



هفتمین کنگره مشترک سیستم‌های فازی و هوشمند ایران

همایش کنفرانس سیستم‌های فازی و همایش کنفرانس سیستم‌های هوشمند

۱۱-۹ بهمن ۱۳۹۷ بجنورد، ایران



## FNHSM\_HRS: Hybrid recommender system using fuzzy clustering and heuristic similarity measure

Mostafa Khalaji\*, Faculty of Computer Engineering, K. N. Toosi University of Technology, Tehran, Iran, Email: [Khalaji@email.kntu.ac.ir](mailto:Khalaji@email.kntu.ac.ir), [ORCID: 0000-0002-5019-1824](https://orcid.org/0000-0002-5019-1824)  
Chitra Dadkhah, Faculty of Computer Engineering, K. N. Toosi University of Technology, Tehran, Iran

### Abstract

Nowadays, Recommender Systems have become a comprehensive system for helping and guiding users in a huge amount of data on the Internet. Collaborative Filtering offers to active users based on the rating of a set of users. One of the simplest and most comprehensible and successful models is to find users with a taste in recommender systems. In this model, with increasing number of users and items, the system is faced to scalability problem. On the other hand, improving system performance when there is little information available from ratings, that is important. In this paper, a hybrid recommender system called FNHSM\_HRS which is based on the new heuristic similarity measure (NHSM) along with a fuzzy clustering is presented. Using the fuzzy clustering method in the proposed system improves the scalability problem and increases the accuracy of system recommendations. The proposed system is based on the collaborative filtering model and is partnered with the heuristic similarity measure to improve the system's performance and accuracy. The evaluation of the proposed system based results on the MovieLens dataset carried out the results using MAE, Recall, Precision and Accuracy measures Indicating improvement in system performance and increasing the accuracy of recommendation to collaborative filtering methods which use other measures to find similarities.

Conference: 7th Iranian Joint Congress on Fuzzy and Intelligent Systems, 18th Conference on Fuzzy Systems and 17th Conference on Intelligent Systems, Bojnord, Iran, University of Bojnord, p.p. 562-568, January 2019.

## FNHSM\_HRS: سیستم توصیه‌گر ترکیبی مبتنی بر خوشه‌بندی فازی و معیار شباهت اکتشافی

مصطفی خلجی\*، دانشکده مهندسی کامپیوتر، دانشگاه صنعتی خواجه نصیرالدین طوسی، تهران  
چیترا دادخواه، دانشکده مهندسی کامپیوتر، دانشگاه صنعتی خواجه نصیرالدین طوسی، تهران



**چکیده:** امروزه سیستم‌های توصیه‌گر به یک سیستم فراگیر برای راهنمایی و هدایت کاربران در حجم عظیمی از داده در اینترنت، تبدیل شده است. پالایش همکارانه که پیشنهادات به کاربر فعال را براساس امتیازدهی مجموعه‌ای از کاربران ارائه می‌دهد، یکی از مدل‌های ساده و قابل درک و موفق برای پیدا کردن افراد هم سلیقه در سیستم‌های توصیه‌گر است. در این مدل، با افزایش تعداد کاربران و اقلام، سیستم دچار مشکل مقیاس پذیری می‌شود. از طرفی دیگر بهبود عملکرد سیستم در مواقعی که اطلاعات کمی از امتیازات در دسترس داریم، امری مهم است. در این مقاله یک سیستم توصیه‌گر ترکیبی به نام FNHSM\_HRS که مبتنی بر معیار شباهت اکتشافی (NHSM) به همراه خوشه‌بندی فازی است، ارائه شده است. استفاده از روش خوشه‌بندی فازی در سیستم پیشنهادی باعث بهبود پذیری مساله مقیاس پذیری گشته و دقت پیشنهادات سیستم را افزایش می‌دهد. سیستم پیشنهادی مبتنی بر مدل پالایش همکارانه بوده و با استفاده از معیار شباهت اکتشافی، عملکرد و صحت سیستم را ارتقا می‌دهد. ارزیابی نتایج سیستم پیشنهادی بر روی مجموعه داده MovieLens صورت گرفته، نتایج ارزیابی با استفاده از معیارهای MAE، Accuracy، Precision، Recall بیانگر بهبود کارایی سیستم و افزایش دقت پیشنهادات نسبت به روش‌های پالایش همکارانه‌ای که از معیارهای دیگری برای پیدا کردن شباهت استفاده می‌نمایند، می‌باشد.

