Journal of Soft Computing and Information Technology (JSCIT)
Babol Noshirvani University of Technology, Babol, Iran
Journal Homepage: www.jscit.nit.ac.ir
DOI:
Volume 9, Number 2, Summer 2020, pp. 1-15
Received: 03/26/2019, Revised: 10/02/2019, Accepted: 10/23/2019

A Novel Similarity Measure for Fuzzy Recommender Systems.

Mansoore Shojaei, Hassan Saneifar[2*]
1- Faculty of Mechanics, Electrical Power and Computer, Science and research Branch, Islamic Azad University, Tehran, Iran.
2*- Department of Computer Engineering, Raja University of Qazvin, Qazvin, Iran.
[1]mansooreh.shojaee@gmail.com, [2*]hassan.saneifar@raja.ac.ir
Corresponding author address: Hassan Saneifar, Department of Computer Engineering, Raja University of Qazvin, Qazvin, Iran, Post Code: 34148-95834.

Abstract- Due to the development of the internet, there is today explosive growth and diversity of information. Thus, decision making in various fields has faced different challenges. Recommender systems by identifying of the interests of users, data filtering and data management, offer personalized services to users. This is beneficial for marketing and user satisfaction. According to researchers, the Collaborative Filtering (CF) is one of the most successful methods of recommender system. CF based on the similarity between users. We argue that similarity is a fuzzy notion and we get more realistic results in recommender systems by using fuzzy logic. fuzzy logic is kind of new technologies that is based on the multi-value logic. This technology can replace complicated calculations and advanced mathematics in designing and modeling of a system. fuzzy logic is an effective method to identify ambiguities and uncertainty in measuring the similarity of items and users. In this research, we present a new fuzzy similarity measure for recommender systems which is based on popularity and significance. In this work, in order to improve the accuracy and quality of recommendations, we have used a fuzzy logic. To evaluate the contribution of this work, we use MAE, F1, recall, and precision. The MAE value based on the proposed fuzzy similarity measure is equal to 0.46 and outperforms the NHSM respectively by %5. Also, using the proposed fuzzy similarity measure, we obtain a F1 value equal to 0.65, which outperforms the PIP and NHSM respectively by %17 and %20. We have also observed an improvement in recall and precision using the proposed approach.

Keywords- Recommender System, Collaborative Filtering, Similarity Measure, Fuzzy Set, Criteria FSR.

مجله علمی پژوهشی رایانش نرم و فناوری اطلاعات
دانشگاه صنعتی نوشیروانی بابل
صفحه مجله: www.jscit.nit.ac.ir
شناسگر اشیاء دیجیتالی (DOI) :
جلد ۹، شماره ۲، تابستان ۱۳۹۹ ، صفحه ۱ - ۱۵
دریافت: ۱۳۹۸/۰۱/۰۶، بازنگری: ۱۳۹۸/۰۷/۱۰، پذیرش: ۱۳۹۸/۰۸/۰۱

ارائه یک معیار جدید اندازه‌گیری شباهت در سیستم‌های توصیه‌گر فازی

منصوره شجاعی، حسن صانعی‌فر*[۲]
۱- دانشکده مکانیک، برق و کامپیوتر، دانشگاه علوم و تحقیقات، تهران، ایران.
*۲- گروه مهندسی کامپیوتر، دانشگاه رجا قزوین ، قزوین، ایران
[1]mansooreh.shojaee@gmail.com, [2*]hassan.saneifar@raja.ac.ir

* نشانی نویسنده مسئول: حسن صانعی‌فر، قزوین، دانشگاه رجا قزوین، گروه مهندسی کامپیوتر، کد پستی: 34148-95834

چکیده- با توسعه اینترنت، بشر امروزی با رشد انفجاری و تنوع اطلاعات مواجه است و این امر تصمیم‌گیری در زمینه های مختلف را با چالش مواجه کرده است. بدین منظور سیستم‌های توصیه‌گر با شناسایی علایق کاربران، فیلترکردن داده‌ها و مدیریت اطلاعات، باعث صرفه جویی در زمان، شخصی‌سازی تجارت الکترونیک، بهبود بازاریابی و غیره شده است. یکی از موفق‌ترین روش‌های سیستم‌های توصیه‌گر، فیلترینگ همکاری است که مبتنی بر شباهت بین کاربران می‌باشد. سیستم‌های توصیه‌گر همواره با چالش‌هایی از جمله، مشکل شروع سرد، مشکل پراکندگی، مقیاس پذیری، دقت و کیفیت مواجه بوده اند. در سیستم‌های توصیه‌گر، دقت و کیفیت توصیه‌ها اهمیت بسزایی دارند. روش ارائه شده در این پژوهش بر روی دقت و کیفیت سیستم‌های توصیه‌گر متمرکز بوده است. راهکارهای مختلفی مانند استفاده از منطق فازی برای بهبود دقت و کیفیت توصیه‌ها، ارائه شده است. استفاده از منطق فازی، روشی موثر برای برخورد با عدم قطعیت در اندازه‌گیری شباهت آیتم‌ها و کاربران است. در این پژوهش، روشی جدیدی برای اندازه‌گیری شباهت فازی به نام FSR براساس فاکتورهای محبوبیت و اهمیت، در سیستم‌های توصیه‌گر فیلترینگ همکاری ارائه شده است. در روش پیشنهادی، از اعداد فازی برای بهبود دقت و کیفیت توصیه‌ها استفاده شده است. روش پیشنهاد شده، توسط معیارهای MAE، F1، فراخوانی و صحت ارزیابی شده است. مقدار F1 در حالت استفاده از معیار شباهت پیشنهادی نسبت به روش‌های مبتنی بر معیارهای PIP و NHSM به ترتیب ۱۷ و ۲۰ درصد افزایش یافته است. همچنین معیارهای فراخوانی و صحت بدست آمده از این روش نسبت به دو روش مذکور نیز بهبود قابل توجهی یافته‌اند. نتایج نشان می‌دهد روش پیشنهادی FSR نسبت به روش‌های اندازه‌گیری شباهت ارائه شده در سال‌های اخیر مانند PIP و NHSM و همچنین روش‌های قدیمی مانند MSD عملکرد بهتری دارد.

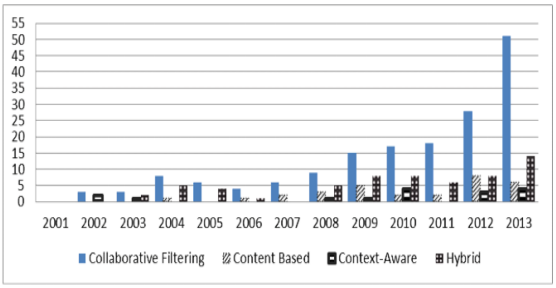
واژه‌های کلیدی: سیستم‌های توصیه‌گر، فیلترینگ همکاری، اندازه‌گیری شباهت، مجموعه فازی، معیار FSR

۱- مقدمه

امروزه بیشتر مردم از تلفن همراه، رایانه و سایر دستگاه‌های هوشمند استفاده می‌کنند. آنها زمان زیادی را در شبکه‌های اجتماعی (مانند فیس‌بوک[۱] و توییتر[۲]) و یا در سایت‌های تجارت الکترونیک (مانند آمازون[۳] و eBay[۴]) می‌گذرانند. در این شبکه‌ها و سایت‌ها، اطلاعات و محصولات بسیاری در دسترس کاربران قرار گرفته است، بنابراین آنها می‌بایست زمان زیادی را برای یافتن اطلاعات و محصولات مورد نظر خود صرف کنند[۱]. در واقع کاربران، هر روز، با حجم فزاینده و رو به رشدی از اطلاعات مواجه هستند که نیازمند ظرفیت پردازشی بزرگی است. هنگامی که حجم اطلاعات برای تصمیم گیری در مورد یک موضوع بیش از حد باشد، پدیده ی " سربار اطلاعاتی[۵] " رخ می‌دهد[۲]. توسعه اینترنت، رشد انفجاری و تنوع اطلاعات موجود، کاربران را در تصمیم‌گیری‌های خود ضعیف کرده است و در دسترس بودن انتخاب‌ها، به جای سودمند بودن، رفاه کاربران را کاهش داده است. برای حل مشکل سربار اطلاعاتی استفاده از سیستم‌های توصیه‌گر با موفقیت‌هایی همراه بوده است و استفاده از آن در سال های اخیر رشد خوبی داشته است[۳].

در واقع سیستم‌های توصیه‌گر به منظور راهنمایی و هدایت افراد برای رسیدن به گزینه‌های مفید و مورد علاقه آنها در میان حجم عظیمی از انتخاب‌ها به کار گرفته می‌شود. هدف اصلی سیستم‌های توصیه‌گر شناسایی علایق کاربران، فیلترکردن داده‌های موجود و در نهایت پیشنهاد گزینه‌های مفید و مورد علاقه به کاربر است. استفاده از سیستم‌های توصیه‌گر مزایای بسیاری دارد مانند: صرفه‌جویی در زمان کاربران در هنگام جستجوی اطلاعات، مهیا نمودن شخصی‌سازی برای مشتریان تجارت الکترونیک، بهبود بازاریابی و بهبود تجربه کاربر در تعامل اجتماعی با دیگر کاربران[۴].

میزان استفاده از انواع روش‌ها در میان سال های ۲۰۰۱ تا ۲۰۱۳ در شکل ۱ نشان داده شده است. همانطور که در شکل دیده می‌شود، در سال‌های اخیر استفاده از روش توصیه مبتنی بر فیلترینگ همکاری، افزایش قابل ملاحظه‌ای نسبت به روش های دیگر داشته است[۱۰].



شکل ۱ : نمودار استفاده از سیستم‌های توصیه‌گر در سال های ۲۰۰۱ تا ۲۰۱۳، به تفکیک نوع الگوریتم[۱۰]

سیستم‌های توصیه‌گر همواره با چالش‌هایی مانند شروع سرد، پراکندگی، مقیاس پذیری، دقت و کیفیت توصیه‌ها مواجه بوده‌اند. روش‌های کلاسیک فیلترینگ همکاری مبتنی بر معیارهای شباهت مانند روش پیرسون و روش کسینوسی، منجر به مشکل شروع سرد شده است و کیفیت توصیه‌ها را هنگامی که سوابق امتیاز کمی در دسترس است را کاهش می‌دهد. به منظور رفع مشکل فوق، هیونگ جون آون[۶] ، روش جدید اندازه‌گیری شباهت برای بهبود عملکرد توصیه در شرایط شروع سرد ارائه کرده است[۵]. مشکل پراکندگی[۷] ، ضعف اصلی سیستم فیلترینگ همکاری است. مشکل پراکندگی به طور ذاتی در این سیستم رخ می‌دهد و به نسبت افزایش تعداد کاربران و آیتم‌ها، افزایش می‌یابد. از این رو، سارایاکانت[۸] و همکاران، روش اندازه‌گیری شباهت جدیدی براساس میانگین اندازه‌گیری واگرایی[۹] ارائه داده‌اند تا مشکل پراکندگی، موجود در سیستم‌های مبتنی بر فیلترینگ همکاری را حل کنند[۶]. همچنین، پولاتیدس و جرجیدیس[۱۰] [۷،۸]، تانر آرسان و همکاران[۱۱] [۹]، برای بهبود دقت و کیفیت توصیه‌ها روش‌هایی را ارائه کرده‌اند.

