

سیستم پیشنهاد دهنده خبر آگاه بر مکان با استفاده از منطق فازی

مهدی نجاتی^(۱) - حمید طباطبایی^(۲) - مهرداد جلالی^(۳)

(۱) کارشناس ارشد - گروه مهندسی کامپیوتر، دانشکده مهندسی، واحد مشهد، دانشگاه آزاد اسلامی، مشهد، ایران

(۲) استادیار - باشگاه پژوهشگران جوان و نخبگان، واحد مشهد، دانشگاه آزاد اسلامی، مشهد، ایران

(۳) استادیار - گروه مهندسی کامپیوتر، دانشکده مهندسی، واحد مشهد، دانشگاه آزاد اسلامی، مشهد، ایران

تاریخ دریافت: ۱۳۹۴/۷/۲۴ تاریخ پذیرش: ۱۳۹۵/۵/۲۳

خلاصه: با انتشار حجم عظیمی از خبرها بر روی اینترنت و تمایل بیشتر کاربران به سرویس‌های خبری تحت وب، نیاز به یک سیستم پیشنهاد دهنده خبر ضروری است. خبرها برای آنکه قابل توجه باشند سرویس‌های خبری از تعدادی معیار به نام ارزش خبر استفاده می‌کنند و مکان کاربر نقش مؤثری در بررسی این معیارها دارد. در این مقاله، سیستم پیشنهاد دهنده خبر با ساختار دو مرحله‌ای به نام LONEF ارائه شده است. در مرحله اول با توجه به مکان کاربران، خبرها اولویت‌دهی می‌شوند و مرحله دوم با در نظر گرفتن خصوصیات همچون اولویت مکانی، تازگی، اعتبار منبع خبر، اولویت گروهی و محبوبیت، خبرها پیشنهاد داده می‌شوند. برای کاهش ابهامات این خصوصیات از دو سیستم در منطق فازی شامل استنتاج فازی ممدانی و تصمیم‌گیری مبتنی بر مورد استفاده شده است. در سیستم استنتاج فازی ممدانی سعی شده است با بهینه‌سازی در انتخاب قوانین و توابع عضویت سرعت سیستم بهبود داده شود و همچنین به علت پیاده‌سازی مبهم بازخوردهای کاربران در پیشنهاد دهنده‌های موجود یک سیستم تصمیم‌گیری بکار رفته است تا بتوان رفتار کاربران را بهتر بازنمایی کرد. میزان کارایی دیدگاه ارائه شده در این پژوهش از طریق نتایج آزمایش‌های مختلف بر روی مجموعه خبری نمایش داده شده است.

کلمات کلیدی: سیستم‌های پیشنهاد دهنده خبر، پیشنهاد دهنده خبر، منطق فازی، پیشنهاد دهنده آگاه بر مکان.

Location-Aware News Recommendation System with Using Fuzzy Logic

Mehdi Nejati⁽¹⁾ – Hamid Tabatabaei⁽²⁾ – Mehrdad Jalali⁽³⁾

(1) MSc – Department of Computer Engineering, Islamic Azad University, Mashhad Branch, Mashhad, Iran
mahdi.nejati@mshdiau.ac.ir

(2) Assistant Professor - young researcher club, Islamic Azad University, Mashhad Branch, Mashhad, Iran
h_Tabatabaei@mshdiau.ac.ir

(3) Assistant Professor - Department of Computer Engineering, Islamic Azad University, Mashhad Branch, Mashhad, Iran
mehرداد.jalali@mshdiau.ac.ir

with release of a huge amount of news on the Internet and the trend of users to Web-based news services. it is necessary to have a recommendation system. To grab attentions to news, news services use a number of criteria that called news values and user location is an important factor for it. In this paper, LONEF is proposed as a tow stage recommendation system. In first stage news are ranked by user's locations and in second stage news are recommended by location Preferences, recency, Trustworthiness, groups priorities and popularity. To reduce ambiguity these properties is used tow Mamdani fuzzy interference and case-based decision systems. In Mamdani fuzzy interference system, it is tried to increase the system speed by optimizing selection of rules and membership functions and because of ambiguous feedback implementation, a decision making system is used to enable better simulation of user's activities. Performance of our proposed approach is demonstrated in the experiments on different news groups.

Index Terms: News Recommendation systems, News Recommendation, fuzzy logic ,location-aware Recommendation

نویسنده مسئول: حمید طباطبایی - استادیار - باشگاه پژوهشگران جوان و نخبگان، واحد مشهد، دانشگاه آزاد اسلامی، مشهد، ایران،
h_Tabatabaei@mshdiau.ac.ir

۱- مقدمه

هر روزه اطلاعات بسیار زیادی ایجاد و روی وب قرار می‌گیرند. برای دسترسی کاربران به اطلاعات دلخواه خود سیستم‌های پیشنهاددهنده با تکنیک‌های مختلف ایجاد شدند تا کاربران بتوانند با کمک این سیستم‌ها به اطلاعاتی مطلوب خود دست پیدا کنند. سیستم‌های پیشنهاد دهنده زیر مجموعه سیستم‌های صافی‌سازی اطلاعات هستند که بر اساس امتیازدهی یا اولویت‌بندی کاربر پیشنهاداتی ارائه می‌دهند. یکی از دلایل اصلی پیدایش سیستم‌های پیشنهاد دهنده، افزایش سود فراهم کنندگان تجارت الکترونیک در وب می‌باشد تا به کاربران محصولاتی را پیشنهاد دهنده که احتمال خرید آنها بالا بوده است. این سیستم‌ها بر اساس روش ایجاد پیشنهاد به سه گروه مبتنی بر محتوا، مشارکتی و ترکیبی دسته‌بندی شده است [۱].

پیشنهاد دهنده خبر یکی از مهم‌ترین ابزارها برای فراهم‌کنندگان محتوا مانند گوگل نیوز و یاهو نیوز می‌باشد. این سرویس‌ها با مشکل ازدحام اطلاعات سروکار دارند که مانع از دستیابی خوانندگان به اطلاعات مورد نظرشان می‌شود. راه حل این مشکل سیستم‌های پیشنهاد دهنده هستند که می‌توانند به طور فزاینده‌ای دستیابی به اطلاعات دقیق و کارآمد را بهبود بخشند.

