



FIRMACA-Fuzzy intelligent recommendation model using ant clustering algorithm for social networking

Sankar Pariserum Perumal¹ · Ganapathy Sannasi² · Kannan Arputharaj³Received: 26 May 2020 / Accepted: 3 September 2020
© Springer Nature Switzerland AG 2020

Abstract

Error rate is a challenging parameter to be met for most of the recommendation systems where there is a high probability of item movement among the data clusters. In social networking applications, it is often necessary to produce apt user recommendations where ant-based clustering techniques provide best optimal solutions to the clustering problems. However, the existing ant clustering algorithms lack in efficient local search procedures. Also, they are in need of intelligent fuzzy rule improvements for effective cluster partitioning. Therefore, a new model named Fuzzy Intelligent Recommendation Model using Ant Clustering Algorithm (FIRMACA) is proposed in this paper where a combination of biased exploration and fuzzy rules is used for ant clustering. Experimental analysis shows improvement in recommendation metrics—precision, recall, normalized discounted cumulative gain (NDCG) ($\leq 5\%$) and significant reduction ($\leq 1\%$) in misclassification error rate (MER) compared to contemporary ACO-based algorithms. The selected clusters are optimized globally and locally to pull out the best possible redefined clusters.

Keywords Fuzzy ant clustering · Fuzzy local search optimization · Pheromone density · Recommendation score · Pheromone trail matrix

1 Introduction

In the advent of recommendation systems, the influence of social networking systems to produce recommendation lists [1] in many domains including e-commerce, e-learning and telemedicine is pivotal. Recommendations made through most reliable social networks play a major role in the e-commerce systems. Moreover, the web-based recommendation systems [2] are helpful to the users for purchasing preferred items in the e-commerce applications, manufacturing the relevant products by manufacturers for sales based on the demand in the market, provision of suitable learning materials to the learners in e-learning systems and the recommendation of preferred medical service centres and hospitals in the healthcare

applications. In such a scenario, the application of soft computing techniques [3] will help in building better and effective recommendation systems.

Consumer behaviours leave a huge impact on recommendations given against online shopping. t-test trails are run by Gohary and Hanzae [4] to find the differences between male and female personalities in the purchase behaviour. Though this model is useful for the managers in business applications to set goals and choose suitable customer groups on their products, their recommendation lists are not optimal. Association rule mining (ARM)-based e-business recommendation systems [5] don't consider diversity and convergence of the solution at large. Hybrid recommendation systems [6, 7] present a mix of content-based, collaborative filtering-based approaches with other

✉ Sankar Pariserum Perumal, p.sankar@auist.net; Ganapathy Sannasi, sganapathy@vit.ac.in; Kannan Arputharaj, kannan.a@vit.ac.in | ¹Chennai, India · ²School of Computing Science and Engineering, VIT University, Chennai, India. ³School of Computing Science and Engineering, VIT Vellore, Vellore, India.



parameters like dissonance, a psychological factor, etc., but are prone to limitations such as less response to user context variations and cross-domain recommendations.

Opinions given by members of a social network community like Facebook, Twitter, YouTube, Instagram, Tumblr, etc., considered for making efficient decisions form the basis of collaborative filtering (CF) approach. It is also necessary to rank such social networking sites [8] based on user clusters and analysing their experiences. The same ranking algorithms help in finding the top-10 or top-20 items based on relevancy to the query based on interactions. By writing suitable regular expressions and deterministic finite automations designed for accepting the domain specific words, the discussions made through social networking sites need to be analysed for developing a better and reliable recommendation system [9].

Swarm intelligence techniques including ant colony optimization (ACO) and particle swarm optimization (PSO) can be used to make most optimal search from the available knowledge base. However, ACO meta-heuristics is more suitable for analysing the social networks since it can traverse the web graphs promisingly using pheromones. The recommendations are made based on the ratings given by genuine social network users who are either directly trusted by the current user or indirectly introduced by another trusted user through suitable trust propagation mechanisms [10]. Pheromone trail updates are important in ACO algorithms where radical manipulations in the rate of pheromone evaporation make the system less reliable for better recommendations.

In spite of the presence of many ant colony optimized recommendation algorithms in the literature, the quality improvement in recommendation scores and prediction accuracy efficiency are not considered in most of the existing protocols. Therefore, a new fuzzy intelligent recommendation model which uses novel global and local search optimization procedures for effective cluster partitioning is proposed in this paper. The main advantages of the proposed algorithm are the increased normalized discounted cumulative gain and reduced misclassification error rate.