**واژه‌های کلیدی:** سیستم‌های توصیه‌گر، پالایش همکارانه، خوشه‌بندی فازی، معیار شباهت اکتشافی

## ۱. مقدمه

با افزایش اطلاعات در اینترنت و فضای مجازی و خریدهای اینترنتی و تعاملات کاربران با یکدیگر، سیستم‌های توصیه‌گر (Recommender Systems) جهت هدایت کاربران به سمت سلاقی یا نیازهایی که دارند، در بیست سال اخیر و به ویژه در دهه‌ی اول قرن بیست و یکم مورد مطالعه قرار گرفته‌اند و پژوهش‌های بسیاری در این زمینه انجام شده است [۱]. سیستم‌های توصیه‌گر براساس نحوه پیشنهاد دهی به مدل‌های مختلفی دسته بندی می‌شوند که یکی از مهم ترین آنها مبتنی بر پالایش همکارانه (Collaborative Filtering) است. عملکرد این روش به این طریق است که براساس شباهت بین کاربران یا اقلام، پیش بینی و پیشنهاد دهی به کاربر فعال را انجام می‌دهد. از این رو شامل تکنیک‌های مبتنی بر حافظه و مبتنی بر مدل است [۲]. مدل مبتنی بر حافظه در ابتدا شباهت بین کاربران را محاسبه و پس از آن کاربران مشابه به کاربر فعال را به عنوان کاربران همسایه انتخاب کرده و در پایان پیشنهادات را براساس این کاربران به کاربر فعال ارائه می‌کند.

پالایش همکارانه مبتنی بر مدل در ابتدا یک مدلی از رفتار کاربران را ایجاد می‌کند و پس از آن امتیازات اقلام مشاهده نشده را براساس آن مدل پیش بینی می‌نماید. شروع سرد (Cold Start) و مقیاس پذیری (Scalability) مشکلاتی در سیستم‌های توصیه‌گر هستند. هنگامی که سیستم توصیه‌گر با کمبود اطلاعات (امتیازات) از سوی سوابق کاربر مواجه می‌شود، مشکل شروع سرد از نوع کاربر جدید در سیستم بوجود می‌آید و در صورتیکه قلم جدیدی وارد سیستم شود از آنجایی که کاربران سیستم این قلم را قبلاً مشاهده ننموده‌اند، مشکل شروع سرد برای قلم جدید رخ خواهد داد. از سوی دیگر با افزایش چشمگیر تعداد کاربران در فضای مجازی، مشکل مقیاس پذیری برای سیستم‌های توصیه‌گر رخ می‌دهد که می‌توان گفت سیستم عملکرد و دقت خود را تا حدی از دست خواهد دهد. سیستم توصیه‌گر از نوع پالایش همکارانه مدل‌های بسیار ساده و قابل فهم و به راحتی قابل پیاده سازی می‌باشند.

در این مقاله تمرکز اصلی سیستم توصیه‌گر پیشنهادی بر روی عملکرد سیستم با استفاده از ترکیبی از روش‌های مبتنی بر حافظه و مبتنی بر مدل جهت رفع مشکل مقیاس پذیری است. هسته اصلی

\* نویسنده مسئول

آدرس پست الکترونیک: [Khalaji@email.kntu.ac.ir](mailto:Khalaji@email.kntu.ac.ir)

Bellogin et.al روش‌هایی را برای بهبود عملکرد سیستم‌های توصیه‌گر معرفی کردند که از روش‌های Herlocker's weighting و McLaughlin's weighting جهت تعیین کاربرانی که از نظر سلیقه به کاربر فعال نزدیک بودند را انتخاب می‌کرد [۷]. همچنین Haifeng et.al یک معیار شباهت اکتشافی با نام NHSM ارائه کردند که سه جنبه مجاورت<sup>۵</sup>، اثر شدید<sup>۶</sup> و محبوبیت<sup>۷</sup> امتیازات کاربران را در هنگام انتخاب کاربران همسایه برای کاربر فعال مد نظر داشت. از سوی دیگر یک معیار ترکیب شده از Jaccard و Mean Squared Difference توسط Bobadila et.al و همکارانش ارائه گردید [۸].