در حوزه خبرنگاری هر خبر توسط هفت معیار ارزیابی می‌شود که به آنها ارزش خبر می‌گویند [۲]. برای طراحی یک سیستم پیشنهاد دهنده خبر کارآمد تا حد امکان باید این معیارها در نظر گرفته شوند. سه معیار شهرت، فراگیری و مجاورت خبر وابستگی زیادی به محل زندگی کاربر دارند. از این رو مکان یا محل زندگی کاربر به عنوان یک شاخص موثر در سیستم‌های خبری باید در نظر گرفته شود. از طرفی دیگر سرویس‌های خبری با منابع خبر بسیاری سرکار دارند. به عنوان نمونه گوگل نیوز، خبرهای خود را از بین ۵۰۰۰ منبع خبری با ۷۲ ویرایش در ۳۰ زبان جمع‌آوری می‌کند [۳]. برای تعیین اعتبار این منابع خبری در سیستم‌های خبری باید از کمیت‌هایی استفاده شود.

پیشنهاد دهنده خبر همچون سیستم‌های پیشنهاد دهنده دیگر نیازمند داده‌هایی از رفتار کاربران و ویژگی‌های آیت‌ها می‌باشد و کارایی آنها وابسته به چگونگی استنباط و تفسیر این داده‌ها است. در این سیستم‌ها، پارامترهای ورودی تازگی، متن و اعتبار منبع خبری و پروفایل کاربر می‌باشد [۴]. کارایی این سیستم‌ها به بازنمایی و تفسیر این ورودی‌ها وابسته می‌باشد که در نمایش و استدلال و ارتباط بین آنها عدم قطعیت و ابهام وجود دارد. برای نمونه، تعیین موثرترین منبع خبری یا بازه زمانی تازگی خبر با ابهام همراه است. منطق فازی یک بینش قوی از روش‌ها برای مدیریت عدم قطعیت‌های ارائه می‌دهد [۵].

ما یک سیستم پیشنهاد دهنده خبری LONEF^۱ ارائه می‌دهیم که از دو ماژول برای حل این مشکلات استفاده می‌کند. در ماژول اول اولویت‌های کاربران نه تنها بر اساس تاریخچه دسترسی خود آنها ایجاد می‌شود بلکه همچنین اولویت‌ها و تاریخچه خبرهای بازدیدی گروه‌هایی از کاربران مورد بررسی قرار می‌گیرد که کاربر به آن گروه تعلق دارد. این گروه‌ها به صورت یک ساختار سلسله مراتبی بر اساس

محل زندگی کاربران ایجاد می‌شوند و کاربر بیشتر خبرهایی اهمیت می‌دهد که نزدیک محل زندگی او باشد [۲]. در مرحله دوم LONEF با توجه به عدم قطعیت‌ها در پارامترهای موجود و به دست آمده از مرحله قبل شامل اولویت مکانی، تازگی، اعتبار منبع خبر، اولویت گروهی و محبوبیت از منطق فازی استفاده می‌کند. برای هر پارامتر کمیت‌های زبانی و مجموعه سیستم‌های صافی‌سازی اطلاعات هستند. عضویت با توجه به کارایی سیستم انتخاب شده است. برای اعمال منطق فازی از دو سیستم شامل استنتاج فازی مددانی و پشتیبانی تصمیم استفاده می‌کند. هدف کلی ما از این سیستم نمایش خبرهای مورد توجه کاربر می‌باشد که برای آن باید دقت در انتخاب خبرهای مورد علاقه کاربر بهبود یابد و از طرف دیگر با داشتن تنوع مطلوب در خبرها مانع از کاهش جلب توجه کاربر شود.

در ادامه مقاله بدین صورت تقسیم‌بندی شده است: بخش دوم کارهای انجام گرفته در زمینه سیستم خبری مورد بررسی قرار می‌دهیم. بخش سوم ساختار سیستم LONEF معرفی می‌کنیم. بخش چهارم نتایج ارزیابی را نشان می‌دهیم و در آخر نتیجه‌گیری ارائه می‌کنیم.

۲- کارهای مرتبط

پیاده‌سازی‌های مختلفی از پیشنهاد دهنده‌های خبر ارائه شده است و همچنین در چند سال اخیر ویژگی مکانی کاربر و منطق فازی در پیشنهاد دهنده‌ها مورد توجه قرار گرفته‌اند.

۲-۱- مبتنی بر محتوا

این پیشنهاد دهنده‌ها به کلی بر مبنای محتوای خبر هستند که شباهت‌های بین پروفایل کاربر و محتوای متنی خبر مورد بررسی قرار می‌دهند. هم پروفایل کاربری و هم محتوای خبر عموماً با استفاده از مدل فضای برداری به عنوان مثال TF-IDF [۶] یا توزیع‌های موضوع توسط مدل‌های زبان PLSI [۷] و LDA [۸] تعیین می‌شوند. بسیاری از سیستم‌های پیشنهاددهنده خبری مبتنی بر محتوا در دهه اخیر پیشنهاد شده‌اند. برای مثال New Dude [۹] یک عامل پیشنهاد دهنده خبر شخصی که TF-IDF با الگوریتم نزدیکترین همسایگی ترکیب می‌کند تا آیت‌های خبری به کاربران خاص پیشنهاد دهد.

۲-۲- پیشنهاد دهنده ترکیبی

چندین سیستم ترکیبی برای پیشنهاد اخبار توسعه داده شده‌اند که برخی از آنها عبارتند از SCENE [۱۰] و PENETRATE [۱۱] اولین مورد یک سیستم دو مرحله‌ای و شخصی شده با قابلیت مقیاس‌پذیری می‌باشد و به صورت یک ساختار سلسله مراتبی دو سطحی می‌توان در نظر گرفته شود که اولین سطح شامل گروه‌های خبری (بر اساس موضوع) است و دومین سطح شامل خبرهای خاص در گروه انتخابی کاربر است. سیستم PENETRATE کاربران را گروه‌بندی می‌کند بر اساس مقالات خبری که توسط آنها خوانده شده است و هر کاربر می‌تواند به بیشتر از یک گروه تعلق پیدا کند. سپس پیشنهاددهنده سیستم از گروه‌های متعلق به کاربر تولید می‌شود. دیدگاه دیگر MONERS [۱۲] است که اولویت‌های کاربر، توسط یک سری اطلاعات آماری از اهمیت، تازگی و الگوهای دسترسی به خبر

۳) از سیستم پشتیبانی تصمیم استفاده کرده‌ایم تا علاوه بر بکارگیری بازخورد کاربران در نتایج، ابهام در استنتاج و شباهت بین گروه‌های کاربری کاهش یابد.