2 Literature survey

There are many works carried out on the development of content-based recommendation systems in the last decade. Rules are used to develop more efficient recommendation systems. In the past, classification systems have been developed by many researchers in which accurate decisions were possible by the application of rules for classification. Ganapathy et al. [11] proposed a new pattern matching-based classification algorithm which was

developed by combining temporal rules with fuzzy logics and neural networks. They also used PSO for rule optimization so that the system optimized the number of rules extracted for participating in the decision-making process.

Bedi et al. [12] proposed a new recommendation system called ant recommendation system (ARS) based on collaborative behaviour of ants for effective top-N recommendations, but did not concentrate on local optimization process to finetune the initial set of clusters. Kao and Cheng [13] presented an algorithm named ant colony optimization for clustering (ACOC) which used both gathered pheromone and the heuristic data to overcome the limitations in traditional clustering techniques that easily converge to a local optimum solution.

Mavrouniotis and Yang [14] constructed an ant colony optimization (ACO) framework. It imparted knowledge from old environments to the pheromone trails based on various immigrants' schemes using which 3 algorithms were proposed. These algorithms were demonstrated on dynamic travelling salesman problems (DTSPs) with traffic parameters. Sobecki and Tomczak [15] made university course recommendations for students based on ACO with graphically represented problem statements with each node in it representing a decision.

Loyola et al. [16] produced an ACO algorithm for web usage pattern predictions. In their paper, they compared the trained ants in the artificial sessions loaded onto new web graphs against those in real sessions present in web log history. Their approach was able to pull out the list of important keywords that authorized the available navigational screens. Distributed systems and various caching methodologies are employed and for each element of such big data system, unique recommendation strategy is proposed [17]. Impact of memory caching in this big data platform is analysed against cassandra data.

Nadi et al. [18] first modelled the existing users' behaviours offline and used that model for online recommendations based on fuzzy ants for obtaining optimal solutions. Mocholi et al. [19] proposed a semantic ACO-based model based on multi-criteria by making use of domain ontologies. Their model was used for the automatic generation of music playlists. Sanadhya and Singh [20] calculated the trust path and trust cycle in online social networks based on ant colony optimization (ACO) where the inputs for those calculations were probabilistic trust rules and social intimacy pheromones. Wang and Pang [21] used ACO for getting approximate solutions on non-deterministic polynomial-time hard (NP-hard) problems of deriving the software products from software product lines (SPLs). They compared it against two other existing algorithms—(i) cartesian flattening and modified heuristic (FCF + M-HEU) algorithm and (ii) genetic algorithm for optimized feature selection (GAFES).

ACO algorithm [22] is also used in data preprocessing phase to deal with multi-dimensional and nonuniform databases. Xiao and Li [23] combined the continuous population-based incremental learning and the differential evolution for continuous domains using a hybrid ant colony optimization to populate the Gaussian probability density functions during evolution.

Bedi and Sharma [24] built a trust-based ant recommender system (TARS) where new users make best use of dynamic trust pheromones that recommends the popular user over a period of time. Chen et al. [25] applied ACO to the stacking ensemble construction process to generate domain-specific configurations for a real-world cost-sensitive data mining problem.

A new test probability is proposed based on the distance and the pheromone deposits in the iterative data transitions and thereby cluster prototypes among these data are identified [26]. Abdelbar et al. [27] adapted ant colony-based ACOR algorithm that is applied for HONEST (High Order Network with Exponential Synaptic links) training and proved to be more stable than gradient-based resilient propagation (R-Prop) algorithm.

SpyNetMiner system populates the major factors that govern the covert network on a larger scale using feature extraction and semantic rule converters for classification [28]. cAnt-Miner, a variation of the classical ant-miner classifier, is then designed [29] mainly for interpreting the numerical gene expression data. Fuzzy linear programming model-induced ACO algorithms enhance the practical realization of the local improvement procedures in various domains [30].

ACO is also deployed in a bimodal knuckle verification system where the chosen optimal fusion parameters are applied at each level of security [31]. Based on the resulting matching scores computed against knuckle database, two classes are defined in a fuzzy binary decision tree (FBDT) using fuzzy gini index—genuine (accept) and imposter (reject). To cope up with the concurrent changes in the artificial immune system (AIS) against that in the geo-position-based recommendation system, new process is proposed [32] for dynamic pattern detection.