در این مقاله، برای بهبود توصیه‌ها در سیستم‌های توصیه‌گر فیلترینگ همکاری یک معیار جدید شباهت به نام FSR[۱۲] با استفاده از پارامترهایی مانند اهمیت[۱۳] ، محبوبیت[۱۴] و روش میانگین اختلاف مربع[۱۵] ارائه شده است. در این معیار جدید شباهت همچنین از منطق فازی برای بهبود نتایج و افزایش دقت و کیفیت سیستم‌های توصیه‌گر فیلترینگ همکاری استفاده شده است. نتایج نشان می‌دهد روش پیشنهادی FSR نسبت به روش‌های اندازه‌گیری شباهت ارائه شده در سال‌های اخیر مانند PIP و NHSM و همچنین روش‌های قدیمی مانند MSD عملکرد بهتری دارد.

ادامه مقاله به صورت زیر تنظیم شده است: در بخش۲، سابقه تحقیق و برخی از مهمترین روش های اندازه‌گیری شباهت بیان می‌گردد. در بخش۳، روش پیشنهادی اندازه‌گیری شباهت شرح داده می‌شود. سپس در بخش۴، نتایج بدست‌آمده بررسی شده و در نهایت نتیجه‌گیری و کارهای آینده ارائه می‌گردد.

۲- تحقیقات پیشین

سیستم توصیه‌گر، سیستمی است که انتخاب کاربران را براساس بیان ترجیحات کاربر، رفتارگذشته و داده‌های دیگر محدود می‌کند[۱۱]. سیستم‌های توصیه‌گر[۱۶] ثابت کرده‌اند که می‌توانند با کاهش اطلاعات زائد به کاربران کمک کنند[۳]. روش‌ها و الگوریتم‌های مختلفی برای سیستم‌های توصیه‌گر ارائه شده است. سیستم‌های توصیه‌گر، براساس نوع الگوریتم به صورت زیر دسته‌بندی می‌شوند:

۱- روش توصیه مبتنی بر محتوا[۱۷]

۲- روش توصیه مبتنی بر فیلترینگ همکاری[۱۸]

۳- روش توصیه مبتنی بر دانش[۱۹]

۴- روش توصیه ترکیبی[۲۰] [۷،۱۲،۱۳]

یکی از موفق‌ترین روش‌ها، فیلترینگ همکاری است که محصولات را براساس شباهت کاربران توصیه می‌کند. در این پژوهش تمرکز بر روی سیستم‌های توصیه‌گر فیلترینگ همکاری می‌باشد. در سیستم‌های توصیه‌گر فیلترینگ همکاری، از تابع شباهت استفاده شده است که در آن توصیه‌ها براساس امتیاز کاربران است. هسته سیستم‌های توصیه‌گر فیلترینگ همکاری مبتنی بر شباهت بین کاربران یا آیتم‌ها است. سیستم‌های توصیه‌گر فیلترینگ همکاری، شامل روش‌های مبتنی بر حافظه و مبتنی بر مدل می‌باشد. در روش مبتنی بر حافظه، ابتدا شباهت‌های بین کاربران محاسبه می‌شود و سپس کاربرانی که بیشترین شباهت را با کاربر فعال دارند به عنوان همسایه‌های کاربر فعال انتخاب می‌کند و در نهایت، توصیه براساس همسایگان کاربر فعال انجام می‌شود. در روش مبتنی بر مدل، ابتدا

مدلی را برای توصیف رفتار کاربران ارائه می‌دهند و سپس پیش بینی امتیاز آیتم را انجام می‌دهند. در روش مبتنی بر حافظه توصیه‌ها با دقت قابل توجهی، ارائه می‌شود، در حالی که زمان محاسبات با افزایش تعداد کاربران و آیتم‌ها به سرعت افزایش می‌یابد، اما به علت ساخت مدل در زمان خارج از خط در روش مبتنی بر مدل، زمان پیش‌بینی نسبت به روش مبتنی بر حافظه، کاهش می‌یابد. در سیستم‌های توصیه‌گر فیلترینگ همکاری، دقت توصیه‌های پیش‌بینی شده، موضوع مهم و با اهمیتی است[۸].

سیستم‌های توصیه‌گر در ارائه توصیه‌ها با چالش‌هایی مواجه هستند. مشکل شروع سرد[۲۱] ، مشکل پراکندگی[۲۲] ، مشکل مقیاس پذیری[۲۳] ، دقت و کیفیت توصیه‌های ارائه شده تعدادی از چالش‌های پیش رو سیستم‌های توصیه‌گر هستند. در سال‌های اخیر، پژوهش‌های زیادی در جهت بهبود و افزایش دقت و کیفیت توصیه‌ها به کاربر انجام شده است. اطمینان از کیفیت توصیه‌ها برای کسب اعتماد مشتری الزامی است. چنانچه مشتریان از یک سیستم توصیه‌گر پیروی کنند، اما نتایج آن مورد پسند آنها نباشد، بعید به نظر می‌رسد که دوباره از این سیستم استفاده کنند[۴]. در سیستم‌های توصیه‌گر شباهت بین آیتم‌ها و شباهت بین کاربران به طور طبیعی فازی است و تلاش‌های زیادی برای استفاده از تئوری منطق فازی و روابط فازی در سیستم‌های توصیه‌گر صورت گرفته است. برای حل مشکلات راهکارهای متفاوتی ارایه شده است. یک راه‌حل، استفاده از مجموعه فازی است. مجموعه فازی در حل مشکل مقیاس پذیری موفق نبوده است ولی برای مشکل شروع سرد و مشکل پراکندگی عملکرد قابل قبولی داشته است[۱۱].

در سیستم‌های فیلترینگ همکاری، روش‌های اندازه‌گیری شباهت متفاوتی وجود دارد. تعدادی از روش‌های کلاسیک اندازه گیری شباهت عبارتند از : ضریب همبستگی پیرسون[۲۴] ، شباهت کسینوسی[۲۵] ، روش Jaccard و روش میانگین اختلاف مربع(MSD)[۱]. همچنین، در سال‌های اخیر روش‌های جدید اندازه‌گیری شباهت مانند PIP [۵] و NHSM [۱] برای بهبود نتایج و رفع چالش‌های سیستم‌های توصیه‌گر ارائه شده است.

در ادامه، برای شرح چگونگی عملکرد معیارهای شباهت و مقایسه بهتر، از یک نمونه کوچک از جدول امتیاز کاربران به آیتم‌ها در سیستم‌های توصیه‌گر به عنوان مثال استفاده می‌شود. سپس، مقدار شباهت بدست آمده در هر روش براساس مثال ذکر شده محاسبه شده و اشکالات این روش‌ها با استفاده از مثال مطرح شده بیان می‌گردد.

۲-۱- تعریف مثال

در این مثال فرض کردیم تعداد کاربرها برابر ۵ و تعداد آیتم ها برابر ۴ باشد. ماتریس کاربر-آیتم به صورت جدول ۱ تعریف شده است. سلول‌های این ماتریس، امتیاز کاربران (امتیازها در بازه [۵-۱] قرار دارد) برای آیتم‌ها را نشان می‌دهد. رفتار کاربرها نسبت به یکدیگر نیز در جدول ۲ نشان داده شده است.

جدول ۱ : ماتریس امتیاز کاربر-آیتم

	Item1	Item2	Item3	Item4
User1	۴	۳	۵	۴
User2	۵	۳	-	-
User3	۴	۳	۳	۴
User4	۲	۱	-	-
User5	۴	۲	-	-

همانطور که در جدول ۲ مشاهده می‌شود، به دلیل داشتن ۳ امتیاز برابر در آیتم‌ها، کاربر۱ و کاربر۳ بیشترین شباهت را در بین کاربران دارند. با مقایسه زوج (کاربر۱ و کاربر۲) و زوج (کاربر۱ و کاربر۵) مشاهده می‌شود که در آیتم اول دارای امتیاز برابر و در آیتم دوم ۱ امتیاز اختلاف دارند همچنین زوج (کاربر۲ و کاربر۳) و زوج (کاربر۳ و کاربر۵) در آیتم دوم دارای امتیاز یکسان و در آیتم اول دارای ۱ امتیاز اختلاف هستند. اختلاف این ۲ حالت، تنها در شمارش امتیاز آیتم است. در این حالت شباهت زوج (کاربر۱و کاربر۲) و زوج (کاربر۱ و کاربر۵) از شباهت زوج(کاربر۲ و کاربر۳) و زوج (کاربر۳ و کاربر۵) بیشتر است. همانطور که در جدول ۲ دیده می‌شود، کمترین شباهت بین کاربر۲ و کاربر۴ است. بدین دلیل که در آیتم اول ۳ امتیاز اختلاف و در آیتم دوم ۲ امتیاز اختلاف وجود دارد.

جدول ۲ : رفتار کاربرها نسبت به یکدیگر

کاربر ۱ با کاربر ۲	U1=(4,3,5,4) – U2=(5,3,-,-)
کاربر ۱ با کاربر ۳	U1=(4,3,5,4) – U3=(4,3,3,4)
کاربر ۱ با کاربر ۴	U1=(4,3,5,4) – U4=(4,2,-,-)
کاربر ۱ با کاربر ۵	U1=(4,3,5,4) – U5=(5,3,-,-)
کاربر ۲ با کاربر ۳	U2=(5,3,-,-) – U3=(4,3,3,4)
کاربر ۲ با کاربر ۴	U2=(5,3,-,-) – U4=(4,2,-,-)
کاربر ۲ با کاربر ۵	U2=(5,3,-,-) – U5=(5,3,-,-)
کاربر ۳ با کاربر ۴	U3=(4,3,3,4) – U4=(4,2,-,-)
کاربر ۳ با کاربر ۵	U3=(4,3,3,4) – U5=(5,3,-,-)
کاربر ۴ با کاربر ۵	U4=(4,2,-,-) – U5=(5,3,-,-)

در ادامه به معرفی تعدادی از اصلی ترین معیارهای شباهت مورد استفاده در سیستم‌های توصیه‌گر مبتنی بر فیلترینگ همکاری و بررسی کارایی آنها با استفاده از مثال ذکر شده می‌پردازیم.

