت (در بسیاری از محیط‌های تعاملی، کاربران فقط به اسناد معدود با ربط و با مانعیت بالا نیاز دارند)، زمان لازم برای یادگیری کار با نظام، زمان لازم برای برای دستیابی به هدفها در وظایف معیار، نرخ خطا، و تداوم استفاده از محیط رابط در طی زمان و... توجه می‌شود.

در مدل استاندارد، فرایند دسترسی به اطلاعات، این واقعت مورد توجه قرار نمی‌گیرد که بسیاری از کاربران دوست ندارند با سیاهه‌ای طولانی و به هم ریخته از نتایج بازیابی (که به طور مستقیم به نیاز اطلاعاتی‌شان نمی‌پردازد) مواجه شوند. همچنین، حاوی این فرض اساسی است که نیاز اطلاعاتی کاربر، ایستاست و فرایند اطلاع‌جویی یک فرایند اصلاح‌پایی است، تا جایی که فقط اسناد مربوط به نیازهای اطلاعاتی اصلی، بازیابی شوند. تعاملی که در هنگام مرور واژه‌های پیشنهادشده به عنوان نتیجه بازخورد ربط، مرور ساختارهای اصطلاحنامه‌ای، یا مشاهده نمایه موضوعی گردآوردهای اسناد به وسیله کاربر رخ می‌دهد، در مدل استاندارد چندان مورد توجه

1. Information-seeking behavior.

2. Relevance feedback.

3. Jarvelin, K., & Ingwersen, P.

4. Berry-picking.

قرار نمی‌گیرد.

چنان‌که ملاحظه می‌شود، یکی از مسائل مهم در این نوع نظامها، بازخورد ربط است. بازخورد ربط در شکل اصلی آن به یک چرخه تعامل اطلاق می‌شود که در آن، کاربر مجموعه کوچکی از اسناد را که به نظر می‌رسد با پرسش ربط دارند، انتخاب می‌کند. سپس، نظام از ویژگیهای برگرفته شده از اسناد انتخاب شده مرتب، برای بازنگری در پرسش اصلی استفاده می‌کند. بعد، این پرسش بازنگری شده اجرا و مجموعه جدیدی از اسناد برگردانده می‌شود. اسناد مجموعه اصلی ممکن است در سیاهه یافته‌های جدید نیز ظاهر شوند، هرچند ممکن است رتبه متفاوتی داشته باشند. از جمله روشهای اعمال بازخورد ربط می‌توان به محیط رابط برای بازخورد استاندارد ربط (جعبه‌های علامتگذاری، اسناد دیگر شبیه به این، بازخورد تصاعدی)، قرار دادن اطلاعات با ربط در پس زمینه (کنترل پیشرفت و رفتار کاربران در دوره‌های تعاملی طولانی مدت)، قضاوت ربط گروهی (استفاده از قضاوت‌های ربط گروه بزرگی از کاربران مختلف، برای رتبه‌بندی یا درجه‌بندی اطلاعات عمومی)، و بازخورد شبه ربط اشاره کرد (در این روش، نظام به جای اتکا بر کاربر برای انتخاب K سند با ربط برتر، فرض می‌کند که اسناد رتبه‌بندی شده صورتی‌اش با ربط هستند، و از این اسناد برای طولانی کردن پرسش با یک الگوریتم رتبه‌بندی بازخورد ربط، استفاده می‌کند) (بائزا - بیتس و ریبرو - نتو، جلد دوم، ۱۳۸۹).

بررسی یک مدل ترکیبی

در پژوهش مورد بررسی (وانگ و فورگیونه، ۲۰۰۶) سعی شده است معیاری جامع و چند وجهی برای ارزیابی نظامهای بازیابی اطلاعات عرضه شود. مبنای این پژوهش بر فرایند تحلیل سلسله مراتبی^۱ در بافتار نظامهای بازیابی قرار گرفته است. بحث نو این پژوهش، تمرکز بر جنبه کاربری و کاربرد نظریه‌های تصمیم‌گیری^۲ در حوزه بازیابی

1. Analytic Hierarchy Process(AHP).
2. Decision making.

اطلاعات است.