Zhang et al. [33] gave a new ACO approach based on Jaccard similarity index calculation, names JCACO strategy for resolving travelling salesman problem (TSP). This approach makes use of basic ant colony system and max–min ant system for managing multi-ant colonies. For the path to converge fast, this approach made use of the pheromone reward technique, and then the appropriate population is recommended based on the calculated Jaccard similarity index and collaborative filtering. Elite reverse learning algorithm is used to get rid of the local optima.

Osche et al. [34] show that multi-agent ant colony systems utilize both success and failure pheromones for better recommendations. Cosine similarity-based item preferences and graph population per user are used in this approach for handling larger datasets.

HealthOne recommendation system [35] makes use of page rank based on ant colony decision tree and proposes an updated pheromone density review process based on Link Like, Category Entire Like, Link Hits and Category Link Hits for each user. A digital app is also developed for this system, and each health information content is ranked based on category it belongs to. K-means clustering is used for grouping similar users according to their personal health records into various levels of risks.

Multi-depot vehicle routing problem (MDVRP) [36] based on ant colony optimization used probabilistic and exact methods using simulated annealing. This method also makes use of single colony and mutual colony local optimization techniques and distributed processing for individual colonies for convergence.

For improving task allocation efficiency in agricultural wireless sensor networks, an elite immune ant colony optimization (EIACO) [37] was presented that leveraged both clone and elite operations for better task allocation management. Convergence speed of this algorithm improves along with the number of generations in average profits for temperature sensor nodes. Sankar and Chandra [38] proposed an ACIACO method for prediction of software test cases and their design based on stable tables. Their ant colony pheromone review step is dependent on parameters of visibility and desirability of the chosen path and the automated test cases cover all state transitions at least once. There are also quite a few ACO applications [39, 40] in healthcare domain to analyse CT scan images and perform feature selection for disease diagnosis.

Most of these ACO-based recommendation systems discussed above lack in adaptability of their algorithms and efficient cluster judgement for large number of ants. For addressing these issues, efficient local search optimization procedures [33, 36] are employed for a better trade-off between convergence and diversity of the algorithms and improved global optimization techniques [35] to closely monitor the significant changes in the pheromone evaporation rates that directly impact the recommendation contents. This paper focuses on enhanced global as well as local optimization procedures that improve recommendation quality with the help of updated pheromone density review process.

3 Proposed system

The proposed model consists of an offline computation model including normalized fuzzy rate matrix population, iterative fuzzy ants clustering judgements, ant colony optimization-based cluster selection and cluster pheromone density computation, and an online computation model including online user profile-based recommendation selection and pheromone density review processes as shown in Fig. 1. The proposed model enhances the pheromone density review process compared to other ACO-based algorithms and also presents an updated cluster modal function for determining the fitness of the cluster judgement records, which in turn, refines the final recommendation list with respect to the online user profile.

3.1 Normalized fuzzy rate matrix population

Input for the offline computation model is the user—social network portal rating matrix $U \times W$ where U is the

list of users who have provided their ratings and W is the list of social networking websites against which ratings are given.

Since these ratings follow a nonuniform cycle among the users, these are normalized on the total rating values in the matrix to introduce greater objectivity. Figure 2 displays the membership function where the ratings are better represented in fuzzy 2-tuples. So, the ratings present in $U \times W$ matrix are updated as fuzzy attribute values, where each of them is given a membership degree for computational accuracy.

3.2 Iterative fuzzy ants clustering solutions

Each iteration (ant) in this algorithm looks out for an arbitrary grouping of users list and such information is gathered in a global information hub named pheromone trail matrix. Also, successive iterations (ants) utilize this pheromone trail matrix down the line and generate multiple clustering judgements that keep this matrix updated subsequently. This algorithm groups U users into one of K clusters using T ants in a way that each user is not so far away from its cluster centre. For illustration purposes, a