بکارگیری روش‌های خوشه‌بندی برای گروه‌بندی کاربران مشابه از نظر سلیقه کمک شایانی به حل مسائل مقیاس‌پذیری می‌کند. از این رو Koochi et.al با بکارگیری خوشه‌بندی فازی و روش فازی‌زادایی<sup>۸</sup> Max، به صورت تخصیص کاربران به تمام خوشه‌ها با درجه عضویت متفاوت و استفاده از معیار شباهت Pearson برای پیدا کردن نزدیک‌ترین همسایه، نشان دادند که عملکرد سیستم آنها نسبت به استفاده از روش‌های K-Means و SOM بهبود بخشیده شده است [۹].

### ۳. سیستم FNHSM\_HRS

در شکل شماره ۱ ساختار سیستم پیشنهادی FNHSM\_HRS نشان داده شده است. سیستم FNHSM\_HRS دارای دو بخش برون خط و برخط است که در بخش برون خط مدل سیستم براساس اطلاعات داخل ماتریس امتیاز دهی کاربر-قلم آموزش داده می‌شود. در بخش برخط پیشنهادات سیستم براساس مدل بدست آمده در مرحله برون خط برای کاربر فعال ارائه می‌گردد.

#### ۱.۳. بخش برون خط

روش‌های سنتی سیستم‌های توصیه‌گر، تنها به پیدا کردن کاربرانی که با کاربر فعال هم سلیقه هستند به انواع روش‌های معیار شباهت، اکتفا می‌کردند که برای مواقعی که ماتریس امتیازات کاربران، تنگ<sup>۹</sup> است، کارا نمی‌باشد. از سوی دیگر اکثر این معیارها از مشکل

این سیستم در ابتدا از روش خوشه‌بندی فازی<sup>۲</sup> و الگوریتم‌های پالایش همکارانه مشتق شده است که از معیار شباهت اکتشافی که در مرجع [۳] با نام NHSM<sup>۳</sup> معرفی شده است، جهت تعیین کاربران همسایه با کاربر فعال استفاده می‌نماید. استفاده از روش‌های خوشه‌بندی در این نوع سیستم‌های توصیه‌گر باعث می‌شود تا کاربرانی که از نظر سلیقه به یکدیگر دارای شباهت‌هایی هستند در گروه‌های خاصی خوشه‌بندی شوند و پس از آن عملیات پیش‌بینی و پیشنهاد دهی اقلام در هر خوشه به صورت مجزا صورت گیرد و همین امر باعث بهبود عملکرد سیستم توصیه‌گر از نظر زمان و دقت می‌شود.

ساختار مقاله به این صورت می‌باشد که در بخش دوم به مروری بر کارهای محققین می‌پردازیم. در بخش سوم سیستم توصیه‌گر پیشنهادی FNHSM\_HRS<sup>۴</sup> معرفی می‌شود. بخش چهارم، بخش ارزیابی و نتایج پیاده‌سازی می‌باشد که در آن به نتایج آزمایشات و مقایسه با روش‌های دیگر پرداخته می‌شود. در نهایت در بخش آخر نیز نتیجه‌گیری ارائه می‌گردد.

## ۲. مروری بر کارهای محققین

سیستم‌های توصیه‌گر در ابتدا توسط Goldberg et.al معرفی شدند [۴]. پس از آن چندین روش براساس نحوه پیشنهاد دهی ارائه شدند که یکی از آنها روش پالایش همکارانه است. پالایش همکارانه به عنوان یکی از محبوب‌ترین روش پیشنهاد دهی خصوصی سازی شده برای کاربران است که در بسیاری از حوزه‌ها مورد استفاده قرار می‌گیرد. این روش براساس یک سری معیارهای شباهت و یکسری مدل‌های از پیش تعریف شده، عملیات خود را انجام می‌دهند. اگرچه این روش از مشکلاتی از قبیل شروع سرد، پراکندگی داده و مقیاس‌پذیری رنج می‌برد اما فهم و پیاده‌سازی آنها بسیار راحت است و از مدل‌های پایه در سیستم‌های توصیه‌گر می‌باشند.