۳- ساختار کلی سیستم

ساختار کلی سیستم پیشنهادی ما LONEF در شکل (۲) نمایش داده شده است. داده‌های ورودی سیستم خبر و اطلاعات کاربر می‌باشند. در LONEF شباهت بین پروفایل کاربر و خبرها و همچنین معیارهای خبر، شامل تازگی و موثق بودن بررسی می‌گردد. سپس خروجی آن به ماژول منطق فازی ارسال و به کمک استنتاج فازی ممدانی عدم قطعیت‌ها در خروجی تنظیم می‌شود. مکان کاربر توسط سیستم دریافت و به ماژول بخش‌بندی سلسله مراتبی ارسال می‌شود. در این ماژول گروه کاربری مرتبط با مکان کاربر یافته و از طریق آن مدل مکانی از خبرهای مشابه برای کاربر تهیه می‌شود. مدل مکانی به ماژول فازی ارسال می‌شود تا تئوری سیستم فازی به دست آید و در نهایت ماژول منطق فازی لیستی از خبرهای پیشنهادی برای کاربر تهیه می‌کند. در ادامه هر ماژول را به طور کامل شرح می‌دهیم.

۳-۱- ماژول بخش‌بندی مکانی

ماژول بخش‌بندی از مکان کاربر استفاده می‌کند تا معیار نزدیکی به عنوان یک پارامتر در تولید پیشنهاد نقش مؤثری داشته باشد. کاربران بر اساس موقعیت مکانی مانند شهر یا کشور محل زندگی خود به چندین گروه یا بخش تقسیم می‌شوند و با توجه به این گروه‌ها، خبرها بر اساس تعلق که به این گروه‌ها دارند در بخش‌هایی با ساختار سلسله مراتبی تقسیم می‌شوند. در اینجا ساختار سلسله مراتبی همان چاردرخت ناحیه‌ای با تفاوت جزئی است که ناحیه‌ها روی نقشه تقسیم‌بندی‌های جغرافیایی مختلفی همچون شهر، استان، ایالت، کشور، قاره می‌تواند باشد و اولویت مکانی هر سطح بالاتر مجموع اولویت‌های مکانی چهار بخش سطح پایین می‌باشد. در بالاترین سطح تنها یک بخش وجود دارد یعنی اولویت‌های مکانی خبرها یکسان است. هر بخش مدلی از اولویت‌های مکانی می‌سازد که شامل محبوب‌ترین خبرها یک مکان می‌باشد. روند ساخت مدل مکانی در ماژول فازی توضیح خواهیم داد.

۳-۱-۱ پیاده سازی بخش‌بندی مکانی

در شکل (۱) نمای کلی از این ماژول آمده است. پیاده‌سازی این ماژول شبیه چاردرخت ناحیه‌ای است که در آن امتیازهای کاربران بر اساس مکان کاربر به ناحیه‌های مکانی به صورت سلسله‌مراتبی تقسیم می‌شود. ما از یک هرم جزئی و جدول برای پیاده‌سازی استفاده می‌کنیم [۲۰] که در آن سلسله مراتب شامل H سطح می‌شود و تعداد سلول‌ها در هر سطح برابر شماره سطح به توانی از چهار می‌باشد. هر سطح ممکن است معرف یک تقسیم‌بندی جغرافیایی مانند سطح استان‌ها و سطح پایین‌تر آن شهرهای درون استان باشد. حداکثر تعداد ناحیه جغرافیایی در هر سطح برابر تعداد سلول‌های آن سطح می‌باشد و جدول نگاشتی بین هرم و مکان کاربر ایجاد می‌کند.

جمع‌آوری و تخمین زده می‌شود. در سیستم Wesomender [۴]، خبر از نگاه حوزه خبری مورد بررسی قرار گرفته است که با در نظر گرفتن پارمترهای پروفایل کاربر، فاصله فیزیکی خبر با مکان کاربر و اعتمادپذیر منبع خبر و ترکیب آنها با یک ساختار انطباقی از چندین الگوریتم مشارکتی، خبرها اولویت‌دهی و پیشنهاد می‌شوند.

۲-۳- منطق فازی در سیستم پیشنهاد دهنده

منطق فازی پایه تئوری مناسبی برای توسعه و گسترش سیستم‌های پیشنهاد دهنده فراهم می‌کند. منبع اصلی اطلاعات برای استفاده در این سیستم‌ها دانش درباره آیت‌ها است. سودمندی این اطلاعات وابسته به نمایش درست و دقیق از آیت‌ها می‌باشد. در ابتدا چارچوب‌های مفهومی و پایه‌ای برای نمایش و استنتاج منطق فازی در سیستم‌های پیشنهاد دهنده توسط [۱۳، ۱۴] و معرفی شد که در آنها روابط فازی بیان‌کننده تاثیر آیت‌ها بر کاربر می‌باشد و درجه تعلق هر آیت‌ها به ویژگی‌ها و میزان تاثیر این ویژگی‌ها بر علائق کاربر در بستر منطق فازی بیان شد. علائق کاربری در [۱۵] به صورت تمایل چهارگانه از درجه تعلق مثبت و منفی تعریف شده است. تعدادی از پژوهش‌ها سعی کرده‌اند از منطق فازی برای سیستم‌های پیشنهاد دهنده در حوزه‌های مختلف همچون تجارت الکترونیکی، کتابخانه‌های دیجیتال و سایت‌های خبری استفاده کنند و دیدگاه کلی آنها دسترسی سریع‌تر کاربران به آیت‌های مورد نظرشان با افزایش دقت می‌باشد. در ادامه چند مورد آورده شده است. در مرجع [۱۶] کتابخانه دیجیتال با سیستم زبانی فازی و ساختار سلسله مراتبی نمایش داده است که در آن نیازهای کاربر و ویژگی‌های منابع کتابخانه با متغیر زبانی فازی مشخص شده‌اند. اولویت‌بندی سلسله مراتبی ارتباط بین کاربر و آیت مشخص می‌کند که پیشنهادها از وزن‌دهی آنها به دست می‌آید. در [۱۷] نیازهای کاربر و خصوصیات آیت‌ها با عدد فازی مثلثی برای یک فروشگاه بررسی شده است. با استفاده از محاسبه شباهت توسط معیار فاصله فازی آیت‌ها انتخاب می‌شوند. در [۱۸] پیاده‌سازی عملی از تئوری منطق فازی برای سیستم پیشنهاددهنده روی مجموعه از آیت‌های فیلم ارائه شده است و با استفاده از آزمایش‌های مختلف میزان تاثیر این منطق را نمایش می‌دهد.