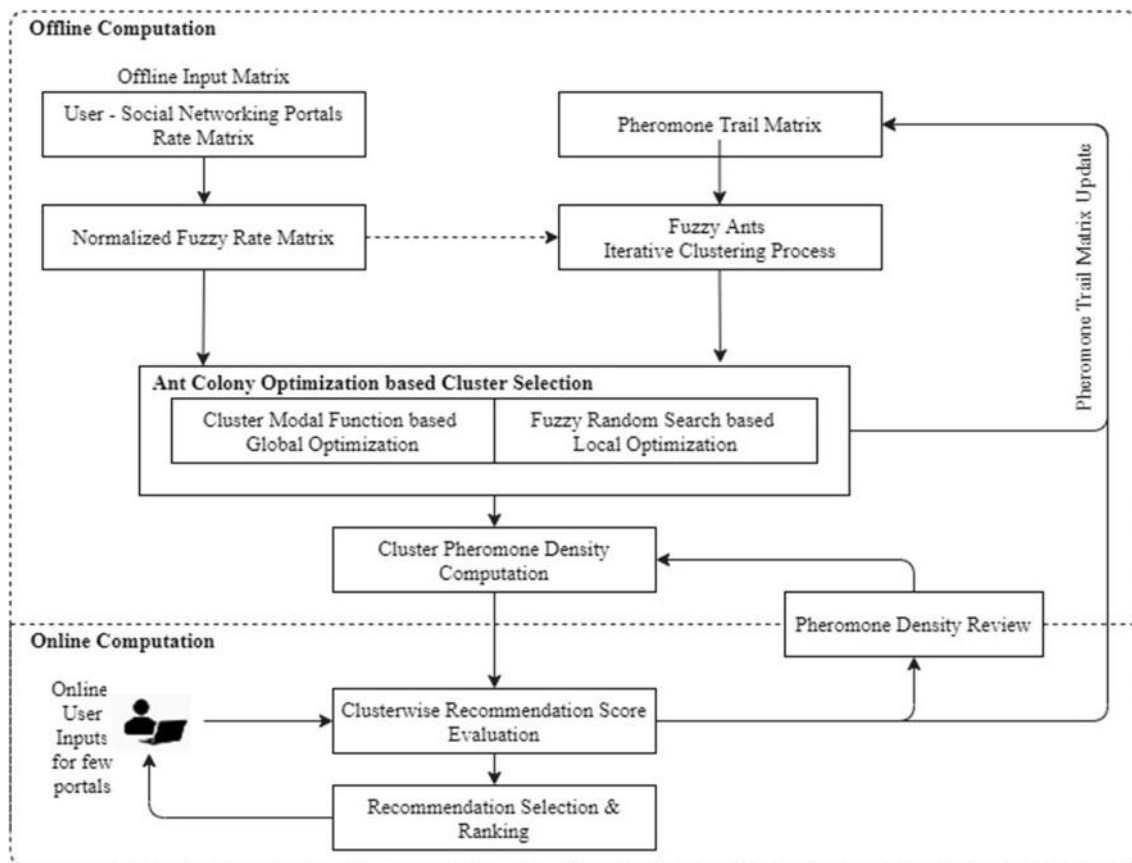


Fig. 1 Proposed FIRMACA Model Architecture

Fig. 2 Fuzzy 2-tuple representation for social networking rating scale

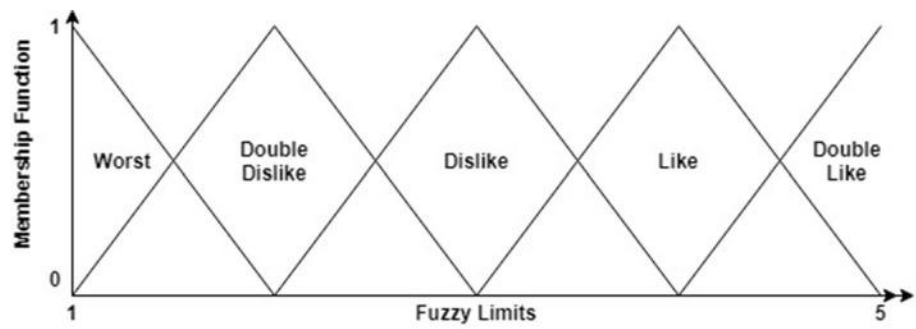


Table 1 Initialized Pheromone Trail Matrix $U \times K$ for the first ant $T=1$

Users	Clusters		
	$K=1$	$K=2$	$K=3$
U_1	0.01837	0.01251	0.01968
U_2	0.01251	0.01968	0.01837
U_3	0.01837	0.01968	0.01251
U_4	0.01251	0.01837	0.01968
U_5	0.01968	0.01837	0.01251
U_6	0.01837	0.01968	0.01251
U_7	0.01968	0.01251	0.01837
U_8	0.01251	0.01837	0.01968
U_9	0.01968	0.01837	0.01251
U_{10}	0.01837	0.01251	0.01968
U_{11}	0.01968	0.01251	0.01837