این پژوهش فرایند جستجوی اطلاعات^۱ از منظر کاربر را در ۶ مرحله (شروع، گزینش، شناسایی، تدوین عبارت جستجو^۲، گردآوری و ارائه) مطابق مدل «کولثاو» (۱۹۹۱) در نظر گرفته و آن را با فرایند تصمیم‌گیری مطابقت داده است. وی معتقد است تفکر در مورد جستجوی اطلاعات به عنوان فرایند ساخت تصمیم به طور خاص در زمان ارزیابی نظامهای بازیابی اطلاعات، اهمیت ویژه‌ای دارد و از این راه می‌توان مشکل مرتبط با رویکردهای متفرق ارزیابی نظامها را هم در مورد برون‌داد و هم فرایند حل کرد. این تطابق در جدول ۱ تشریح شده است.

جدول ۱. مراحل تصمیم‌گیری در فرایند جستجوی اطلاعات (مبتنی بر مدل وانگ و فورگیونه، ۲۰۰۶)

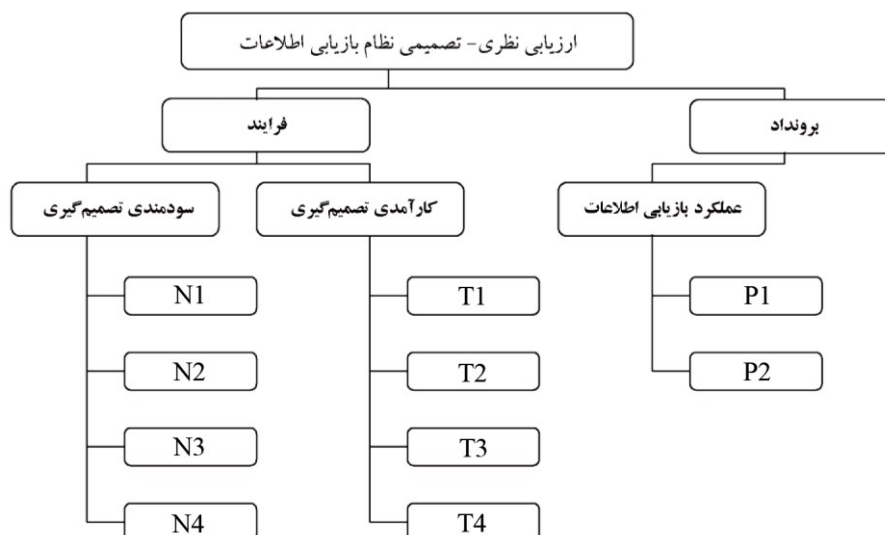
مراحل فرایند جستجوی اطلاعات	مراحل تصمیم‌گیری در فرایند جستجوی اطلاعات	مراحل عام تصمیم‌گیری
شروع به کار	بازشناسی مسئله	بازشناسی مسئله
انتخاب موضوع	ارزیابی جایگزینهای موضوع عام تشخیص یک موضوع عام	تشخیص عینیت‌ها و برقراری معیارها
شناسایی اولیه	جمع‌آوری اطلاعات مرتبط با موضوع عام	جمع‌آوری داده و تولید
	بررسی ارتباط بین موضوعات جایگزین و موضوع عام	جایگزینها
تمرکز بر فرمول‌بندی	برقراری معیارها (تمرکز)	تعیین روابط بین معیارها و جایگزینها
گردآوری اطلاعات	ارزیابی جایگزینهای مقاله	ارزیابی جایگزینها و توسعه