برای بهبود عملکرد سیستم، بسیاری از پژوهشگران انواع روش‌های معیار شباهت را معرفی کرده‌اند. این معیارها شامل معیار شباهت Pearson، Cosine، PIP و غیره [۵ و ۶] است.

<sup>۲</sup> Fuzzy C-Means

<sup>۳</sup> New Heuristic Similarity Measure

<sup>۴</sup> Fuzzy NHSM Hybrid Recommender System

<sup>۵</sup> Proximity

<sup>۶</sup> Impact

<sup>۷</sup> Popularity

<sup>۸</sup> Defuzzification

<sup>۹</sup> Sparse

۲.۳. بخش بر خط

در این بخش، کاربر فعال که قرار است سیستم پیشنهاداتی در رابطه با فیلم‌های مشاهده نشده به او ارائه نماید، وارد سیستم شده و براساس سوابق خود کاربر فعال، عملیات محاسبات نزدیک ترین همسایه در خوشه کاربر فعال توسط معیار شباهت NHSM صورت میگیرد. این معیار دارای دو ضریب اصلی است که در رابطه (۱) مشخص شده است.

$$NHSM\_Sim(u, v) = JPSS\_Sim(u, v) \cdot URP_{Sim(u, v)} \quad (1)$$

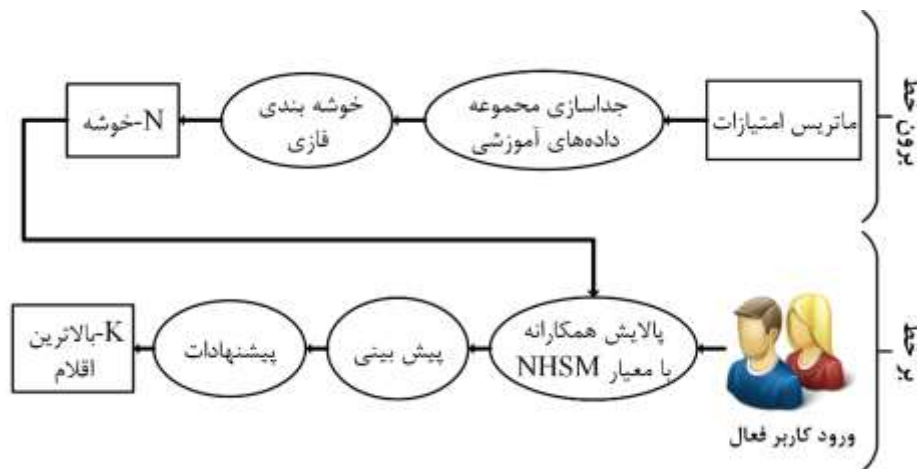
برای محاسبه معیار شباهت NHSM\_Sim باید در ابتدا معیار شباهت JPSS\_Sim محاسبه شود که خود نیز از دو معیار شباهت مشتق شده است که در روابط (۲) و (۳) ذکر شده است.

$$JPSS\_Sim(u, v) = PSS\_Sim(u, v) \cdot Jaccard'_{Sim(u, v)} \quad (2)$$

$$Jaccard'_{Sim(u, v)} = \frac{|I_u \cap I_v|}{|I_u| \times |I_v|} \quad (3)$$

پیچیدگی زمانی رنج می‌برند. برای حل این مشکلات، با خوشه‌بندی کردن کاربران در خوشه‌های متفاوت در سیستم پیشنهادی، باعث تسریع در روند عملکرد سیستم شده و بیانگر این مطلب می‌باشد که کاربران موجود در هر خوشه از نظر سلیقه باهم دارای شباهت‌هایی هستند.

در این بخش ماتریس امتیازات به عنوان ورودی به سیستم توصیه‌گر داده شده است؛ ۸۰ درصد ماتریس امتیازات به عنوان مجموعه داده آموزشی انتخاب و با استفاده از روش خوشه‌بندی فازی کاربران سیستم به تعدادی خوشه تقسیم می‌شوند. در سیستم FNHSM\_HRS با توجه به پیشنهادات محققین [۹]، ۳ خوشه در نظر گرفته شده است. الگوریتم خوشه‌بندی فازی با تخصیص درجه عضویت متفاوت به کاربران در هر خوشه، تاثیر کاربران در هر خوشه را مشخص می‌نماید. برای رتبه‌بندی درجه کاربران و انتخاب کاربران در مرحله پیش بینی سیستم از روش فازی زدایی درجات مبتنی بر مرکز ثقل (COG) استفاده شد.