۲-۲- معیار شباهت میانگین اختلاف مربع(MSD)

معیار شباهت MSD براساس میانگین اختلاف مربع امتیاز کاربرها بدست می‌آید. معیار شباهت *MSD* به صورت زیر تعریف می‌شود[۹]:

$$sim(u,v)^{MSD} = 1 - \frac{\sum_{p\epsilon I}(r_{u.p} - r_{v.p})^2}{|I|} \qquad (۱)$$

که $|I|$ مربع مجموعه امتیاز مشترک بین دو کاربر است[۱].
در *MSD* تنها امتیازهای مطلق در نظر گرفته می‌شود. در این روش تعداد امتیازهای مشترک و موقعیت امتیازهای مشترک نادیده گرفته می‌شوند. ماتریس شباهت اندازه‌گیری شده براساس این روش با استفاده از مثال ذکر شده ، در شکل۲ نشان داده شده است. همانطور که در جدول ۲ دیده می‌شود، بیشترین شباهت در مثال مطرح شده بین زوج (کاربر۱ و کاربر۳) است ولی براساس مقادیر بدست آمده از *MSD*، این دو کاربر بیشترین مقدار شباهت را ندارند. همچنین با توجه به شکل ۲، مقدار شباهت‌های (کاربر۱ و کاربر۲) و (کاربر۲ و کاربر۳) با استفاده از *MSD* یکسان و برابر ۰/۹۸ می‌باشد در حالیکه براساس جدول۱ زوج (کاربر۱ و کاربر۲) و زوج (کاربر۲ و کاربر۳) در امتیاز آیتم۲ یکسان و در امتیاز آیتم۱، یک واحد اختلاف دارند. همچنین برای زوج (کاربر۱ و کاربر۵) و زوج (کاربر۳ و کاربر۵) بدین شکل است که در امتیاز آیتم۱، برابرند و در امتیاز آیتم ۲، یک واحد اختلاف دارند اما مقدار شباهت *MSD* آنها یکسان و برابر ۰/۹۸ می‌باشد. عملکرد روش *MSD* تنها برای حالتی که فقط براساس اختلاف آیتم‌ها باشد، قابل قبول است.

	U_2	U_3	U_4	U_5
U_1	۰/۹۸	۰/۹۶	۰/۷۳	۰/۹۸
U_2		۰/۹۸	۰/۶۷	۰/۹۶
U_3			۰/۷۳	۰/۹۸
U_4				۰/۸

شکل۲ : ماتریس شباهت براساس روش MSD در مثال مطرح شده

۲-۳- معیار شباهت PIP

این روش شباهت براساس سه فاکتور نزدیکی[۲۶] ، تأثیر[۲۷] و محبوبیت[۲۸] محاسبه می‌شود. روش شباهت PIP برای دو کاربر u و v به صورت زیر تعریف می‌شود[۵] :

$$sim(u,v)^{PIP} = \sum_{p\epsilon I} PIP(r_{u,p}, r_{v,p}) \qquad (۲)$$

در فرمول بالا، $r_{u,p}$ و $r_{v,p}$ به ترتیب امتیاز کاربر u و v به آیتم p، I مجموعه آیتم‌های مشترک کاربر u وv و PIP نمره امتیاز کاربر u وv است که براساس رابطه۳ محاسبه می‌شود.

$$\begin{aligned}\mathrm{PIP}(r_{u,p}, r_{v,p}) = \mathrm{Proximity}(r_{u,p}, r_{v,p}) \times \\ \mathrm{Impact}(r_{u,p}, r_{v,p}) \times \mathrm{Popularity}(r_{u,p}, r_{v,p})\end{aligned} \qquad (۳)$$

شکل۳ ماتریس شباهت روش PIP را براساس مثال ذکر شده نشان می‌دهد. در این روش نسبت امتیازهای مشترک در نظر گرفته نمی‌شود. زوج (کاربر۴ و کاربر۵) و زوج (کاربر۲ و کاربر۵) امتیاز مشترک ندارند ولی مقدار شباهت آنها از زوج (کاربر۱ و کاربر۲)، زوج (کاربر۲ و کاربر۳)، زوج (کاربر۱ و کاربر۵) و زوج (کاربر۳ و کاربر۵) که یک امتیاز مشترک دارند، بیشتر است. این روش اولویت کلی کاربر را در نظر نمی‌گیرد. برای نمونه زوج (کاربر۱ و کاربر۲)، زوج (کاربر۲ و کاربر۳) دارای مقدار شباهت PIP یکسان ۰/۷۴۳ می‌باشند، ولی در آیتم۱، یک واحد اختلاف و در آیتم۲، امتیاز برابر دارند. همچنین زوج (کاربر۱ و کاربر۵) و زوج (کاربر۳ و کاربر۵) که در آیتم۱، امتیاز برابر و در آیتم۲، یک واحد اختلاف امتیاز دارند، دارای مقدار شباهت PIP ۰/۵۰۶ هستند که باید بیشتر از زوج (کاربر۱ و کاربر۲) و زوج (کاربر۲ و کاربر۳) باشد که در روش PIP برعکس شده است.
اشکال دیگر این روش، پیچیدگی محاسبات است. در این روش برای محاسبه هر قسمت دو حالت در نظر گرفته شده است که پیاده سازی را با مشکل مواجه می‌کند.

	U_2	U_3	U_4	U_5
U_1	۰/۷۴۳	۱/۰	۰/۱۶۷	۰/۵۰۶
U_2		۰/۷۴۳	۰/۱۶۲	۰/۷۶۳
U_3			۰/۱۶۷	۰/۵۰۶
U_4				۰/۷۶۷

شکل ۳ : ماتریس شباهت براساس روش PIP در مثال مطرح شده

۲-۴- معیار شباهت NHSM

در این روش، شباهت براساس سه فاکتور نزدیکی[۲۹] ، اهمیت[۳۰] و یکتایی[۳۱] محاسبه می‌گردد. روش NHSM که از محاسبات غیر خطی استفاده می‌کند حالت بهبود داده شده روش PIP است. شباهت NHSM دو کاربر را می‌توان به صورت زیر محاسبه کرد[۱]:

$$sim(u,v)^{NHSM} = sim(u,v)^{JPSS}.sim(u,v)^{URP} \qquad (۴)$$

شکل۴ ماتریس شباهت روش NHSM را براساس مثال ذکر شده را نشان می‌دهد. با توجه به نتایج بدست آمده این روش موفق نبوده است. همانطور که در شکل۴ دیده می‌شود، مقدار شباهت زوج (کاربر۲ و کاربر۵) برابر ۰/۰۳۵۶۱ بدست آمده است که از مقدار شباهت زوج (کاربر۱ و کاربر۵) بیشتر می‌باشد. در حالیکه زوج

(کاربر۱ و کاربر۵) در آیتم۱، امتیاز برابر دارند و در آیتم۲، یک واحد اختلاف دارند ولی زوج (کاربر۲ و کاربر۵) در هر دو آیتم۱ و آیتم۲ یک واحد اختلاف دارند.

اشکال دیگر این روش این است که با مشاهده امتیازهای آیتم‌های زوج (کاربر۱ و کاربر۵) و زوج (کاربر۳ و کاربر۵) دیده می‌شود که این دو حالت کاملا شبیه یکدیگر هستند پس مقدار شباهت‌ها باید برابر باشد ولی همانطور که در شکل۴ نشان داده شده است، مقدار شباهت آنها براساس NHSM برابر نیست.

	U_2	U_3	U_4	U_5
U_1	۰/۰۲۰۸۹	۰/۰۵۵۲۰	۰/۰۰۴۷۵	۰/۰۲۴۴۰
U_2		۰/۰۱۸۳	۰/۰۰۴۶۴	۰/۰۳۵۶۱
U_3			۰/۰۰۶۳۶	۰/۰۲۵۰۰
U_4				۰/۰۱۵۳۱

شکل۴ : ماتریس شباهت براساس روش NHSM در مثال مطرح

در جدول ۳ برخی از اشکالات روش‌های MSD، PIP وNHSM بیان شده است. با توجه به مطالب مطرح شده در مورد روش‌های اخیر، هدف ما ارائه روشی است که مشکلات بیان شده را حل کند.

جدول ۳ : بررسی نقاط ضعف برخی از معیارهای شباهت

نام روش	نقاط ضعف
MSD (شاردانند (۱۹۹۴))	۱- فقط امتیازهای مشترک را در نظر می‌گیرد. ۲- تعداد امتیازهای مشترک و موقعیت امتیازهای مشترک را در نظر نمی‌گیرد.
PIP (آه جون هیونگ و همکاران (۲۰۰۸))	۱- نسبت امتیازهای مشترک را در نظر نمی‌گیرد. ۲- الویت کلی کاربر را در نظر نمی‌گیرد. ۳- روش پیچیدگی محاسباتی زیادی دارد.
NHSM (لیو هایفنگ و همکاران (۲۰۱۴))	۱- تمام عوامل Impact در هم ضرب می‌شوند و تشخیص نسبت عوامل Impact در محاسبات در نظر گرفته نمی‌شود. ۲- برای حل مشکل نسبت امتیازهای مشترک از روش 'Jaccard' استفاده شده، روش 'Jaccard' برای این مشکل خیلی موفق نبوده است. ۳- روش پیچیدگی محاسباتی زیادی دارد.

از آنجا که در روش پیشنهادی از منطق فازی به منظور افزایش دقت استفاده شده است در ادامه ابتدا برخی از مفاهیم اولیه در حوزه فازی شرح داده می‌شود و سپس در فصل ۳، معیار شباهت پیشنهادی معرفی می‌گردد.

۲-۵- منطق فازی

در جهان پیرامون ما، مفاهیم نادقیق بسیاری وجود دارند که آنها را به صورت روزمره در قالب عبارتهای مختلف بیان می‌کنیم. به عنوان مثال: «هوا خوب است.» هیچ کمیتی برای خوب بودن هوا مطرح نیست تا آن را به طور دقیق اندازه‌گیری نماییم، در حالی که حسی کیفی است. در واقع مغز انسان با در نظر گرفتن عوامل گوناگون و بر پایه تفکر استنتاجی جملات را تعریف و ارزش‌گذاری می‌نماید که الگوبندی آنها به زبان و فرمول‌های ریاضی کاری بسیار پیچیده خواهد بود. منطق فازی، فناوری جدیدی است که در آن برای طراحی و مدل سازی یک سیستم، ریاضیات پیچیده و پیشرفته، با استفاده از مقادیر زبانی و دانش فرد خبره جایگزین می‌گردد. استفاده از منطق فازی، روشی موثر برای شناسایی ابهامات و عدم قطعیت و اندازه‌گیری شباهت آیتم‌ها و کاربران است[۱۴].

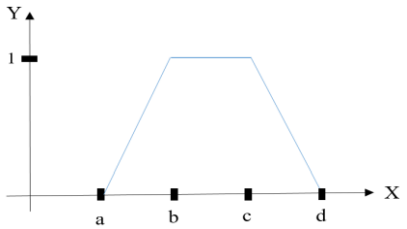
در ادامه به معرفی اعداد فازی و تابع عضویت و مجموعه فازی استفاده شده در این مقاله می پردازیم.