در [۱۹] از منطق فازی برای یک سیستم خبری استفاده شده است. نویسنده سیستم استنتاج فازی ممدانی را به کار می‌برد. ورودی‌های این سیستم میزان شباهت پروفایل کاربری و تازگی خبر می‌باشد و با یک متغیر زبانی شامل سه عدد فازی گوسی تعریف شده‌اند. برای خروجی سیستم توابع عضویت مثلثی در نظر گرفته شده است.

مقاله [۱۹] نزدیکترین کار نسبت به سیستم LONEF می‌باشد اما سیستم ما تفاوتی اساسی با آن دارد:

(۱) چون سرویس‌های خبری با تعداد منابع خبری زیادی سرکار دارند ما از یک پارامتر میزان اعتمادپذیری منبع خبر استفاده کرده‌ایم.

(۲) از آنجایی که خبرهای نزدیک کاربر بیشتر مورد توجه آنها می‌باشد ما با یک ساختار گروه‌بندی از مکان کاربر به عنوان عامل موثر در فرآیند پیشنهاد استفاده کرده‌ایم.

۳-۲-۱- سیستم استنتاج فازی ممدانی

یک سیستم استنتاج فازی به کمک منطق فازی، نگاهی از ورودی به خروجی را فرموله می‌کند. این سیستم از تعدادی عبارت «اگر-آنگاه» ساخته شده است. در هر سیستم استنتاج فازی مهمترین قسمت قواعد فازی (موتور استنتاج فازی)، توابع عضویت و روش غیرفازی سازی است. توابع عضویت مختلفی از قبیل مثلثی، گوسی، دوزنقه، شبه گوسی وجود دارد. ممدانی و سوگونو، دو نوع سیستم استنتاج فازی پرکاربرد موجود هستند.

برای فازی سازی ورودی و خروجی با توجه به نوع داده (تا حدی وابسته به زمان) و مرجع [۱۹]، کاهش پیچیدگی بدون تاثیر بر معیار دقت پیش بینی از تابع عضویت دوزنقه‌ای شکل به کار می‌رود.

در این جا برای ویژگی داده‌های سیستم پیشنهاد دهنده خبر متغیرهای فازی و توابع عضویت آنها را باید تعریف کنیم که این داده‌ها همان خبرها و کاربران هستند. روابط یا ویژگی‌ها مورد بررسی به شرح زیر می‌باشند و تمام آنها به صورت نرمال شده هستند:

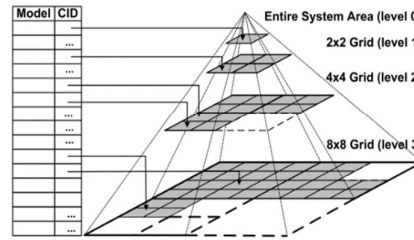
شباهت موضوعی: میزان ارتباط خبر جدید یا دیده نشده توسط کاربر و پروفایل کاربری را تعیین می‌کند. به علت تمرکز بیشتر موضوع خبر در تیتراژ آن می‌آید برای نمایش خبرها و پروفایل کاربری ما برداری از کلمات کلیدی درون تیتراژ آنها استفاده می‌کنیم. پروفایل کاربری شامل تمام خبرهایی می‌باشد که کاربر دیده است. کلمات کلیدی خبر را با کلمات کلیدی پروفایل کاربری مقایسه و درصد ارتباط آن را تعیین می‌کند. متغیر زبانی آن را TS و متغیرهای فازی با اصطلاحات خیلی کم، کم، زیاد، خیلی زیاد تعریف می‌کنیم. در شکل (۳) توابع عضویت این متغیر را نمایش داده‌ایم.

تازگی خبر: معرف فاصله زمانی بین انتشار خبر و زمان کنونی می‌باشد و ما از نرمال شده‌ی آن استفاده می‌کنیم. متغیر زبانی، R در نظر می‌گیریم و برای فاصله زمانی برچسبها به صورت یکسان، نزدیک، دور، خیلی دور تعریف می‌شوند. شکل (۴) توابع عضویت آن را نمایش می‌دهد.

باور پذیری: معرف میزان اعتماد و اعتبار منبع خبر می‌باشد. ما از اطلاعات تراست نیوز برای بررسی این پارامتر استفاده می‌کنیم. شکل (۵) متغیر زبانی TW با مقادیر کاملاً قابل اعتماد، قابل اعتماد، نسبتاً قابل اعتماد، غیر قابل اعتماد نمایش می‌دهد.

خروجی: در سیستم پیشنهاد دهنده خروجی میزان اولویت یا امتیازی هر خبر می‌باشد. بر این اساس ما امتیازها را به صورت کم، خیلی کم، زیاد و خیلی زیاد برای سیستم فازی تعریف کرده‌ایم و نام آن را NR و به صورت شکل (۶) نمایش می‌دهیم.