شکل شماره ۱: سیستم پیشنهادی FNHSM\_HRS

$$Significance(r_{u,p}, r_{v,p}) = \frac{1}{1 + \exp(-|r_{u,p} - r_{med}| \cdot |r_{v,p} - r_{med}|)} \quad (6)$$

$$Singularity(r_{u,p}, r_{v,p}) = 1 - \frac{1}{1 + \exp(-\frac{r_{u,p} + r_{v,p} - \mu_p}{2})} \quad (7)$$

$r_{u,p}$  امتیاز قلم  $p$  توسط کاربر فعال  $u$  و  $r_{v,p}$  امتیاز قلم  $p$  توسط کاربر  $v$  است.  $r_{med}$  هم میانگین امتیاز بین امتیازاتی که کاربر موظف است برای امتیازدهی به یک قلم وارد کند، می‌باشد. ماتریس امتیازدهی در سیستم توصیه‌گر FNHSM\_HRS دارای محدوده امتیازات بین عدد ۱ تا ۵ است که متوسط آن ۳ می‌باشد.  $\mu_p$  هم میانگین امتیازات قلم  $p$  توسط کاربران است.

$|I_u \cap I_v|$  بیانگر تعداد اقلام مشابه ای است که کاربر  $u$  و  $v$  مشاهده نموده اند و  $|I_u|$  بیانگر تعداد اقلامی که کاربر فعال ( $u$ ) امتیاز داده است و  $|I_v|$  بیانگر تعداد اقلامی که کاربر دیگر ( $v$ ) امتیاز داده است. معیار  $PSS\_Sim$  از طریق رابطه (۴) بدست می‌آید.

$$PSS_{Sim}(r_{u,p}, r_{v,p}) = Proximity(r_{u,p}, r_{v,p}) \cdot Significance(r_{u,p}, r_{v,p}) \cdot Singularity(r_{u,p}, r_{v,p}) \quad (4)$$

$Proximity$ ،  $Significance$  و  $Singularity$  از طریق روابط (۵) و (۶) و (۷) بدست می‌آید.

$$Proximity(r_{u,p}, r_{v,p}) = 1 - \frac{1}{1 + \exp(-|r_{u,p} - r_{v,p}|)} \quad (5)$$



Accuracy, Precision و Recall طبق رابطه‌های (۱۱-۱۴) بر روی داده‌های آزمایش محاسبه شده است که در جدول شماره ۱ ماتریس اغتشاش<sup>۱۱</sup> مربوط را مشاهده می‌نمایید [۱۱].

Actual / Predicted	Negative	Positive
Negative	A	B
Positive	C	D

جدول شماره ۱: ماتریس اغتشاش [۱۱]

$$MAE = \frac{\sum_{i=1}^n |r_{u,i} - \hat{r}_{u,i}|}{n} \quad (11)$$

$$Accuracy = \frac{Correct\ Recommendation}{Total\ Possible\ Recommendation} = \frac{A + D}{A + B + C + D} \quad (12)$$

$$Precision = \frac{Correctly\ Recommended\ Items}{Total\ Recommended\ Items} = \frac{D}{B + D} \quad (13)$$

$$Recall = \frac{Correctly\ Recommended\ Items}{Total\ Useful\ Recommendations} = \frac{D}{C + D} \quad (14)$$

همانطور که قبلاً اشاره نموده‌ایم در سیستم FNHSM\_HRS تعداد خوشه‌ها را ۳ و تعداد کاربران همسایه را ۵۰ در نظر گرفتیم. سیستم ۵ بار مستقل برای هر Top-N جداگانه اجرا و نتایج ذخیره گردیده است. در سیستم FNHSM\_HRS از معیار شباهت اکتشافی استفاده شده است که در نتیجه با ۵ روش دیگری: روش Pearson، روش Cosine، روش وزن‌دار McLaughlin's weighting [۷]، روش وزن‌دار Herlocker's weighting [۷] و روش RA [۱۲] مقایسه و ارزیابی شده است که نتایج را در جداول شماره (۲) و شماره (۳) مشاهده می‌نمایید.