1. Information search process (ISP).
2. Formulation.

مراحل فرایند جستجوی اطلاعات	مراحل تصمیم‌گیری در فرایند جستجوی اطلاعات	مراحل عام تصمیم‌گیری
	تشخیص مقاله‌های مرتبط	اعمال توصیه شده در بیشترین اشتراک با معیارهای تصمیم
اتمام جستجو	اتخاذ انتخابهای نهایی	توسعه یک نقشه پیاده‌سازی و به‌کاربردن آن در عمل

بنابراین، مبانی این پژوهش را می‌توان به صورت زیر دسته‌بندی کرد. نخست، مطالعه‌ای است در مورد مراحل تصمیم‌گیری درگیر در فرایند جستجوی اطلاعات توسط کاربران و نقش نظریه‌های تصمیم‌گیری در ارزیابی نظامهای بازیابی اطلاعات. دوم، به تشریح مدلی نظری - تصمیمی که سنجه‌های رویکردهای برون‌دادی و فرایندی را ترکیب می‌کند، برای ارزیابی نظامهای بازیابی اطلاعات می‌پردازد. کاربرد معیار چند وجهی مبتنی بر روش تصمیم‌گیری که فرایند سلسله مراتبی تحلیل نامیده می‌شود، ارائه مدل ارزیابی و سپس کاربرد آن در حوزه خاص کاربرد بازیابی اطلاعات، قدمهای بعدی این پژوهش است. مدل ارائه شده مبتنی بر این رویکرد، مطابق شکل ۱ است.

به اعتقاد محققان طراح، این مدل سنجه‌های مختلف ارزیابی مرتبط به زمینه جستجوی اطلاعات را به شکل مجتمع به هم مرتبط می‌سازد. به دلیل اینکه یک نظام بازیابی اطلاعات بندرت در تمامی مراحل و رویه‌های ساخت تصمیم به کاربر یاری می‌رساند، تنها سنجه‌های مرتبط، در ارزیابی باید در نظر گرفته شوند. بعضی سنجه‌ها به فرایند و بعضی به برون‌داد می‌پردازند که به تفکیک بیان شده‌اند.



شکل ۱. مدل چند معیاره ارزیابی نظام بازیابی اطلاعات (وانگ و فورگیونه، ۲۰۰۶)

که در آن حروف ذکر شده (سنجه‌های خاص ارزیابی) نمایانگر موارد زیر هستند:

- T1 زمان صرف شده برای بازشناسی مسئله
- N1 تعداد جایگزینهای موضوعی عام
- T2 زمان صرف شده برای تشخیص موضوع عام
- N2 تعداد مقاله‌های تولید شده برای موضوع عام
- T3 زمان صرف شده برای ایجاد معیار
- T4 زمان صرف شده برای تشخیص مقاله مرتبط
- N3 تعداد مقاله‌های مرتبط تشخیص داده شده
- N4 تعداد مقاله‌های اضافی تشخیص داده شده در واریسی دوباره
- P1 سودمندی نظام
- P2 یادگیری کاربر