حال قوانین فازی را بین ورودی و خروجی تعریف می‌کنیم. برای استنتاج در حالت کلی امکان ایجاد ۲۵۶ قانون می‌باشد که از بین آنها تنها ۲۱۳ مورد کارایی لازم را دارند. در جدول (۱) تعدادی از این قوانین آمده است. خروجی حاصل از قوانین فازی را توسط تابع ترکیبی



شکل (۱): ساختار هرمی بخش بندی مکان

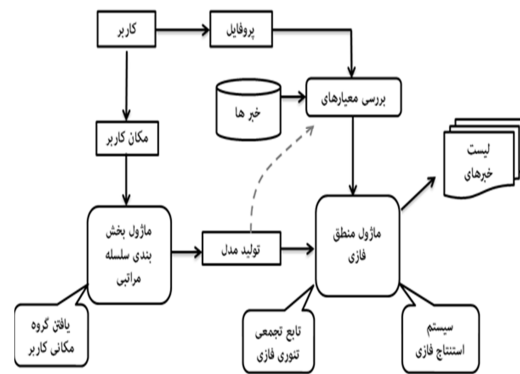
Fig. (1): Location segmentation pyramid structure

۳-۲-۲- ماژول منطق فازی

پیشنهاد دهنده خبر همچون سیستم‌های پیشنهاد دهنده دیگر نیازمند داده‌هایی از رفتار کاربران و ویژگی‌های آیتم‌ها می‌باشند و کارایی آنها وابسته به چگونگی استنباط و تفسیر این داده‌ها است. نحوه نمایش و استنتاج درباره رفتار کاربران و ویژگی‌های آیتم‌ها مسائلی چالش برانگیز می‌باشد.

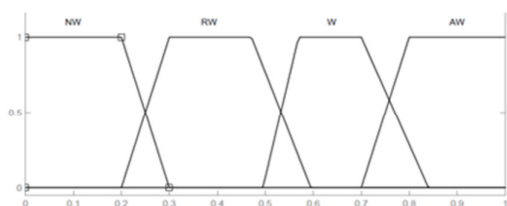
ویژگی‌های آیتم و رفتار کاربر موضوعاتی مهم و نامعین هستند و باعث وجود عدم قطعیت در نمایش و ارتباط بین آنها می‌شود. چنین عدم قطعیت غیر تصادفی است و از دامنه دانش ناشی می‌شود. به طور کلی عدم قطعیت در سیستم‌های پیشنهاددهنده به روابط موجود در آن وابسته است. که این روابط شامل ارتباط بین رفتار کاربر و ویژگی‌های آیتم، بین کاربران و بین آیتم‌ها می‌باشد.

منطق فازی برای مدیریت عدم قطعیت‌های، نمایش دقیق اطلاعات و بهبود نتایج سیستم‌ها ارائه شده است. در منطق فازی تمام ورودی‌ها و اطلاعات به صورت مجموعه فازی تعریف می‌شوند. پس در ابتدا اعداد فازی، برای ورودی‌ها باید تعریف شوند. مرحله بعد، تکنیک‌های اعمال منطق فازی تعریف می‌شوند که در اینجا از سیستم‌های تصمیم‌گیری فازی و استنتاج ممدانی استفاده شده است و در مرحله آخر نتایج فازی باید به مقادیر قطعی تبدیل شوند. این ماژول از دو تکنیک مجزا تشکیل شده است که لیست خبرهای پیشنهادی نتیجه مجموع وزن دار آنها می‌باشد.

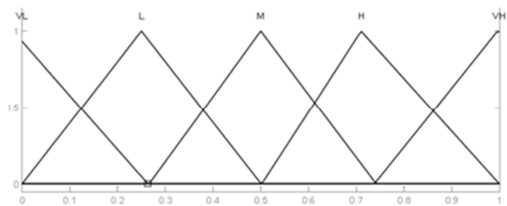


شکل (۲): نمای کلی سیستم LONEF

Fig. (2): Fuzzy variable of topic similarity (TS)



شکل (۵): متغیر فازی باورپذیری TW
Fig. (5): Fuzzy variable of trustworthiness (TW)



شکل (۶): متغیر فازی امتیاز خبر NR
Fig. (6): Fuzzy variable of rate

در نتیجه اولویت به عنوان یک عدد فازی می‌توان در نظر گرفت [۱۸]. عدد فازی مورد نظر یک نیمه مثلثی بین بیشترین و کمترین مقدار سیستم امتیازدهی قرار می‌گیرد و با تابع عضویت زیر مشخص می‌کنیم.

$$\mu_A(I_i) = \frac{r - \text{Min}}{\text{Max} - \text{Min}} \quad (2)$$

پس از تعیین تابع عضویت باید میزان شباهت بین آیت‌ها به دست آید. معیارهای شباهت مختلفی در مجموعه فازی معرفی شده است. ما از معیار شباهت معرفی شده در مرجع [۲۲] استفاده می‌کنیم. بر اساس آنکه دو آیت I_j یا خبر I_k دو مجموعه فازی $I_j = \{(x_i, \mu_{x_i}(I_j))\}$ و $I_k = \{(x_i, \mu_{x_i}(I_k))\}$ و $i = 1, \dots, N$ و N تعداد کاربران سیستم در نظر گرفته ایم این معیار بین دو آیت به صورت زیر تعریف می‌شود.

$$S(I_k, I_j) = \frac{\sum_i \min(\mu_{x_i}(I_k), \mu_{x_i}(I_j))}{\sum_i \max(\mu_{x_i}(I_k), \mu_{x_i}(I_j))} \quad (3)$$

از روش تجمعی حداکثر-حداقل برای به دست آوردن امتیاز آیت‌های پیشنهادی استفاده می‌کنیم.

$$LR = \max_k \{ \min(S(I_j, I_k), \mu_E(I_k)) \} \quad (4)$$

لیست خبرهای پیشنهادی از مجموع خروجی‌ها به دست می‌آید:

$$R = w * VR + (1 - w) * LR \quad (5)$$

۴- ارزیابی

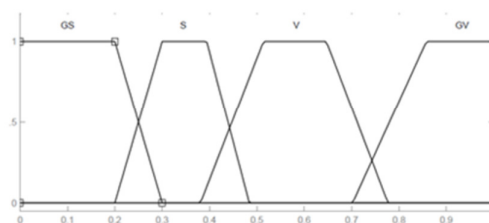
برای ارزیابی از کل سیستم پیشنهادی اطلاعات زمان انتشار خبر، زمان دسترسی کاربر و موقعیت مکانی کاربر نیاز است و ۳۲۱۰۰ خبر تحت شرایط و نیاز فراهم می‌کنیم و کلمات کلیدی آنها به کمک ابزار zemanta استخراج شده‌اند و خبرهای مرتبط و غیر مرتبط مشخص هستند. تعداد ۳۴۰۰ کاربر در نظر گرفته شده است و جدول بین کاربر و خبر شامل خبرهای بازدیدی کاربر، امتیاز آنها و زمان مشاهده خبر می‌باشد. ما ۸۰ درصد داده‌ها را مجموعه یادگیری و ۲۰ درصد باقی

حداقل ایجاد می‌کنیم و خروجی فازی از عملگر تجمعی حداکثر حاصل می‌شود.