۲-۵-۱- اعداد فازی ذوزنقه‌ای

عدد فازی $M = (a,b,c,d)$، عدد فازی ذوزنقه‌ای گفته می‌شود اگر تابع عضویت x در M به صورت زیر است:

$$\mu_M(x) = \begin{cases} 0 & ,x \le a \\ \frac{x-a}{b-a} & ,a \le x < b \\ 1 & ,b \le x \le c \\ \frac{c-x}{d-c} & ,c < x \le d \\ 0 & ,x \ge d \end{cases} \qquad (۵)$$

عدد فازی ذوزنقه‌ای M به وسیله چهار عدد a < b < c < d تعریف می‌شود که در بازه [a,d] قاعده و x=b,c درجه عضویت یک می‌باشد. اعداد فازی ذوزنقه‌ای به صورت M=(a,b,c,d) نوشته می‌شوند. برای مثال، عدد فازی M=(a,b,c,d) در شکل ۵ نشان داده شده است[۱۵]. به منظور اندازه‌گیری شباهت بین کاربران در یک سیستم فازی، انجام برخی محاسبات فازی نیاز است. در این قسمت محاسبات فازی برای چهار عمل اصلی شرح داده شده است[۱۶].



شکل ۵ : اعداد فازی ذوزنقه‌ای[۱۵].

فرض کنید که دو عدد فازی ذوزنقه‌ای مانند A و B وجود دارد به طوری که :

$$A = (a_1, a_2, a_3, a_4 : 1) \quad (۶)$$

$$B = (b_1, b_2, b_3, b_4 : 1) \quad (۷)$$

که در آن $i = \{1,2,3,4\}\ ;\ 0 \le a_i{}'b_i \le 1$.

بعضی از عملیات ریاضی پایه، بین اعداد فازی ذوزنقه‌ای به صورت زیر تعریف شده است :

۱- جمع دو عدد فازی ذوزنقه‌ای:

$$A + B = (a_1 + b_1, a_2 + b_2, a_3 + b_3, a_4 + b_4) \quad (۸)$$

۲- تفریق دو عدد فازی ذوزنقه‌ای :

$$A - B = A + (-B) =$$

$$(a_1 - b_4, a_2 - b_3, a_3 - b_2, a_4 - b_1) \quad (۹)$$

۳- ضرب دو عدد فازی ذوزنقه‌ای:

$$A \times B = (a_1 \times b_1, a_2 \times b_2, a_3 \times b_3, a_4 \times b_4) \quad (۱۰)$$

۴- تقسیم دو عدد فازی ذوزنقه‌ای:

$$A \div B = (a_1/b_4,\ a_2/b_3,\ a_3/b_2,\ a_4/b_1) \quad (۱۱)$$

$$\text{where } {}^{a}/_{b} = \begin{cases} \dfrac{a}{b} & , \quad a < b \\ 1 & , \text{otherwise} \end{cases}$$

در این مقاله از اصطلاح زبانشناسی معرفی شده در جدول ۴ که در مقاله(۱۷) ارائه شده است، استفاده شده است[۲۰-۱۷]. براساس این مدل از اصطلاح زبانی و اعداد فازی ذوزنقه‌ای استفاده شده است. به عنوان مثال، اگر کاربر به آیتمی در مجموعه داده غیرفازی امتیاز ۵ بدهد یا اگر گزینه بسیار خوب در اصطلاح زبانشناسی انتخاب کند،

جدول ۴ : اصطلاح زبانی اعداد فازی ذوزنقه‌ای[۲۰-۱۷]

اعداد فازی ذوزنقه‌ای	اصطلاح زبانشناسی	امتیاز واقعی
(۱: ۱، ۱، ۰/۹۸، ۰/۹۳)	بسیار خوب	۵
(۱: ۰/۹۷، ۰/۹۲، ۰/۸، ۰/۵۸)	خوب	۴
(۱: ۰/۶۵، ۰/۵۸، ۰/۴۱، ۰/۳۲)	بدون نظر	۳
(۱: ۰/۴۲، ۰/۳۶، ۰/۱۸، ۰/۰۴)	بد	۲
(۱: ۰/۰۷، ۰/۰۲، ۰، ۰)	بسیار بد	۱

در مجموعه اعداد فازی با عدد فازی (۱: ۱، ۱، ۰/۹۸، ۰/۹۳) معادل است.

۳- روش پیشنهادی

در این قسمت روش پیشنهادی انداه‌گیری شباهت در سیستم‌های توصیه‌گر که FSR نامیده شده است به صورت کامل شرح داده می‌شود. در روش اندازه‌گیری شباهت پیشنهادی، از معیارهای اهمیت، محبوبیت و روش میانگین اختلاف مربع استفاده می‌شود. همچنین، به علت وجود عدم قطعیت و نسبی بودن امتیازهای کاربران به آیتم‌ها، به منظور بهبود دقت از اعداد فازی استفاده می‌کنیم. به علت استفاده از اعداد فازی در روش پیشنهادی، به منظور کاهش پیچیدگی و زمان محاسبات از محاسبات خطی استفاده شده است. در شکل۶، دیاگرام مراحل اندازه‌گیری شباهت بر اساس روش پیشنهادی FSR نشان داده شده است. در ادامه تمام مراحل و پارامترها به صورت کامل شرح داده می‌شود.

۳-۱-۱- محاسبه معیار شباهت پیشنهادی (FSR)

شکل ۶ : دیاگرام مراحل اندازه‌گیری شباهت براساس روش پیشنهادی (FSR)

در محاسبه شباهت ، اولین فاکتور استفاده شده فاکتور اهمیت[۳۲] است در هنگامی که امتیاز ۲ آیتم برابر باشد، اهمیت بیشتری را ایجاد می‌کند و در حالاتی که امتیاز ۲ آیتم برابر نیست، امتیازهایی که به متوسط امتیاز کاربران نزدیکتر است اهمیت بیشتری پیدا می‌کند. در این بخش سعی شده است، فرمول به گونه‌ای باشد که نتایج حاصل در بهترین شرایط برابر ۱ و در دیگر حالات کمتر از ۱ باشد. برای محاسبه فاکتور اهمیت، از فرمول زیر استفاده می‌شود:

$$if(r_1 == r_2) \quad \text{Significance}(r_1, r_2) = 1 \quad (۱۲)$$

$$else \quad \text{Significance}(r_1,r_2) = \frac{1}{(|r_1 - r_{med}| + 1).(|r_2 - r_{med}| + 1)}$$

در رابطه (۱۲)، $|r_1 - r_{med}|$ برابر قدرمطلق اختلاف امتیاز کاربر با متوسط امتیاز کاربران می‌باشد. در رابطه بالا، نتایج مخرج به کوچکترین عدد طبیعی یعنی عدد ۱ اضافه می‌شود تا از صفر شدن مخرج جلوگیری کند. همچنین اعداد فازی استفاده شده در بازه [۰-۱] قرار دارند و در پیاده سازی این روش با اعداد فازی، مقدار عدد ۱ در مخرج با مقدار ۰/۰۰۱ که کوچکترین عددی است که از صفر شدن مخرج جلوگیری و همچنین کمترین تاثیر را بر روی نتایج دارد، جایگزین می‌شود، در نتیجه مقدار نهایی فاکتور اهمیت به صورت زیر محاسبه می‌گردد:

$$if(r_1 == r_2) \quad \text{Significance}(r_1,r_2) = 1 \qquad (۱۳)$$

$$else \quad \text{Significance}(r_1,r_2) = \frac{1}{(|r_1 - r_{med}| + 0/001).(|r_2 - r_{med}| + 0/001)}$$

دومین فاکتور استفاده شده در محاسبه شباهت فاکتور محبوبیت[۳۳] است. اگر معدل امتیاز دو کاربر از میانگین امتیاز آیتم مورد نظر بیشتر باشد، فاکتور محبوبیت برای شباهت کاربر، ارزش بیشتری ایجاد می‌کند[۱, ۵]. برای محاسبه فاکتور محبوبیت از رابطه زیر استفاده شده است[۵]:

$$if\big((r_1 > \mu_k \; and \; r_2 > \mu_k) \; or \; (r_1 < \mu_k \; and \; r_2 < \mu_k)\big) \qquad (۱۴)$$

$$Popularity = 1 + \left(\frac{r_1 + r_2}{2} - \mu_k\right)^2$$
$$else \quad Popularity = 1$$

که μ_k میانگین امتیاز آیتم‌ها است. همانطور که در رابطه (۱۴) مشاهده می‌شود، اگر امتیاز آیتم‌های هر دو کاربر بزرگتر و یا کوچکتر از میانگین امتیاز آیتم‌ها باشد نتیجه بدست آمده بزرگتر از ۱ می‌شود و در غیر این صورت مقدار این فاکتور برابر ۱ می‌شود. در نتیجه این فاکتور کمترین مقداری که می‌گیرد برابر ۱ است.

در ادامه برای بهبود نتایج و استفاده از تاثیر عامل نسبت امتیازهای مشترک رابطه (۱۵) طراحی شده است.

$$SPI = \frac{(\text{Significance} \times Popularity)}{I} \qquad (۱۵)$$

که I تعداد امتیاز مشترک بین دو کاربر است.

در ادامه تاثیر استفاده از رابطه (۱۵) با توجه به مثال مطرح شده بخش دوم برای روش پیشنهادی به صورت زیر بیان می‌شود. در این قسمت کاربر۱ و کاربر۳، هر دو برای چهار آیتم امتیاز دارند که در سه آیتم امتیاز برابر و در آیتم بعدی ۲ واحد اختلاف امتیاز دارند. در حالی که، کاربر۱ و کاربر۲ فقط برای دو آیتم دارای امتیاز هستند و در این دو آیتم، در اولی مقدار امتیاز برابر و در دیگری ۱ واحد اختلاف امتیاز دارند. با توجه به توضیحات بیان شده مشخص است که مقدار شباهت کاربر۱ و کاربر۳ چون دارای سه آیتم برابر هستند باید بیشتر از مقدار شباهت کاربر۱ و کاربر۲ باشد که در یک آیتم برابر هستند. در روش پیشنهادی با توجه به شکل ۷ مقدار شباهت اندازه‌گیری شده برای کاربر۱ و کاربر۳ و کاربر۱ و کاربر۲ به ترتیب برابر ۰/۸۰۴۶ و ۰/۶۱۳۶ است.

در نهایت برای محاسبه شباهت باید اختلاف امتیاز آیتم‌ها در نظر گرفته شود. با توجه به نتایج بدست آمده از روش‌های موجود، مشاهده می‌شود که روش MSD بهترین عملکرد برای نمایش اختلاف امتیاز آیتم‌ها دارد. برای افزایش دقت شباهت اندازه‌گیری شده، در روش پیشنهادی برای محاسبه اختلاف امتیاز آیتم‌ها، از روش MSD استفاده شده است.

$$Sim_{NPFSM} = SPI \times MSD \qquad (۱۶)$$

رابطه نهایی برای اندازه‌گیری شباهت در این پژوهش از حاصل ضرب نتایج قبلی در روش MSD بدست می‌آید که در معادله (۱۶) نشان داده شده است. نتیجه روش پیشنهادی برای مثال مطرح شده در شکل۷ نشان داده شده است.