تشریح فرایند سلسله مراتبی تحلیل، به عنوان یکی از پر کاربردترین روشهای مورد

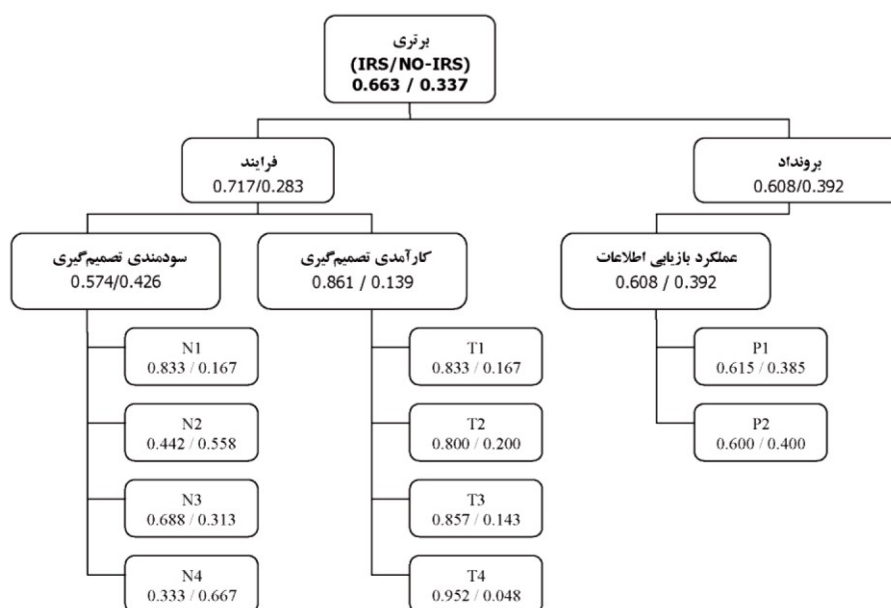
استفاده در پرداختن به مسائل دارای معیارهای چند وجهی تصمیم‌گیری^۱ (که ارزیابی نظام‌های بازیابی اطلاعات نیز از این دسته مسائل است) در بخش بعدی مورد توجه قرار گرفته و در نهایت، مدل ارائه شده را در یک نظام بازیابی اطلاعات (شبکه پیشگیری از خودکشی)^۲ به کار برده است و نتایج را در دو حالت (با استفاده یا بدون استفاده) از این نظام، مقایسه کرده است. با به کارگیری فرایند سلسله مراتبی تحلیل، این دو حالت در معیارهای فرایند و برون‌داد می‌توانند مقایسه شوند. لازم به ذکر است، برای تولید داده‌های مورد نیاز فرایند سلسله مراتبی تحلیل، از یک سناریوی شبیه‌سازی شده استفاده شده است. نتایج در جدول ۲ ارائه شده‌اند.

جدول ۲. ارزش سنج‌های ارزیابی در موارد با استفاده یا بدون استفاده از نظام بازیابی اطلاعات

بدون استفاده از نظام بازیابی اطلاعات	با استفاده از نظام بازیابی اطلاعات	سنج‌های ارزیابی
۱۵۰ ثانیه	۳۰ ثانیه	T1
۴	۲۰	N1
۶۰ ثانیه	۱۵ ثانیه	T2
۲۹	۲۳	N2
۳۶۰ ثانیه	۶۰ ثانیه	T3
۳۰۰ ثانیه	۱۵ ثانیه	T4
۵	۱۱	N3
۲	۱	N4
۵	۸	P1
۶	۷	P2

1. Multiple criteria decision making(MCDM).
2. Preventing Suicide Network(PSN).

ارزشهای عددی نتایج به دست آمده در جدول ۲، پس از تغییر به وزنهای مساوی به منظور سنجش برتری هرکدام از دو روش، در شکل ۲ آورده شده است. اگرچه شاید برخی معیارها مهم تر از بقیه باشند، به دلیل پرهیز از سوگیری اوزان یکسان در نظر گرفته شده‌اند.



نتایج ارائه شده در شکل ۲، نمایانگر معناداری نسبی تفاوت میان دو جایگزین در بهتر شدن ساخت تصمیم کاربر در خلال جستجوی اطلاعات است.

همچنین، نشان می‌دهد که سودمندی تصمیم‌گیری به میزان قابل توجهی با استفاده کاربر از نظام بازیابی اطلاعات، در مقایسه با زمانی که از آن استفاده نمی‌کند، بالا می‌رود (۰/۵۷۴ در مقابل ۰/۴۲۶) و کارآمدی تصمیم‌گیری به میزان بسیار زیاد با نظام بازیابی بهتر می‌شود (۰/۸۶۱ در مقابل ۰/۱۳۹). همچنین، عملکرد بازیابی اطلاعات نیز به طور معناداری با نظام بازیابی اطلاعات افزایش می‌یابد (۰/۶۰۸ در مقابل ۰/۳۹۲). این نظام در هر دو بخش معیارهای فرایندی و بروندادی نشان می‌دهد و بیشترین