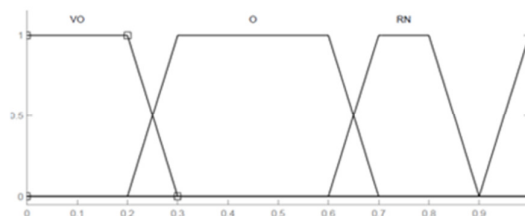
۳-۲- بازخورد و مدل مکانی فازی

مدل مکانی شامل یک لیست از آیت‌های مشابه می‌باشد و برای تولید مدل از الگوریتم‌های شباهت بین آیت‌ها استفاده می‌شود. در این روش نگاشتی از مفاهیم فازی به مجموعه‌ای از برچسب‌های مجزا استفاده شده است تا فرآیند امتیازدهی کاربر با کمترین ابهام همراه باشد. همانطور که از قبل بیان شد بازخورد و امتیازدهی کاربر با بررسی مقادیر قطعی در یک مجموعه اعداد همچون $S = \{1, 2, \dots, 5\}$ ابهاماتی با خود دارد. مثلاً آیتی که کاربر شدیداً علاقمند است با مقدار ۵ در سیستم پیشنهاد دهنده ذخیره می‌شود. ما فرض می‌کنیم که شدیداً علاقمند توسط یک مجموعه فازی $\{0/1, 0/2, 0/3, 0.5/4, 1/5\}$ شرح داده می‌شود که این مجموعه مانع از آن است که درجه رضایت کاربر توسط مقدار ۴ بتواند بیشتر یا کمتر توصیف شود. در ابتدا ما امتیاز کاربر یا اولویت‌های کاربر به عنوان مشاهدات فازی از مجموعه S را بررسی می‌کنیم. یعنی هر امتیاز به عنوان یک زیرمجموعه فازی از S در نظر گرفته می‌شود. متغیر زبانی DI را با متغیرهای فازی شامل شدیداً علاقمند، علاقمند، بی تفاوت، متنفر، شدیداً متنفر تعریف می‌کنیم و به صورت مجموعه فازی $S_L = \{sd, d, i, l, sl\}$ برای S نمایش می‌دهیم [۲۱، ۱۸].

روش ما چون به صورت یک چهارچوب احتمال می‌باشد. نیاز به یک توزیع احتمال روی این متغیرهای فازی داریم. بر اساس مفهوم متغیر فازی، امتیاز کاربر به وسیله یک تابع توزیع احتمال بازنمایی و قابل تفسیر می‌باشد



شکل (۳): متغیر فازی باورپذیری TW
Fig. (3): Fuzzy variable of topic similarity (TS)



شکل (۴): متغیر فازی تازگی خبر R
Fig. (4): Fuzzy variable of news recency (R)

غیر مرتبط با کاربر دارد. در نهایت معیار F-measure برای LONEF بهترین نتیجه را نسبت به دیگر روش‌ها به دنبال دارد.

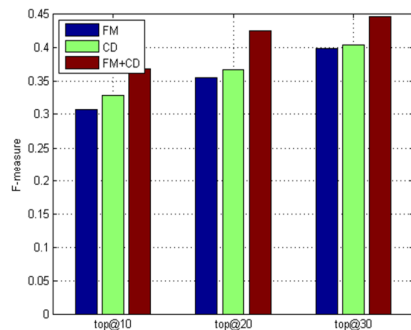
۴-۲- ارزیابی ماژول منطق فازی

ما از دو سیستم فازی شامل استنتاج فازی ممدانی و تصمیم‌گیری مبتنی بر مورد استفاده کرده‌ایم. به منظور ارزیابی کارایی دیدگاه ترکیبی این دو سیستم، تاثیر هر کدام را به صورت جدا آزمایش می‌کنیم. در صورتی که سیستم تنها با سیستم تصمیم‌گیری پیاده‌سازی شود (CD). خروجی سیستم استنتاج فازی ممدانی را با حاصل ضرب ویژگی‌ها به صورت غیرفازی در نظر می‌گیریم و جایگزین تصمیم‌گیری مبتنی بر مورد از روش مشارکتی با شباهت کسینوسی استفاده می‌کنیم (FM). شرایط این آزمایش همانند قبلی می‌باشد.

شکل (۷) نتایج آزمایش را نمایش می‌دهد و آنچه که از این آزمایش به دست آمده است شامل: (۱) دیدگاه ترکیبی ما همیشه کارایی بهتری دارد. (۲) لیست آیتم‌های پیشنهادی مبتنی بر یکی از سیستم‌های فازی نمی‌تواند بهتر از ترکیب آنها شود و علت آن کاملاً مشخص است: چون ابهام بین تمام روابط یک سیستم پیشنهاد وجود دارد و هر کدام باعث کنترل بخشی از ابهامات درون سیستم می‌شود. (۳) تاثیر هر دو سیستم روی کارایی به یک اندازه می‌باشد چون هر کدام روی روابط مستقلی استدلال می‌کنند.

Table (1): A number of Fuzzy inference rules
جدول (۱): تعدادی از قوانین استنتاج

۱	If TS is GV and TW is AW and R is N then NR is VH
۲	If TS is V and TW is RW and R is O then NR is L
۳	If TS is V and TW is W and R is RN then NR is H
۴	If TS is S and TW is AW and R is RN then NR is M
۵	If TS is GS and TW is RW and R is VO then NR is VL



شکل (۷): مقایسه F-measure روی سیستم‌های مختلف فازی. FM: سیستم

استنتاج فازی ممدانی، CD: سیستم تصمیم‌گیری مبتنی بر مورد

Fig. (7): Compression F-measure of fuzzy systems. FM: Mamdani Fuzzy Inference System, CD: Case-based decision system

مانده مجموعه امتحان استفاده می‌کنیم. معیارهای ارزیابی Precision، Recall و F-measure می‌باشند.