	U_2	U_3	U_4	U_5
U_1	۰/۶۱۳۶	۰/۸۰۴۶	۰/۲۱۳۹	۰/۷۳۳۳
U_2		۰/۶۱۳۶	۰/۱۶۶۷	۰/۳۶۰۳
U_3			۰/۲۱۳۹	۰/۷۳۳۳
U_4				۰/۲۲۰۷

شکل ۷ : ماتریس شباهت براساس روش پیشنهادی در مثال مطرح

شکل ۸ ، بیشترین و کمترین شباهت براساس مثال مطرح شده در بخش قبل را نشان می‌دهد. همچنین، شکل ۷ نتایج شباهت بدست آمده از مثال مطرح شده را براساس روش پیشنهادی نشان می‌دهد. با دقت بر روی نتایج شکل ۷ مشاهده می‌شود که ترتیب بیشترین تا کمترین مقدار شباهت بدست آمده دقیقا مانند شکل ۸ می‌باشد.

۴- ارزیابی روش پیشنهادی FSR

در این بخش، ابتدا مجموعه داده‌ها، روش‌های ارزیابی و معیارهای سنجش معرفی می‌گردند. سپس، روش اندازه‌گیری شباهت

شکل۸ : ترتیب کمترین و بیشترین شباهت در مثال مطرح شده

پیشنهادی ارزیابی شده و در نهایت نتایج بدست آمده از روش پیشنهادی با سایر روش‌های مطرح شده، مقایسه می‌گردد.

٤-١- مجموعه داده‌ها[۳۴]

در این پژوهش به منظور افزایش کیفیت ارزیابی روش پیشنهادی از دو مجموعه داده متفاوت از سری مجموعه داده‌های MovieLens[۳۵] استفاده شده است. مجموعه داده‌های MovieLens، مجموعه داده واقعی تهیه شده در دانشگاه مینه سوتا[۳۶] هستند که شامل امتیاز کاربران به فیلم‌ها می‌باشد که به صورت آنلاین جمع آوری شده است[21,5,1]. در این پژوهش از مجموعه‌داده MovieLens-1M، که شامل ۳۹۵۲ فیلم، ۶۰۴۰ کاربر و ۱۰۰۰۲۰۹ امتیاز و مجموعه‌داده MovieLens-100K، که شامل ۹۴۳ فیلم، ۱۶۸۲ کاربر و ۱۰۰۰۰۰ امتیاز استفاده شده است. داده‌های مورد استفاده شامل userId، itemId و rating می‌باشد.

٤-٢- معیارهای سنجش

معیارهای مختلفی برای ارزیابی سیستم‌های توصیه‌گر وجود دارند. این معیارها شامل صحت (نسبت تشخیص[۳۷])، حساسیت[۳۸] (یا فراخوانی[۳۹])، وضوح[۴۰] ، دقت[۴۱] ، F_1 و F_B می‌شوند. معیارهای دقت و فراخوانی به طور گسترده‌ای مورد استفاده قرار می‌گیرند. معیار دقت به عنوان معیار درستی و معیار فراخوانی به عنوان معیار تمامیت در نظر گرفته می‌شود[۴].

همچنین در ارزیابی سیستم‌های توصیه‌گر معیارهای خطا از جمله: میانگین خطای مطلق[۴۲] (MAE)، خطای جذر میانگین مربعات[۴۳] (RMSE) و میانگین خطای مطلق نرمال[۴۴] (NMAE) نیز مورد استفاده قرار می‌گیرد[۲۳].

در این پژوهش از معیارهای فراخوانی، صحت،F1 و MAE برای ارزیابی استفاده می‌شود.

در این پژوهش از روش اعتبار سنجی متقابل[۴۵] (k=10) استفاده شده است[۴].

٤-٣- شرح روش ارزیابی

روش ارزیابی در این پژوهش به صورت زیر است:

در ابتدا با استفاده از مقدار شباهت بدست آمده بین کاربران، k همسایه نزدیک به کاربر فعال انتخاب می‌شود. برای کاربر فعال، تمامی کاربران براساس میزان شباهت بدست آمده مرتب شده و در ادامه k کاربری که بیشترین شباهت را با کاربر فعال دارند انتخاب می‌شود.

در ادامه، بعد از انتخاب k کاربر شبیه به کاربر فعال، برای بدست آوردن پیش‌بینی مقدار آیتم i برای کاربر فعال، روش مجموع وزنی مورد استفاده قرار گرفته است. برای محاسبه معیارهای فراخوانی، دقت، F_1 مطابق رابطه‌های زیر استفاده شده است :

$$Precision = \frac{1}{\#U}\sum_{u \in U}\frac{\#\{i \in Z_u | r_{u,i} \geq \theta\}}{N} \qquad (۱۷)$$

$$Recall = \frac{1}{\#U}\sum_{u \in U}\frac{\#\{i \in Z_u | r_{u,i} \geq \theta\}}{\#\{i \in Z_u | r_{u,i} \geq \theta\}+\#\{i \in Z_u^c | r_{u,i} \neq \bullet \wedge r_{u,i} \geq \theta\}} \qquad (۱۸)$$

در روابط بالا، Z_u مجموعه N توصیه به کاربر، u و θ حد آستانه و N ثابت در نظر گرفته شده است.

$$F_1 = \frac{2*Precision*Recall}{Precision+Recall} \qquad (۱۹)$$

همانطور که در معادله (۱۹) نشان داده شده است برای محاسبه F1، از مقادیر Precision و Recall استفاده گردیده است[۲۳]. معیار میانگین خطای مطلق(MAE) میزان انحراف امتیاز پیش بینی شده با امتیاز واقعی را محاسبه می‌کند. برای محاسبه معیار میانگین خطای مطلق(MAE) از رابطه زیر استفاده می‌شود[۲۴]:

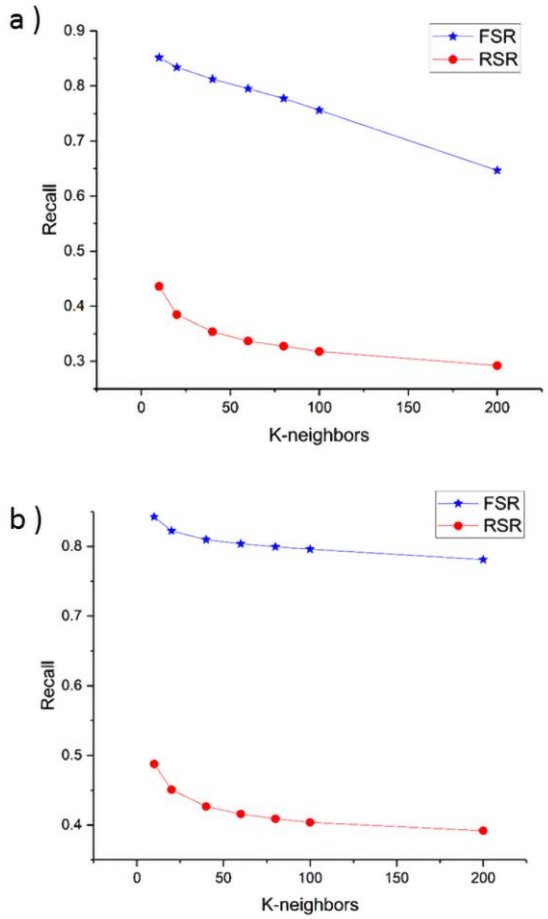
$$MAE = \frac{1}{M_A}\sum_{i=1}^{M_A}\left(\frac{1}{N_i^p}\sum_{k=1}^{N_i^p}|p_{i,k} - r_{i,k}|\right) \qquad (۲۰)$$

در رابطه بالا، M_A تعداد کل کاربران فعال، N_i^p تعدادکل موارد پیش بینی شده برای کاربر u_i، $p_{i,k}$ مقدار پیش بینی شده، $r_{i,k}$ امتیاز واقعی، $rmax$ بیشترین مقدار امتیاز و $rmin$ کمترین مقدار امتیاز است.

در ادامه، ابتدا به منظور بررسی تاثیر استفاده از اعداد فازی، روش پیشنهادی با استفاده از اعداد فازی را با روش پیشنهادی با استفاده از اعداد واقعی مورد ارزیابی قرار می‌دهیم. سپس در ادامه روش پیشنهادی با سایر روش‌های اخیر مورد مقایسه قرار می‌گیرد.

٤-٤- بررسی روش پیشنهادی

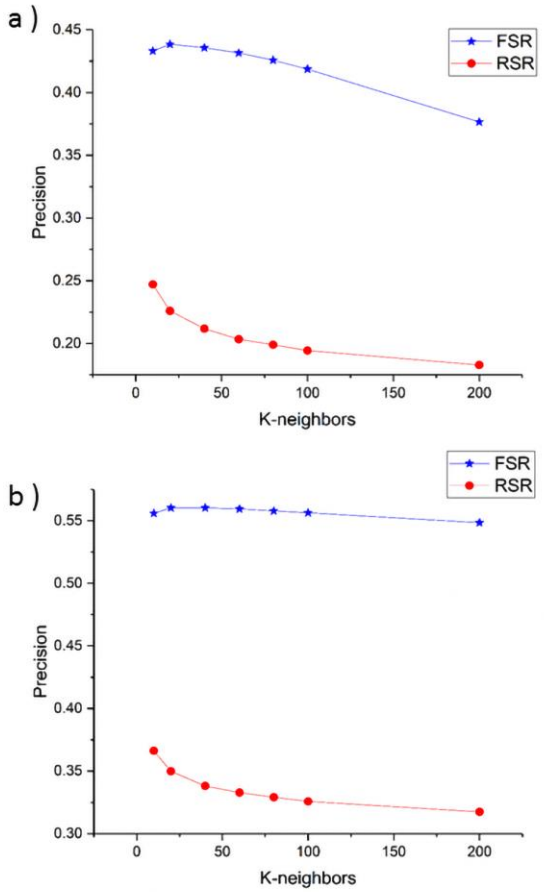
همانگونه که در فصل ۳ بیان گردید در روش پیشنهادی FSR به منظور افزایش دقت از منطق فازی استفاده گردیده است. به منظور ارزیابی تاثیر استفاده از منطق فازی، یک بار روش پیشنهادی با استفاده از اعداد فازی پیاده سازی و مورد ارزیابی قرار گرفته است و یک بار با اعداد واقعی. سپس، نتایج بدست آمده با هم مقایسه می‌گردد. نتایج بدست آمده با اعداد فازی با FSR و نتایج با اعداد واقعی با RSR[۴۶] نمایش داده شده است و بر روی دو مجموعه داده MovieLens-1M و MovieLens-100K محاسبه گردیده است.