برتری در فرایند است. روی هم رفته، امتیاز به دست آمده ۰/۶۶۳ برای نظام بازیابی اطلاعات در مقایسه با امتیاز ۰/۳۳۷ برای مورد استفاده نکردن از این نظام، نشان می‌دهد نظام شبکه پیشگیری از خودکشی (PSN) قدرت تصمیم‌گیری کاربر را در بازیافت اطلاعات مرتبط در مورد پیشگیری از خودکشی، افزایش می‌دهد.

محقق سپس به بحث و نتیجه‌گیری می‌پردازد. وی معتقد است این روش برای بازیافت اطلاعات حقوقی و کتابخانه‌ای نیز قابل استفاده است که می‌تواند در پژوهشهایی مورد سنجش قرار گیرد (وانگ و فورگیونه، ۲۰۰۶).

نتیجه و پیشنهاد

چنان‌که گفته شد، تحقیقات ارزیابی در زمینه بازیابی اطلاعات بر یک جنبه (رویکرد نظام مدار) عمدتاً متمرکز شده و در عین حال از هم گسیخته است. مطالعات ارزیابی در رویکردهای مختلف از هم جدا افتاده‌اند. به اعتقاد «ساراسویک» (۱۹۹۵) این جدا افتادگی در سطوح ارزیابی می‌تواند به عنوان کمبود بنیادی در تمامی ارزیابیهای بازیابی اطلاعات در نظر گرفته شود. نظامهای بازیابی به عنوان فعالیتی اجتماعی، با هدف پیوند و ممکن ساختن تعاملات میان تهیه‌کنندگان/ نویسندگان اطلاعات و کاربران/ خوانندگان اطلاعات، ایجاد می‌شوند (ساراسویک، ۱۹۹۵). این فعالیت، ماهیتی پیچیده، شناختی و ذهنی دارد (بورلند، ۲۰۰۳). بنابراین، ارزیابی یک نظام نه تنها در زمینه فنی و سیستمی، بلکه در زمینه‌های کاربردی باید سنجش شود. علاوه بر این، در ارزیابی باید ترکیبی از معیارهای بروندادی و فرایندی، برای انعکاس ماهیت پویا و چند بعدی ربط استفاده شود. چالش عمده در راه تلاش برای ارزیابی بازیابی اطلاعات، کاربرد مقیاسهای متنوع به روشی متحد شده و جامع است (ساراسویک، ۱۹۹۵) که مسیر پژوهشهای آینده را روشن می‌کند. به اعتقاد برخی از محققان، چالش آینده بر تشریح ابعاد چندگانه و اجزای نسل جدید نظامهای بازیابی اطلاعات از طریق آزمایش آنها با رویکردهای ارزیابانه متنوع، متمرکز است (رابرتسون و هنکاک بولیو، ۱۹۹۲). در این نوشته، پژوهشی مبتنی بر این رویکرد به عنوان نمونه معرفی شد که می‌تواند مسیر را روشن تر کند.