آخرین مرحله از رابطه (۱)، محاسبه معیار شباهت URP\_Sim است که با استفاده از رابطه (۸) بدست می‌آید.

$$URP_{Sim}(r_{u,p}, r_{v,p}) = 1 - \frac{1}{1 + \exp(-|\mu_u - \mu_v| \cdot |\sigma_u - \sigma_v|)} \quad (8)$$

$\mu_u$  میانگین امتیازات و  $\sigma_u$  انحراف معیار کاربر فعال است که طبق رابطه (۹) بدست می‌آید.

$$\sigma_u = \sqrt{\sum_{p \in I_u} (r_{u,p} - \bar{r}_u)^2 / |I_u|} \quad (9)$$

در مرحله بعدی تعداد کاربران همسایه براساس بالاترین درجه شباهت کاربران همسایه به کاربر فعال در خوشه مربوطه، مشخص و انتخاب می‌شود. از این رو با توجه به رابطه (۱۰) عملیات پیش بینی امتیازات اقلام مشاهده نشده براساس پالایش همکارانه در سیستم FNHSM\_HRS محاسبه می‌شود.

$$Predict(u, i) = \mu_u + \frac{\sum_{j=1}^m (r_{v,j} - \mu_v) \cdot NHSM_{Sim}(u, v_j)}{\sum_{j=1}^m |NHSM_{Sim}(u, v_j)|} \quad (10)$$

در اینجا  $u$ ، کاربر فعال و  $i$  فیلمی است که سیستم FNHSM\_HRS قرار است برای آن، امتیازی را پیش بینی نماید. سیستم FNHSM\_HRS برای کلیه اقلام مشاهده نشده توسط کاربر فعال این امتیازات را محاسبه نموده و  $k$  قلم برتر را پیشنهاد می‌دهد. در رابطه ۱۰،  $\mu_u$ ، میانگین امتیازات کاربر فعال و  $m$  تعداد کاربران همسایه داخل خوشه‌ی کاربر فعال است.  $NHSM_{Sim}(u, v_j)$  میزان شباهت بدست آمده کاربر فعال  $u$  با کاربر  $v_j$  است.  $r_{v,i}$  امتیاز کاربر  $v$  به فیلم  $i$  است و  $\mu_v$  میانگین امتیازات کاربر  $v$  است.

#### ۴. ارزیابی سیستم FNHSM\_HRS

داده‌های ورودی سیستم FNHSM\_HRS مجموعه داده MovieLens شامل ۹۴۳ کاربر و ۱۶۸۲ فیلم با ۱۰۰ هزار تا امتیاز کاربر به فیلم‌ها، در نظر گرفته شده است [۱۰]. بازه امتیاز دهی در این مجموعه داده از ۱ تا ۵ است که به ترتیب یک نشان دهنده مورد پسند نبودن و عدد ۵ نشان دهنده مورد پسند بودن کاربر به یک فیلم خاص می‌باشد. برای ارزیابی عملکرد سیستم، از 5-fold مجموعه داده‌ها به روش Cross-validation استفاده شده است که ۸۰ درصد داده‌ها برای آموزش و ایجاد مدل سیستم پیشنهادی و ۲۰ درصد داده‌ها برای آزمایش سیستم در نظر گرفته شده است. ارزیابی سیستم براساس معیارهای MAE،