شکل ۹ : مقادیر Recall با استفاده از روش پیشنهادی در حالت FSR و در حالت RSR

a) بر روی مجموعه داده MovieLens-100K b) بر روی مجموعه داده MovieLens-1M

در شکل ۹(a) و ۹(b)، مقادیر معیار Recall بدست آمده براساس FSR و RSR را بر روی مجموعه داده MovieLens-100K و MovieLens-1M به ترتیب نشان می‌دهد. در این حالت هر چقدر مقدار Recall بدست آمده بیشتر باشد عملکرد سیستم بهتر است[۴, ۲۳]. مقادیر معیار Recall در حالت K-neighbors =۱۰۰ به ترتیب برای روش‌های RSR و FSR با استفاده از مجموعه داده MovieLens-100K برابر ۰/۳۱ و ۰/۷۵ و با استفاده از مجموعه داده MovieLens-1M برابر ۰/۴۰ و ۰/۷۹ است. نتایج نشان می‌دهد که روش FSR نسبت به روش RSR در هر دو مجموعه داده دارای مقدار فراخوانی بهتری می‌باشد.

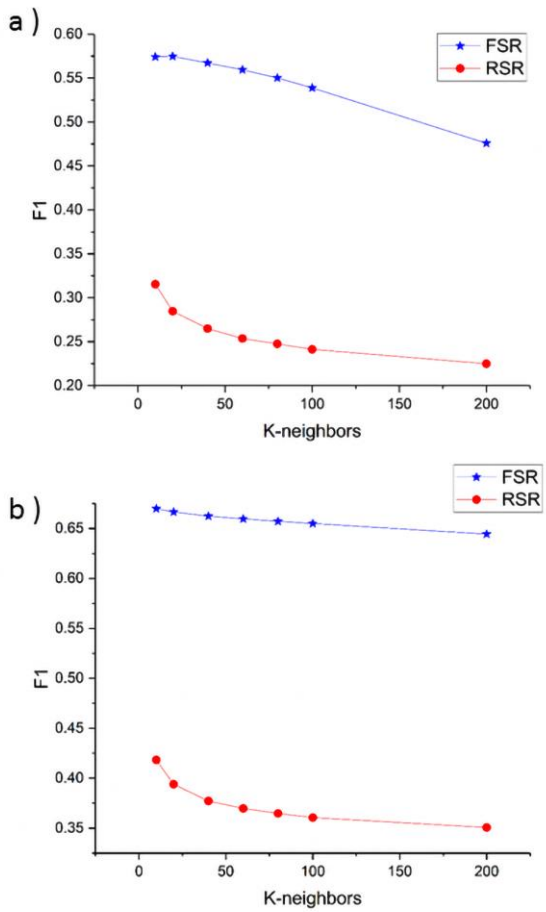


شکل ۱۰ : مقادیر Precision با استفاده از روش پیشنهادی در حالت FSR و در حالت RSR

a) بر روی مجموعه داده MovieLens-100K b) بر روی مجموعه داده MovieLens-1M

در شکل ۱۰(a) و ۱۰(b)، مقادیر معیار Precision بدست آمده براساس FSR و RSR را بر روی مجموعه داده MovieLens-100K و MovieLens-1M به ترتیب نشان می‌دهد. در این حالت هر چقدر مقدار Precision بدست آمده بیشتر باشد عملکرد سیستم بهتر است[۴, ۲۳]. مقادیر معیار Precision در حالت K-neighbors =۱۰۰ به ترتیب برای روش‌های RSR و FSR با استفاده از مجموعه داده MovieLens-100K برابر ۰/۱۹ و ۰/۴۲ و با استفاده از مجموعه داده

MovieLens-1M برابر ۰/۳۲ و ۰/۵۵ است. با توجه به نتایج بدست آمده بر اساس معیار Precision روش FSR نسبت به روش RSR در هر دو مجموعه داده دارای دقت بهتری می‌باشد.



شکل ۱۱ : مقادیر F1 با استفاده از روش پیشنهادی در حالت FSR و در حالت RSR

a) بر روی مجموعه داده MovieLens-100K b) بر روی مجموعه داده MovieLens-1M

در شکل (a)۱۱ و (b)۱۱، مقادیر معیار F1 بدست آمده براساس FSR و RSR را بر روی مجموعه داده MovieLens-100K و MovieLens-1M به ترتیب نشان می‌دهد. در این حالت هر چقدر مقدارF1 بدست آمده بیشتر باشد عملکرد سیستم بهتر است[۲۳,۴]. مقادیر معیار F1 در حالت K-neighbors =۱۰۰ به ترتیب برای روش‌های RSR و FSR با استفاده از مجموعه داده MovieLens-100K برابر ۰/۲۴ و ۰/۵۳ و با استفاده از مجموعه داده MovieLens-1M برابر ۰/۳۶ و ۰/۶۵ است. با توجه به نتایج بدست آمده بر اساس معیار F1 روش FSR نسبت به روش RSR در هر دو مجموعه داده عملکرد بهتری دارد.

با توجه به نتایج، جدول ۵ و ۶ میزان بهبود روش پیشنهادی در حالت FSR نسبت به روش پیشنهادی در حالت RSR را به ترتیب بر روی مجموعه داده MovieLens-100K و MovieLens-1M نشان می‌دهد. برای ارزیابی بهتر و انجام مقایسه در شرایط مختلف، روش پیشنهادی در حالت FSR نسبت به روش پیشنهادی در حالت RSR در K-neighbors برابر ۴۰، ۱۰۰ و ۲۰۰ بررسی کردیم که نتایج در جدول ۵ و ۶ نشان داده شده است. با توجه به نتایج، مشاهده می‌شود که در حالت K-neighbors=۱۰۰ و با استفاده از مجموعه داده MovieLens-1M، مقدار Recall در روش پیشنهادی در حالت FSR نسبت به روش پیشنهادی در حالت RSR، ۴۹٪ و با استفاده از مجموعه داده MovieLens-۱۰۰K، ۵۸٪ بهبود داشته است. همچنین، در حالت K-neighbors=۱۰۰، مقدار Precision در روش پیشنهادی در حالت FSR نسبت به روش پیشنهادی در حالت RSR، ۴۱٪ با استفاده از مجموعه داده MovieLens-1M و ۵۳٪ با استفاده از مجموعه داده MovieLens-100K بهبود یافته است. مقدار F1 نیز در روش پیشنهادی در حالت FSR نسبت به روش پیشنهادی در

جدول ۵ : میزان بهبود روش پیشنهادی FSR نسبت روش RSR بر روی مجموعه داده MovieLens-100K

		درصد بهبود Recall	درصد بهبود Precision	درصد بهبود F1
FSR نسبت RSR	K=۴۰	۵۶٪	۵۱٪	۵۳٪
	K=۱۰۰	**۵۸٪**	**۵۳٪**	**۵۵٪**
	K=۲۰۰	۵۵٪	۵۱٪	۵۳٪

جدول ۶ : میزان بهبود روش پیشنهادی FSR نسبت روش RSR بر روی مجموعه داده MovieLens-1M

		درصد بهبود Recall	درصد بهبود Precision	درصد بهبود F1
FSR نسبت RSR	K=۴۰	۴۷٪	۴۰٪	۴۳٪
	K=۱۰۰	۴۹٪	۴۱٪	**۴۵٪**
	K=۲۰۰	**۵۰٪**	**۴۲٪**	**۴۵٪**

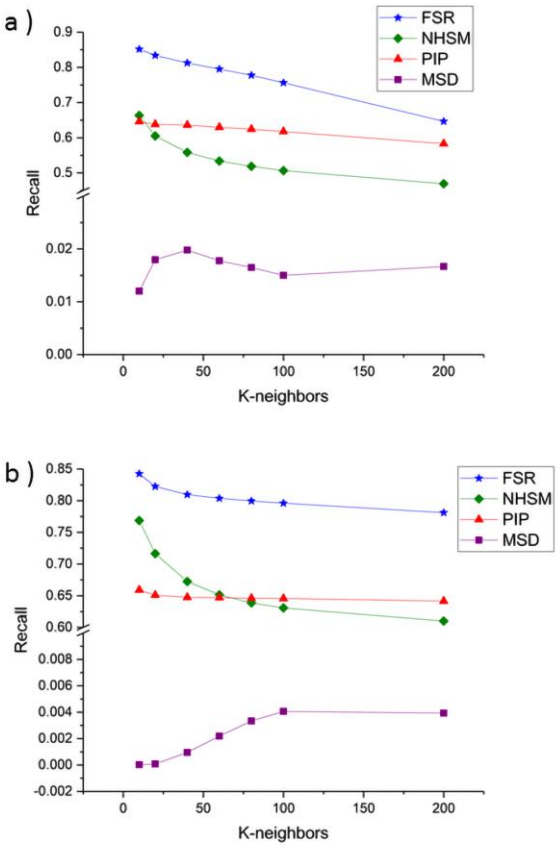
حالت RSR، ۴۵٪ با استفاده از مجموعه داده MovieLens-1M و ۵۵٪ با استفاده از مجموعه داده MovieLens-100K، بهبود داشته است.

با توجه به نتایج ارائه شده در این قسمت، روش پیشنهادی FSR براساس سه معیار Recall، Precision وF1 عملکرد خیلی بهتری نسبت به روش RSR دارد. به طور کلی این نتایج نشان می‌دهد که

استفاده از اعداد فازی در روش پیشنهادی موجب افزایش دقت و بهبود کلی معیار شباهت پیشنهادی FSR شده است. در ادامه این فصل، معیار شباهت پیشنهادی FSR با سایر روش‌های اخیر محاسبه شباهت یعنی PIP و NHSM و همچنین روش پایه MSD مورد مقایسه قرار می‌گیرد.

٤-٥- مقایسه روش پیشنهادی با سایر معیارهای شباهت

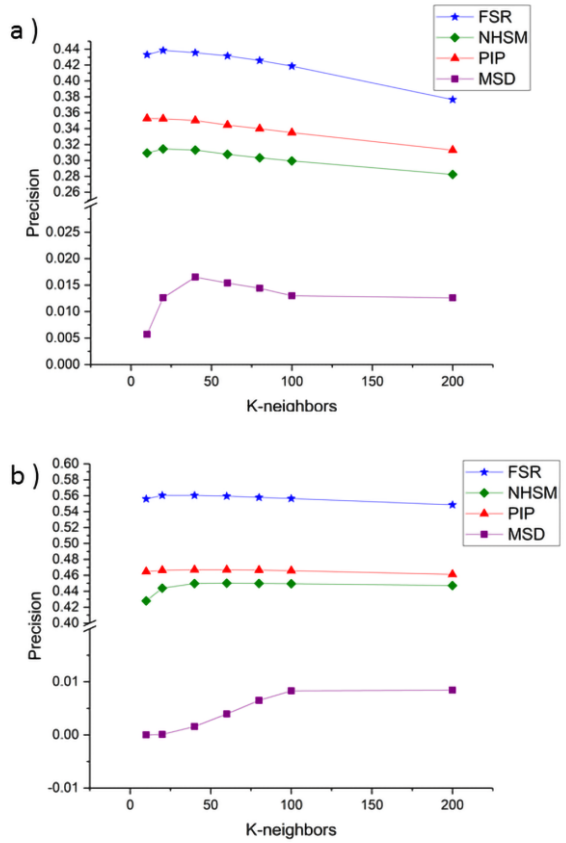
در این قسمت، روش پیشنهادی فازی FSR، با روش پایه MSD و همچنین روش‌های جدید اندازه‌گیری شباهت در سیستم‌های توصیه‌گر یعنی PIP و NHSM مقایسه شده است.