منابع

- پائو، میراندا لی (۱۳۷۸). **مفاهیم بازیابی اطلاعات**. ترجمه اسدالله آزاد و رحمت‌الله فتاحی. مشهد: دانشگاه فردوسی مشهد.
- بانزا - بیتس، ریکاردو؛ ریبرو - نتو، برتیه (۱۳۸۵). **قلمروهای نو در بازیابی اطلاعات**. جلد اول. ترجمه علی حسین قاسمی با همکاری سیروس آزادی و علی جوامع. تهران: چاپار؛ دبیزش.
- _____، (۱۳۸۹). **قلمروهای نو در بازیابی اطلاعات**. جلد دوم. ترجمه علی حسین قاسمی با همکاری سیروس آزادی و علی جوامع. تهران: چاپار؛ دبیزش.
- Belkin, N. J. (1980). Anomalous states of knowledge as a basis for information retrieval. **Canadian Journal of Information Science**, 5, 133-143.
 - Blair, D. C., & Maron, M. E. (1990). Full-text information retrieval: further analysis and clarification. **Information Processing & Management**, 26(2), 437-447.
 - Blair, D. C., & Maron, M. E. (1985). An evaluation of retrieval effectiveness for a full-text document retrieval system. **Communications of the ACM**, 28, 289-299.
 - Borlund, P. (2000). Experimental components for the evaluation of interactive information retrieval systems. **Journal of Documentation**, 56(1), 71-90.
 - Borlund, P. (2003). The concept of relevance in IR. *Journal of the American Society for Information Science and Technology*, 54(10), 913-925.
 - Cleverdon, Cyril (1997). The Cranfield Tests on Index Language Devices. In: Sparck-Jones, Karen; Willett, Peter (Eds.): **Readings in Information Retrieval**. Morgan Kaufman, 47-59.
 - Cleverdon, C.W., Mills, J., & Keen, E.M. (1966). **An inquiry in testing of information retrieval systems**. (2 vols.). Cranfield, U.K.:Aslib Cranfield Research Project, College of Aeronautics.
 - Ellis, D. (1996). The dilemma of measurement in information retrieval research. **Journal of the American Society for Information Science**, 47(1), 23-36.
 - Ingwersen, P. (1996). Cognitive perspectives of information retrieval interaction: Elements of a cognitive IR theory. *Journal of Documentation*, 52(1), 3-50.
 - Jarvelin, K., & Ingwersen, P. (2004). Information seeking research needs extension toward tasks and technology. **Information Research**, 10(1), paper no. 212. Available from <http://InformationR.net/ir/10-1/paper212.html>.
 - Kekalainen, J., & Jarvelin, K. (2002). Evaluating information retrieval systems under the challenges of interaction and multidimensional dynamic relevance. In: Bruce, H., Fidel, R., Ingwersen, P., Vakkari, P. (Eds.), **Proceedings of the 4th CoLIS Conference** (pp. 253-270). Greenwood Village, CP: Libraries Unlimited.
 - Kuhlthau, C. (1991). Inside the Search Process: Information Seeking from the User's Perspective. **Journal of the American Society for Information Science**, 42(5), 361-371.
 - Mandl, T. (2008). Recent Developments in the Evaluation of Information Retrieval Systems: Moving Towards Diversity and Practical Relevance. **Informatica**, 32, 27-38.
 - Robertson, S. E., & Hancock-Beaulieu, M. M. (1992). On the evaluation of IR systems. **Information Processing & Management**, 28(4), 457-466.
 - Salton, G. (1992). The state of retrieval system evaluation. **Information Processing & Management**, 28(4), 441-449.
 - Salton, G. (1989). **Automatic text processing: the transformation, analysis and retrieval of information by computer**. Reading, MA: Addison-Wesley.

- Salton, G. (1971). *The SMART retrieval systems: experiments in automatic document processing*. Englewood Cliffs, NJ: Prentice-Hall.
- Saracevic, T. (1995). Evaluation of Evaluation in Information Retrieval. **Proceedings of the Eighteenth Annual International ACM SIGIR Conference on Research and Development in Information Retrieval**, 138-146.
- Saracevic, T. (1996). Relevance reconsidered. In Ingwersen, P. and Pors, N.O. (Eds.), **Information Science: Integration in Perspective: Proceedings of the 2nd International Conference on Conceptions of Library and Information Science (CoLIS 2)**, 201-218.
- Schamber, L. (1994). Relevance and information behavior. **Annual review of information science and technology**, 29, 3- 48.
- Wang, Y.D., & Forgionne, G. (2006). A decision-theoretic approach to the evaluation of information retrieval systems. *Information Processing & Management*, 42, 863-874.