## هفتمین کنگره مشترک سیستم‌های فازی و هوشمند ایران

جدیدترین کنفرانس سیستم‌های فازی و هوشمند ایران

۹-۱۱ بهمن ۱۳۹۷ بجنورد، ایران



نام روش	معیار ارزیابی	۵ بالاترین ارقام	۱۰ بالاترین ارقام	۱۵ بالاترین ارقام	۲۰ بالاترین ارقام	۳۰ بالاترین ارقام
FNHSM_HRS	Accuracy	۶۰,۲۰۳	۵۹,۳۷۳	۵۶,۲۰۰	۵۵,۵۵۷	۵۵,۱۵۵
	Precision	۹۱,۳۹۱	۹۱,۱۴۸	۹۰,۹۶۲	۹۰,۹۰۲	۹۰,۸۴۴
	Recall	۵۸,۶۴۳	۵۱,۰۲۹	۴۷,۵۶۰	۴۵,۴۸۶	۴۴,۳۷۹
	MAE	۰,۷۵۶	۰,۷۶۶	۰,۷۷۸	۰,۷۸۲	۰,۷۸۹
F_CF with Pearson	Accuracy	۵۹,۵۳۷	۵۶,۷۵۶	۵۵,۷۸۴	۵۵,۱۶۰	۵۴,۷۶۲
	Precision	۹۰,۶۷۴	۹۰,۳۵۲	۹۰,۲۳۰	۹۰,۱۴۷	۹۰,۰۸۲
	Recall	۵۴,۵۳۷	۵۰,۰۹۰	۴۶,۷۹۰	۴۴,۸۳۴	۴۲,۷۶۸
	MAE	۰,۷۶۸	۰,۷۷۵	۰,۷۸۲	۰,۷۸۷	۰,۷۹۴
F_CF with Cosine	Accuracy	۶۰,۰۱۰	۵۷,۲۷۵	۵۶,۰۳۵	۵۵,۰۷۳	۵۳,۶۱۰
	Precision	۸۲,۰۲۰	۸۱,۳۷۱	۸۰,۹۳۵	۸۰,۶۸۳	۸۰,۴۴۱
	Recall	۵۰,۲۷۸	۵۰,۲۰۶	۴۷,۰۴۱	۴۵,۴۲۱	۴۴,۰۰۳
	MAE	۰,۸۲۷	۰,۸۲۸	۰,۸۲۹	۰,۸۳۴	۰,۸۳۹
F_MW	Accuracy	۵۹,۵۰۷	۵۷,۷۱۷	۵۶,۱۱۳	۵۵,۷۳۴	۵۵,۰۵۶
	Precision	۹۱,۰۹۸	۹۰,۸۵۶	۹۰,۶۹۱	۹۰,۶۱۷	۹۰,۵۵۰
	Recall	۵۸,۰۳۹	۵۱,۰۱۵	۴۷,۵۷۶	۴۵,۴۳۳	۴۳,۳۸۳
	MAE	۰,۷۵۸	۰,۷۶۸	۰,۷۷۸	۰,۷۸۳	۰,۷۹۱
F_HW	Accuracy	۵۸,۱۷۲	۵۷,۰۰۲	۵۵,۹۲۹	۵۵,۴۰۶	۵۴,۹۰۴
	Precision	۹۰,۸۸۱	۹۰,۵۶۲	۹۰,۵۰۰	۹۰,۴۲۹	۹۰,۳۶۴
	Recall	۵۸,۲۲۰	۵۰,۶۰۴	۴۷,۰۷۴	۴۵,۱۳۶	۴۳,۰۴۲
	MAE	۰,۷۷۷	۰,۷۸۵	۰,۷۹۱	۰,۷۹۶	۰,۸۰۳
F_RA	Accuracy	۵۹,۸۸۷	۵۸,۶۷۴	۵۶,۲۸۹	۵۴,۶۰۱	۵۴,۰۳۵
	Precision	۸۹,۶۵۰	۸۹,۳۵۵	۸۹,۲۰۹	۸۹,۱۵۰	۸۹,۰۸۸
	Recall	۵۷,۵۱۳	۵۰,۲۹۷	۴۹,۵۵۴	۴۴,۳۸۵	۴۴,۰۶۷
	MAE	۰,۷۶۰	۰,۷۶۸	۰,۷۷۷	۰,۷۸۳	۰,۷۹۱

جدول شماره ۲: نتایج ارزیابی سیستم FNHSM\_HRS

نام روش	معیار ارزیابی	میانگین بالاترین n ها	نام روش	معیار ارزیابی	میانگین بالاترین n ها
FNHSM_HRS	Accuracy	۵۷,۲۹۷۶	F_MW	Accuracy	۵۶,۸۲۵۴
	Precision	۹۱,۰۴۹۴		Precision	۹۰,۷۶۲۴
	Recall	۴۹,۴۱۹۴		Recall	۴۹,۰۸۹۲
	MAE	۰,۷۷۴۲		MAE	۰,۷۷۵۶
F_CF with Pearson	Accuracy	۵۶,۳۹۹۸	F_HW	Accuracy	۵۶,۲۸۲۶
	Precision	۹۰,۲۹۷		Precision	۹۰,۵۴۷۲
	Recall	۴۷,۸۰۳۸		Recall	۴۸,۸۱۵۲
	MAE	۰,۷۸۱۲		MAE	۰,۷۹۰۴
F_CF with Cosine	Accuracy	۵۶,۴۰۰۶	F_RA	Accuracy	۵۶,۶۹۷۲
	Precision	۸۱,۰۹		Precision	۸۹,۳۹۰۴
	Recall	۴۷,۳۸۹۸		Recall	۴۹,۱۶۳۲
	MAE	۰,۸۳۱۲		MAE	۰,۷۷۵۸