شکل ۱۲: مقادیر Recall با استفاده از روش پیشنهادی در حالت FSR و در حالت RSR

a) بر روی مجموعه داده MovieLens-100K b) بر روی مجموعه داده MovieLens-1M

در شکل ۱۲(a) و ۱۲(b)، مقادیر معیار Recall بدست آمده براساس FSR، NHSM، PIP و MSD را بر روی مجموعه داده MovieLens-100K و MovieLens-1M به ترتیب نشان می‌دهد. مقادیر معیار Recall در حالت K-neighbors =۱۰۰ به ترتیب برای روش‌های NHSM و PIP و MSD با استفاده از مجموعه داده MovieLens-100K برابر ۰/۵۲ و ۰/۶۱ و ۰/۰۱ برای روش FSR برابر ۰/۷۵ و با استفاده از مجموعه داده MovieLens-1M برابر ۰/۶۳ و ۰/۶۴ و ۰/۰۰ برای روش FSR برابر ۰/۷۹ است. براساس نتایج



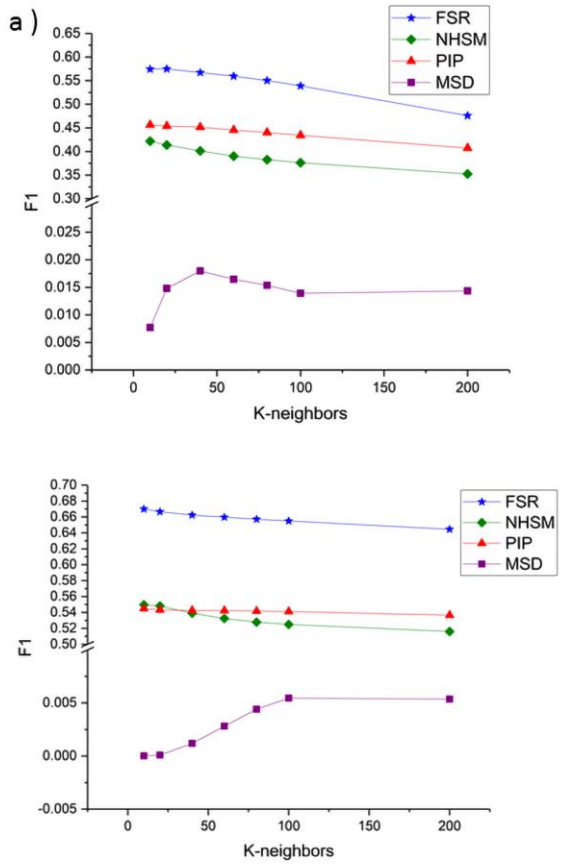
شکل ۱۳ : مقادیر Precision با استفاده از روش پیشنهادی در حالت FSR و در حالت RSR

a) بر روی مجموعه داده MovieLens-100K b) بر روی مجموعه داده MovieLens-1M

بدست آمده، روش FSR براساس معیار Recall بهترین عملکرد را در تمامی حالات نسبت به سایر روش‌ها دارا می‌باشد.

در شکل ۱۳(a) و ۱۳(b)، مقادیر معیار Precision بدست آمده براساس FSR، NHSM، PIP و MSD را بر روی مجموعه داده MovieLens-100K و MovieLens-1M به ترتیب نشان می‌دهد. مقادیر معیار Precision در حالت K-neighbors =۱۰۰ به ترتیب برای روش‌های NHSM و PIP و MSD با استفاده از مجموعه داده MovieLens-100K برابر ۰/۲۹ و ۰/۳۳ و ۰/۰۱ برای روش FSR برابر ۰/۴۲ و با استفاده از مجموعه داده MovieLens-1M برابر ۰/۴۴ و ۰/۴۶ و ۰/۰۱ برای روش FSR برابر ۰/۵۵ است. همانطور که در شکل ۱۳ دیده‌می‌شود، ، روش پیشنهادی FSR بهترین دقت را در تمامی حالات نسبت به سایر روش‌ها دارا می‌باشد. روش PIP نیز بعد

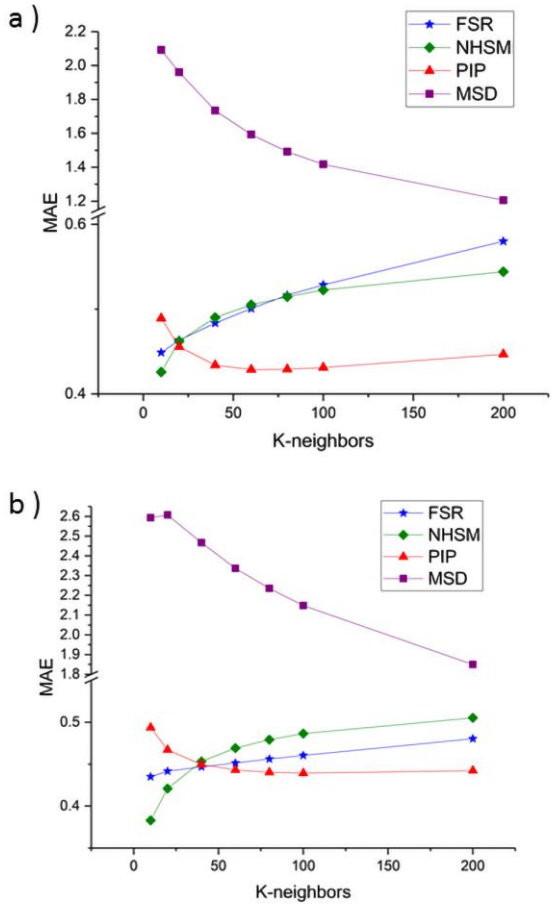
از روش پیشنهادی FSR، عملکرد بهتری نسبت به روش NHSM دارد.



شکل ۱۴ : مقادیر F1 با استفاده از روش پیشنهادی در حالت FSR و در حالت RSR

a) بر روی مجموعه داده b MovieLens-100K) بر روی مجموعه داده MovieLens-1M

در شکل ۱۴(a) و ۱۴(b)، مقادیر معیار F1 بدست آمده براساس FSR، NHSM، PIP و MSD را بر روی مجموعه داده MovieLens-100K و MovieLens-1M به ترتیب نشان می‌دهد. مقادیر معیار F1 در حالت K-neighbors =۱۰۰ به ترتیب برای روش‌های NHSM و PIP و MSD با استفاده از مجموعه داده MovieLens-100K برابر ۰/۳۷ و ۰/۴۳ و ۰/۰۱ برای روش FSR برابر ۰/۵۳ و با استفاده از مجموعه داده MovieLens-1M برابر ۰/۵۲ و ۰/۵۴ و ۰/۰۰ برای روش FSR برابر ۰/۶۵ است. معیار F1 از معیارهای Recall و Precision بدست می‌آید و با توجه به عملکرد یکسان دو معیار، همان طور که در شکل ۱۴ دیده می‌شود، روش پیشنهادی FSR با فاصله زیادی عملکرد بهتری نسبت به سایر روش‌های مورد بررسی دارد.



شکل ۱۵ : مقادیر MAE با استفاده از روش پیشنهادی در حالت FSR و در حالت RSR

a) بر روی مجموعه داده b MovieLens-100K) بر روی مجموعه داده MovieLens-1M

در شکل ۱۵(a) و ۱۵(b)، مقادیر معیار MAE بدست آمده براساس FSR، NHSM، PIP و MSD را بر روی مجموعه داده MovieLens-100K و MovieLens-1M به ترتیب نشان می‌دهد. در این حالت هر چقدر مقدار MAE بدست آمده کمتر باشد عملکرد سیستم بهتر است[۴,۲۳]. مقادیر معیار MAE در حالت K-neighbors =۱۰۰ به ترتیب برای روش‌های NHSM و PIP و MSD با استفاده از مجموعه داده MovieLens-100K برابر ۰/۵۲ و ۰/۴۳ و ۱/۴۱ و برای روش FSR برابر ۰/۵۲ و با استفاده از مجموعه داده MovieLens-1M برابر ۰/۴۸ و ۰/۴۳ و ۲/۱۴ و برای روش FSR برابر ۰/۴۶ است. روش PIP در K-neighbors =۱۰۰ بالاترین مقدار MAE را دربین روش‌های FSR و NHSM را دارست و با افزایش K-neighbors مقدار MAE کاهش می‌یابد. همچنین، روش FSR و روش NHSM براساس معیار MAE عملکرد مشابهی دارند.

جدول ۷ : میزان بهبود روش پیشنهادی FSR نسبت به سایر روش‌ها بر روی مجموعه داده MovieLens-100K

		درصد بهبود Recall	درصد بهبود Precision	درصد بهبود F1	درصد بهبود MAE
نسبت FSR PIP	K=۴۰	**۲۲٪**	**۲۰٪**	**۲۰٪**	-
	K=۱۰۰	۱۸٪	**۲۰٪**	۱۹٪	-
	K=۲۰۰	۱۶٪	۱۷٪	۱۴٪	-
نسبت FSR NHSM	K=۴۰	**۴۰٪**	**۲۸٪**	۲۹٪	۱٪
	K=۱۰۰	۳۱٪	**۲۸٪**	**۳۰٪**	-
	K=۲۰۰	۱۶٪	۲۵٪	۲۶٪	-

جدول ۸ : میزان بهبود روش پیشنهادی FSR نسبت به سایر روش‌ها بر روی مجموعه داده MovieLens-1M

		درصد بهبود Recall	درصد بهبود Precision	درصد بهبود F1	درصد بهبود MAE
نسبت FSR PIP	K=۴۰	**۲۰٪**	**۱۷٪**	**۱۸٪**	۱٪
	K=۱۰۰	۱۹٪	۱۶٪	۱۷٪	-
	K=۲۰۰	۱۸٪	۱۶٪	۱۷٪	-
نسبت FSR NHSM	K=۴۰	۱۷٪	**۲۰٪**	۱۹٪	۱٪
	K=۱۰۰	۲۱٪	۱۹٪	**۲۰٪**	**۵٪**
	K=۲۰۰	**۲۲٪**	۱۸٪	**۲۰٪**	**۵٪**

در ادامه با توجه به نتایج بدست آمده، جمع بندی از میزان بهبودهای بدست آمده ارائه می‌کنیم. جدول ۷ و ۸ میزان بهبود روش پیشنهادی FSR نسبت به سایر روش‌ها را به ترتیب بر روی مجموعه داده MovieLens-100K و MovieLens-1M نشان می‌دهد. با توجه به نتایج گزارش شده و اختلاف روش پایه MSD با سایر روش‌ها و همچنین تمرکز بر روی روش‌های جدید، نتایج روش MSD در جداول بهبود حذف شده است. برای ارزیابی بهتر و انجام مقایسه در شرایط مختلف، تمامی روش‌ها را با روش پیشنهادی FSR در K-neighbors برابر ۴۰، ۱۰۰ و ۲۰۰ بررسی کردیم که نتایج در جدول ۷ و ۸ نشان داده شده است. با توجه به نتایج، مشاهده می‌شود که در حالت K-neighbors=۱۰۰ و با استفاده از مجموعه داده MovieLens-1M، مقدار Recall در روش پیشنهادی FSR نسبت به روش‌های PIP و NHSM، ۱۹٪ و ۲۱٪ و با استفاده از مجموعه داده MovieLens-100K، ۱۸٪ و ۳۱٪ بهبود داشته است. همچنین، در حالت K-neighbors=۱۰۰، مقدار Precision در روش پیشنهادی FSR نسبت به روش‌های PIP و NHSM، ۱۶٪ و ۱۹٪ با استفاده از مجموعه داده MovieLens-1M و ۲۰٪ و ۲۸٪ با استفاده از مجموعه داده MovieLens-100K بهبود یافته است.