در کل سیستم LONEF از دو ماژول اصلی تشکیل شده است. ماژول اول با نام FM که تنظیم و اعمال عدم قطعیت روی پارامترها سیستم همچون ویژگی‌های خبر و امتیازهای کاربر بر عهده دارد و ماژول دوم با نام LM مکانیزم یافت خبرهای مرتبط با مکان کاربر می‌باشد.

کار اصلی سیستم‌های پیشنهاد دهنده قرار دادن آیتم‌های کاملاً مرتبط و مناسب در اول لیست پیشنهاد می‌باشد و علت آن توجه کاربران به آیتم‌های اول لیست است. در روند آزمایش از ۳ سطح برای بخش بندی ماژول LM استفاده می‌کنیم.

۴-۱- ارزیابی کیفیت سیستم

در این قسمت ما مقایسه‌ای بین LONEF و LM و دو سیستم دیگر Wesomender [۴]، FCB [۱۹] انجام می‌دهیم. LM معرف پیاده‌سازی سیستم، تنها با ماژول LM می‌باشد که ما خروجی ماژول FM را با حاصل ضرب مقادیر خروجی و فیلترینگ مشارکتی با شباهت کسینوسی به دست می‌آوریم.

۴-۱-۱- تغییر اندازه لیست پیشنهادی

تعداد اولین یا تاپ خبرهای لیست پیشنهاد این آزمایش به ترتیب ۲۰، ۳۰ و ۴۰ در نظر گرفته می‌شود (@Top) و از تعداد کاربران موجود به صورت تصادفی ۱۰۰ کاربر انتخاب می‌شوند. نتایج آزمایش بر روی معیارهای مختلف در جدول (۱) آمده است. بررسی معیار اول دقت Precision نشان می‌دهد که کمترین دقت برای FCB حاصل شده است چون که در این دیدگاه تعداد کمتری ویژگی برای پیش‌بینی نتایج استفاده می‌شود. در مقابل LONEF نسبت به روش‌های دیگر بهترین نتیجه را به همراه دارد زیرا علاوه بر این که ماژول بخش بندی مکانی موجب تاثیر خاص روی انتخاب آیتم‌های نزدیک به کاربر می‌شود استفاده از منطق فازی باعث کاهش ابهام در بازنامی اطلاعات و افزایش خبرهای مورد علاقه کاربر می‌شود. اگر میزان تغییرات دقت با افزایش لیست پیشنهاد مورد بررسی قرار دهیم. دیدگاهی که از منطق فازی استفاده می‌کنند اختلاف تغییرات کمتری دارند به طوری که بر روی اندازه ۳۰ مقدار دقت wesomender از روش FCB کمتر می‌شود. در بررسی معیار Recall برخلاف معیار دقت میزان اختلاف بین LONEF و LM ناچیز است که دلیل آن، منطق فازی تاثیر چندانی روی بازیابی آیتم‌ها نمی‌گذارد و به خوبی در روش FCB مشخص می‌باشد. با افزایش لیست پیشنهادها مشخص است شیب تغییرات در LONEF و FCB کمتر از دو روش دیگر است. علت آنکه LONEF بیشترین Recall را دارد استفاده از ماژول بخش‌بندی مکانی است، چون مدل مکانی در این ماژول سعی در محدودسازی خبرهای

Table (2): Comparison among different recommenders

Method	Top @10			Top@20			Top@30		
	Precision	Recall	F-measure	Precision	Recall	F-measure	Precision	Recall	F-measure
LONEF	0.4590	0.3068	0.3677	0.4362	0.4138	0.4247	0.3989	0.5029	0.4449
LM	0.3932	0.2721	0.3216	0.3228	0.3782	0.3435	0.2851	0.4425	۰.۳۴۶۸
Wesomender	0.3605	0.2587	0.3012	0.2831	0.3248	0.3025	0.2203	0.3998	0.2841
FCB	0.3182	0.2043	0.2488	0.2765	0.2579	0.2669	0.2461	0.3078	0.2735

Table (3): Diversity evaluation on the result list

جدول (۳): ارزیابی تنوع روی لیست نتایج

Top@20	Top@15	Top@10	روش ها
0/2056	0/3074	0/4101	FCB
0/3502	0/4126	0/5037	Wesomender
0/6504	0/6845	0/7203	LONEF

۵- نتیجه گیری و کارهای آینده

در این مقاله، سیستم پیشنهاد دهنده خبر LONEF ارائه داده شده است که در آن معیار مجاورت یا ارتباط خبر به مکان کاربر با کمک روش بخش‌بندی مکانی به صورت سلسله مراتبی مورد توجه قرار گرفت تا کیفیت سیستم بهبود یابد. همچنین منطق فازی برای بهبود دقت به عنوان مهمترین معیار در سیستم‌های پیشنهاد دهنده به کار رفت. با توجه به نتایج آزمایش میزان دقت سیستم تا حد مطلوبی بهبود یافته است و همچنین به دلیل حساس نبودن منطق فازی به تغییرات اندازه حجم داده کارایی کلی سیستم حفظ شده است.

برای کارهای آینده، تغییرات رفتار کاربر یا رانگی و چهارچوب‌های موازی سازی مانند نگاشت-کاهش و اسپارک استفاده همزمان از سطوح مختلف بخش بندی می‌توان بررسی کرد.

پی‌نوشت:

1. Location-aware News Recommendation System with Using Fuzzy Logic
2. <http://www.zemanta.com/>

۳-۴- ارزیابی گوناگونی

تنوع یا گوناگونی خبرهای پیشنهادی توسط LONEF در سطح بسیار مطلوبی می‌باشد. علت آن استفاده از دو ماژول بخش بندی مکانی و منطق فازی است. در بخش بندی مکانی لیست علایق از کاربران مختلف بدست می‌آید و به کمک منطق فازی نزدیک‌ترین لیست با بیشترین تنوع بر اساس بازنمایی بهبود یافته علاقه کاربری حاصل می‌شود. برای ارزیابی چگونگی گوناگونی در لیست خبرهای سیستم، ما از روش بیان شده در مقاله [۲۳] گوناگونی بین نتایج LONEF و سیستم‌های دیگر را مقایسه می‌کنیم. گوناگونی مجموعه خبری به عنوان متوسط تفاوت همه جفت آیتم‌های خبرهای در لیست پیشنهادی تعریف شده است. مخصوصاً، با یک مجموعه معلوم N از خبرها، متوسط تفاوت $f_d(N)$ به صورت زیر تعریف می‌شود:

$$f_d(N) = \frac{2}{p(p-1)} \sum_{n_i \in N} \sum_{n_j \in N, n_j \neq n_i} (1 - \text{Sim}(n_j - n_i)) \quad (۶)$$

که $|N| = p$ و تفاوت بین جفت خبر توسط $1 - \text{Sim}(n_j - n_i)$ نمایش داده شده است. که شباهت بین دو آیتم خبری را $\text{Sim}(n_j - n_i)$ نشان می‌دهد. برای ارزیابی تنوع یا گوناگونی، ما LONEF را با FCB و Wesomender مورد مقایسه قرار دادیم. نتایج را در جدول (۳) آورده‌ایم.