جدول شماره ۳: میانگین Top-N برای هر روش



*Transactions on Knowledge & Data Engineering*, **6**, (2005), 734-749.

[7] A. Bellogín, P. Castells and I. Cantador, Improving memory-based collaborative filtering by neighbour selection based on user preference overlap, *Proceedings of the 10th Conference on Open Research Areas in Information Retrieval*, (2013).

[8] J. Bobadilla, F. Ortega, A. Hernando and J. Bernal, A collaborative filtering a roach to mitigate the new user cold start problem, *Knowledge-Based Systems*, **26**, (2012), 225-238.

[9] H. Koochi and K. Kiani, User based Collaborative Filtering using fuzzy C-means, *Measurement*, **91**, (2016), 134-139.

[10] J. Konstan, B. Miller, D. Maltz, J. Herlocker, L. Gordon and J. Riedl, GroupLens: a lying collaborative filtering to Usenet news, *Communications of the ACM*, **40.3**, (1997), 77-87.

[11] M. Hahsler, Developing and testing top-n recommendation algorithms for 0-1 data using recommenderlab, *NSF Industry University Cooperative Research Center for Net-Centric Software and System*, (2011).

[12] A. Javari, J. Gharibshah and M. Jalili, Recommender systems based on collaborative filtering and resource allocation, *Social Network Analysis and Mining*, **4.1**, (2014).

## ۵. نتیجه گیری

هدف اصلی سیستم‌های توصیه‌گر ارائه یک سری پیشنهادات براساس سلیقه کاربر و پیدا کردن کاربرانی که از نظر سلیقه با کاربر فعال دارای شباهت‌های زیادی است، می‌باشد. از این رو یکی از چالش‌های اصلی این سیستم‌ها دقت و عملکرد آنها با توجه به حجم زیاد اطلاعات (افزایش تعداد کاربران و اقلام سیستم) در کمترین زمان ممکن است. روش ارائه شده در سیستم FNHSM\_HRS که حاصل ترکیب روش‌های خوشه‌بندی فازی و بکارگیری معیار شباهت NHSM است، کمک به سزایی به بهبود Accuracy, Precision, Recall و MAE کرده است. نتایج آزمایشات سیستم پیشنهادی FNHSM\_HRS به صورت میانگین، MAE برابر با ۰,۷۷۴۲، Accuracy برابر با ۵۷,۲۹۷۶، Precision برابر با ۹۱,۰۴۹۴ و Recall برابر با ۴۹,۴۱۹۴ شده است که در مقایسه با سایر معیارهای شباهت، بیانگر بهبود عملکرد سیستم و افزایش دقت می‌باشد. سیستم FNHSM\_HRS مشکل مقیاس پذیری در سیستم‌های توصیه‌گر را به کمک روش خوشه‌بندی حل نموده است.

## ۶. مراجع

[1] C. Aggarwal, *Recommender systems*. Springer International Publishing, (2016).

[2] F. Cacheda, V. Carneiro, D. Fernández and V. Formoso, Comparison of collaborative filtering algorithms: Limitations of current techniques and proposals for scalable, high-performance recommender systems, *ACM Transactions on the Web (TWEB)*, **5.1**, (2011).

[3] 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*, **56** (2014), 156-166.

[4] D. Goldberg, D. Nichols, B. Oki and D. Terry, Using collaborative filtering to weave an information tapestry, *Communications of the ACM*, **35.12**, (1992), 61-70.

[5] H. Ahn, A new similarity measure for collaborative filtering to alleviate the new user cold-starting Problem, *Information Sciences*, **178.1**, (2008), 37-51.

[6] G. Adomavicius and A. Tuzhilin, Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions, *IEEE*