مقدار F1 نیز در روش پیشنهادی FSR نسبت به روش‌های PIP و NHSM، ۱۷٪ و ۲۰٪ با استفاده از مجموعه داده MovieLens-1M و ۱۹٪ و ۳۰٪ با استفاده از مجموعه داده MovieLens-100K، بهبود داشته است. همچنین روش FSR نسبت به روش NHSM مقدار MAE در K-neighbors=۱۰۰ را ۵٪ بهبود داده است. براساس نتایج می‌توان گفت که روش پیشنهادی FSR براساس سه معیار Recall، Precision و F1 از روش‌های PIP و NHSM به مراتب نتایج خیلی بهتری می‌دهد و براساس معیار MAE رفتاری مشابه NHSM و کمی ضعیف‌تر از PIP دارد.

۵- نتیجه‌گیری

امروزه با توسعه اینترنت و افزایش کاربران استفاده از سیستم‌های توصیه‌گر گسترش یافته است. استفاده از سیستم‌های توصیه‌گر باعث ارائه توصیه‌های دقیق و با کیفیت می‌شوند. دقت و کیفیت، ۲ فاکتور مهم در این سیستم‌ها هستند و تمامی روش های ارائه شده در این زمینه در تلاش هستند که این دو فاکتور را افزایش دهند.

در این مقاله، یک روش اندازه‌گیری شباهت فازی در سیستم‌های توصیه‌گر به نام FSR ارائه شده است. در روش پیشنهادی FSR از پارامترهای اهمیت و محبوبیت استفاده شده است. به منظور ارزیابی روش پیشنهادی، در ابتدا، روش پیشنهادی فازی FSR با همان روش در حالت غیر فازی (RSR) مورد مقایسه ارزیابی قرار گرفت. براساس نتایج، مشاهده می‌شود که استفاده از منطق فازی تاثیر زیادی بر افزایش دقت معیار شباهت پیشنهادی دارد و باعث بهبود عملکرد سیستم توصیه‌گر می‌گردد. همچنین، روش پیشنهادی FSR با معیارهای شباهت مورد استفاده در سیستم‌های توصیه‌گر یعنی PIP و NHSM مقایسه گردید. براساس نتایج بدست آمده، به طور مثال در حالت K-neighbors=۱۰۰ و بر روی مجموعه داده MovieLens-1M ، مقدار F1 در روش پیشنهادی FSR نسبت به روش PIP و NHSM به ترتیب ۱۷٪ و ۲۰٪ بهبود داده شده است. همچنین، مقدار MAE در روش پیشنهادی FSR نسبت به روش NHSM، ۵٪ بهبود یافته است. نتایج بدست آمده نشان می‌دهد که معیار شباهت فازی پیشنهاد شده (FSR) نسبت به روش‌های اخیر اندازه‌گیری شباهت در سیستم‌های توصیه‌گر دارای دقت بالاتری می‌باشد و باعث بهبود عملکرد سیستم‌های توصیه‌گر می‌گردد. همچنین، برای بهبود معیار MAE پیشنهاد می‌شود که از روش سلسله مراتبی استفاده شود. در واقع به منظور بهبود نتایج می‌توان مقادیر شباهت محاسبه شده را در یک روش سلسله مراتبی مورد ارزیابی مجدد قرار داد. استفاده از روش سلسله مراتبی به نزدیک‌تر شدن مقدار شباهت پیشنهادی در هر مرتبه کمک می‌کند.

مراجع

[۳] م. م. رهنی, "ارائه چارچوبی برای توسعه ی سامانه ی پیشنهادگر محصولات در وب سایت های تجارت الکترونیک با استفاده از تکنیک های فازی," کارشناسی ار, دانشگاه شیراز –دولتی, ۱۳۹۱ .

[۱۴] ر. داوطلب, "بهبود روش‌های طبقه‌بندی و خوشه‌بندی با تکیه بر شبکه‌های FMM," کارشناسی ارشد, دانشکده فنی و مهندسی, دانشگاه بوعلی سینا ۱۳۹۰ .

[1] H. Liu, Z. Hu, A. Mian, H. Tian, and X. Zhu, "A new user similarity model to improve the accuracy of collaborative filtering," *Knowledge-Based Systems*, vol. 56, pp. 156-166, 2014.

[2] X. Yang, Y. Guo, Y. Liu, and H. Steck, "A survey of collaborative filtering based social recommender systems," *Computer Communications*, vol. 41, pp. 1-10, 2014.

[4] J. Han, J. Pei, and M. Kamber, *Data mining: concepts and techniques*: Elsevier, 2011.

[5] H. J. Ahn, "A new similarity measure for collaborative filtering to alleviate the new user cold-starting problem," *Information Sciences*, vol. 178, pp. 37-51, 2008.

[6] T. Mahara, "A new similarity measure based on mean measure of divergence for collaborative filtering in sparse environment," *Procedia Computer Science*, vol. 89, pp. 450-456, 2016.

[7] N. Polatidis and C. K. Georgiadis, "A multi-level collaborative filtering method that improves recommendations," *Expert Systems with Applications*, vol. 48, pp. 100-110, 2016.

[8] N. Polatidis and C. K. Georgiadis, "A dynamic multi-level collaborative filtering method for improved recommendations," *Computer Standards & Interfaces*, vol. 51, pp. 14-21, 2017.

[9] T. Arsan, E. Köksal, and Z. Bozkus, "Comparison of Collaborative Filtering Algorithms with Various Similarity Measures for Movie Recommendation," *International Journal of Computer Science, Engineering and Applications (IJCSEA)*, vol. 6, pp. 1-20, 2016.

[10] J. Xu, K. Johnson-Wahrmann, and S. Li, "The development, status and trends of recommender systems: a comprehensive and critical literature review," in *Proceedings of International Conference Mathematics and Computers in Science and Industry*, pp. 117-122, 2014.

[11] A. Abbas, L. Zhang, and S. U. Khan, "A survey on context-aware recommender systems based on computational intelligence techniques," *Computing*, vol. 97, pp. 667-69 , 2015.

[12] J. Mancera, M. T. Nguyen, and E. Portmann, "A Fuzzy-Based Recommender System: Case Börsenspiel for Swiss Universities," in *Applying Fuzzy Logic for the Digital Economy and Society*, ed: Springer, pp. 185-207, 2019.

[13] J. Lu, Q. Shambour, Y. Xu, Q. Lin, and G. Zhang, "a web-based personalized business partner recommendation system using fuzzy semantic techniques," *Computational Intelligence*, vol. 29, pp. 37-69, 2013.

[15] A. Panda and M. Pal, "A study on pentagonal fuzzy number and its corresponding matrices," *Pacific Science Review B: Humanities and Social Sciences*, vol. 1, pp. 131-139, 2015.

[16] D. Chakraborty and D. Guha, "Addition two generalized fuzzy numbers," *International Journal of Industrial Mathematics*, vol. 2, pp. 9-20, 2010.

[17] H. A. Khorshidi and S. Nikfalazar, "An improved similarity measure for generalized fuzzy numbers and its application to fuzzy risk analysis," *Applied Soft Computing*, 2016.

[18] S.-J. Chen and S.-M. Chen, "Fuzzy risk analysis based on measures of similarity between interval-valued fuzzy numbers," *Computers & Mathematics with Applications*, vol. 55, pp. 1670-1685, 2008.

[19] S.-M. Chen and J.-H. Chen, "Fuzzy risk analysis based on ranking generalized fuzzy numbers with different heights and different spreads," *Expert systems with applications*, vol. 36, pp. 6833-6842, 2009.

[20] K. U. Madhuri, S. S. Babu, and N. R. Shankar, "Fuzzy risk analysis based on the novel fuzzy ranking with new arithmetic operations of linguistic fuzzy numbers," *Journal of Intelligent & Fuzzy Systems*, vol. 26, pp. 2391-2401, 2014.

[21] M. Ayub, M. A. Ghazanfar, Z. Mehmood, T. Saba, R. Alharbey, A. M. Munshi and M. A. Alrige, "Modeling user rating preference behavior to improve the performance of the collaborative filtering based recommender systems," *PloS one*, vol. 14, p. e0220129, 2019.

[22] M. Ayub, M. A. Ghazanfar, M. Maqsood, and A. Saleem, "A Jaccard base similarity measure to improve performance of CF based recommender systems," in *2018 International Conference on Information Networking (ICOIN)*, pp. 1-6 , 2018.

[23] J. Bobadilla, A. Hernando, F. Ortega, and J. Bernal, "A framework for collaborative filtering recommender systems," *Expert Systems with Applications*, vol. 38, pp. 14609-14623 , 2011.

[24] M. Y. H. Al-Shamri, "User profiling approaches for demographic recommender systems," *Knowledge-Based Systems*, vol. 100, pp. 175-187, 2016.

زیرنویس‌ها:

[1] https://www.facebook.com/
[2] https://twitter.com/
[3] https://www.amazon.com/
[4] https://www.ebay.com/
[5] Information overload
[6] Hyung Jun Ahn
[7] Sparsity
[8] Suryakant
[9] Mean Measure of Divergence
[10] Polatidis and Georgiadis
[11] Taner Arsan at al
[12] Fuzzy Similarity Measure For Recommender System(FSR)
[13] Significance
[14] Popularity
[15] Mean squared difference(MSD)
[16] Recommender System
[17] Content-based
[18] Collaborative filtering-based
[19] Knowledge-based
[20] Hybrid
[21] Cold start
[22] Sparsity
[23] Scalability
[24] Pearson Correlation Coefficient(PCC)
[25] Traditional Cosine

[26] Proximity
[27] Impact
[28] Popularity
[29] Proximity
[30] Significance
[31] Singularity
[32] Significance
[33] Popularity
[34] Datasets
[35] http://www.movielens.umn.edu
[36] University of Minnesota
[37] Recognition Rate
[38] Sensitivity
[39] Recall
[40] Specificity
[41] Precision
[42] Mean Absolute Error
[43] Root Mean Squared Error
[44] Normalized Mean Absolute Error
[45] k-fold cross-validation
[46] Real Similarity Measure For Recommender System(RSR)