از نتایج مشاهده می‌کنیم که (۱) تنوع با افزایش لیست پیشنهادها کاهش می‌یابد. این میزان در روش‌های دیگر بیشتر است به علت آنکه تنوع در آنها در نظر گرفته نشده است. (۲) نتیجه LONEF بسیار مشهودتر است از آنجایی که ما از کاربران یک محل استفاده می‌کنیم.

References

- [1] M. Balabanovic, Y. Shoham, "Content-based, collaborative recommendation," in Communications of the ACM, 1997.
- [2] O. pencer-Thomas, "Writing a press release.," 2012. [Online]. Available: <http://www.owenspencer-thomas.com/journalism/media-tips/writing-a-press-release>. [Accessed 8 August 2015].
- [3] F. Filloux, "Google news: The secret sauce," The guardian, 25 February 2013. [Online]. Available: <http://www.theguardian.com/technology/2013/feb/25/1>. [Accessed 18 8 2015].
- [4] A. Montes-García, J.M. Álvarez-Rodríguez, J.E. Labra-Gayo, M. Martínez-Merino, "Towards a journalist-based news recommendation system: The Wesomender approach," Expert Systems with Applications, Vol. 40, pp. 6735-6741, 2013.
- [5] T. Bilgiç, I. Turksen, "Measurement of membership functions: theoretical and empirical work," In Fundamentals of Fuzzy Sets: Handbook of Fuzzy Sets and Systems, Boston, Dubois, D, p. 195-232, 2000.

- [6] D. Jurafsky, J. Martin, *speech and language processing*, new jersey: prenticce Hall, 2014.
- [7] T. Hofmann, "Probabilistic latent semantic indexing," In Proceedings of the 22nd annual international ACM SIGIR conference on research and development in information retrieval. SIGIR '99, New York, NY, USA: ACM., 1999.
- [8] D.M. Blei, A. Ng, M.I. & Jordan, "Latent Dirichlet allocation," *Journal of Machine Learning Research*, Vol. 3, p. 993-1022, 2003.
- [9] D. Billsus, M. Pazzani, "A personal news agent that talks, learns and explains," *Proceedings of the third annual conference on Autonomous Agents*, pp. 268-275, 1999.
- [10] L. Li, D. Wang, T. Li, D. Knox, B. Padmanabhan, "SCENE: a scalable two-stage personalized news recommendation system.," *Sigir*, pp. 125-134, 2011.
- [11] L. Zheng, L. Li, W. Hong, T. Li, "PENETRATE: Personalized news recommendation using ensemble hierarchical clustering," *Expert Systems with Applications*, Vol. 40, pp. 2127-2136, 2013.
- [12] H.J. Lee, S.J. Park, "MONERS: A news recommender for the mobile web," *Expert Systems with Applications*, Vol. 32, No. 1, pp. 143-150, 2007.
- [13] R.R. Yager, "Fuzzy logic methods in recommender systems," *Fuzzy Sets and Systems*, Vol. 136, No. 2, pp. 133-149, 2003.
- [14] P. Perny, J.-d. Zucker, "Collaborative filtering methods based on fuzzy preference relations," In Proceedings of EUROFUSE-SIC, 1999.
- [15] C. Cornelis, J. Lu, X. Guo, G. Zhang, "One-and-only item recommendation with fuzzy logic techniques," *Information Sciences*, Vol. 177, No. 22, pp. 4906-4921, 2007.
- [16] A. Tejada-Lorente, J. Bernabé-Moreno, C. Porcel, E. Herrera-Viedma, "Integrating quality criteria in a fuzzy linguistic recommender system for digital libraries," *Procedia Computer Science*, Vol. 31, pp. 1036-1043, 2014.
- [17] B.A. Ojokoh, M.O. Omisore, O.W. Samuel, T.O. Ogunniyi, "A fuzzy logic based personalized recommender system," In *International Journal of Computer Science and Information Technology & Security (IJCSITS)*, 2012.
- [18] A. Zenebe, A.F. Norcio, "Representation, similarity measures and aggregation methods using fuzzy sets for content-based recommender systems," *Fuzzy Sets and Systems*, Vol. 160, No. 1, pp. 76-94, 2009.
- [19] M.N.M. Adnan, M.R. Chowdury, I. Taz, T. Ahmed, R.M. Rahman, "Content based news recommendation system based on fuzzy logic," In *2014 International Conference on Informatics, Electronics & Vision (ICIEV)*, 2014.
- [20] J.J. Levandoski, M. Sarwat, A. Eldawy, M.F. Mokbel, "LARS: A location-aware recommender system," In *Proceedings - International Conference on Data Engineering*, 2012.
- [21] L.M. de Campos, J.M. Fernández-Luna, J.F. Huete, "A collaborative recommender system based on probabilistic inference from fuzzy observations," *Fuzzy Sets and Systems*, Vol. 159, No. 12, pp. 1554-1576, 2008.
- [22] C.P. Pappis, N.I. Karacapilidis, "A comparative assessment of measures of similarity of fuzzy values," *Fuzzy Sets and Systems*, pp. 171-174, 1993.
- [23] Z.M, H.N, "Avoiding monotony: improving the diversity of recommendation lists," In *Proceedings of the ACM Conference on Recommender Systems*, 2008.
- [24] L. Zadeh, "Fuzzy Sets," *Inform.Control* 8, p. 338-353, 1965.
- [25] L. Li, W. Chu, "A contextual-bandit approach to personalized news article recommendation," In *WWW* , 2010.