Table 2 Normalized Cumulative Pheromone Weight Ranges $U \times K$ for the first ant $T=1$

Users	Clusters		
	Range 1	Range 2	Range 3
U_1	0–0.36333	0.36333–0.61076	0.61076–1
U_2	0–0.24743	0.24743–0.63667	0.63667–1
U_3	0–0.36333	0.36333–0.75257	0.75257–1
U_4	0–0.24743	0.24743–0.61076	0.61076–1
U_5	0–0.38924	0.38924–0.75257	0.75257–1
U_6	0–0.36333	0.36333–0.75257	0.75257–1
U_7	0–0.38924	0.38924–0.63667	0.63667–1
U_8	0–0.24743	0.24743–0.61076	0.61076–1
U_9	0–0.38924	0.38924–0.75257	0.75257–1
U_{10}	0–0.36333	0.36333–0.61076	0.61076–1
U_{11}	0–0.38924	0.38924–0.63667	0.63667–1

11 × 7 user-social networking portal rating matrix is aimed to be grouped into 3 clusters using 8 ants. To start with, pheromone trail matrix of size $U \times K$ is initialized as shown in Table 1 and gets evolved with every next ant T .

For generating the iteration judgement J , each ant selects the cluster number based on fuzzy IF–THEN rules using random number function values. Fuzzy input variables for cluster number determination are (i) cluster number with higher pheromone value in pheromone trail matrix, (ii) modulus value + 1 of $i + \sum j$ over K , where $i = 1, 2, \dots, U$ and $j = 1, 2, \dots, W$ and (iii) populated random value falling in/out of the range of order $K + 1$ in the normalized cumulative pheromone trail weight. For the third fuzzy input, the pheromone trail matrix shown in Table 1 has to be normalized and cumulative weight ranges have to be determined as shown in Table 2.

Fuzzy IF–THEN rules based on these three input variables are represented in Table 3, output of which is used for determining the cluster number for the current ant iteration $T = 1$. Any random number generation method can be used to populate random real numbers between 0 and 1. In the modulus function population, $K = 3$ is used that represents the total number of clusters initially considered.

These fuzzy IF–THEN rules are applied on the populated pheromone trail matrix shown in Table 1, and Table 4 shows the resulted cluster judgement record J_1 for first ant $T = 1$. It can be specially noted that if the first two conditions in fuzzy rules set in Table 3 are met and numerically same (referring to the row numbers 1, 2, 5, 6, 9 and 10), the resultant cluster is not impacted by the third condition based on random number compatibility with the pheromone weight ranges. Also, if the first two conditions in the fuzzy rules set are not met for cluster K with high pheromone trail, the resultant cluster selection in the judgement record relies upon the third condition. If the random number falls within the weight range of index $K + 1$ in the cyclic order as shown in Table 2, then the resultant cluster in the judgement record is $K + 1$. Otherwise, if the random number does not fall within the weight range of index $K + 1$ as shown in Table 2, then the resultant cluster in the judgement record is $K + 2$ in the cyclic order. Applying these fuzzy rules set on the initial pheromone trail matrix, the cluster judgement record J_1 for the first ant $T = 1$ is formed.

Above methodology is followed by the remaining ten iterations (ants), and they also build their own cluster

Table 3 Fuzzy Rules Set generic to all T ants

Cluster with high pheromone trail	Modulus output $\text{mod}((i+j), K) + 1$	Random number to pheromone weight range compatibility	Chosen Cluster
1	1	Selected random number falls in pheromone weight range2 in Table2	1
1	1	Selected random number does not fall in pheromone weight range2 in Table2	1
1	2 or 3	Selected random number falls in pheromone weight range2 in Table2	2
1	2 or 3	Selected random number does not fall in pheromone weight range2 in Table2	3
2	2	Selected random number falls in pheromone weight range3 in Table2	2
2	2	Selected random number does not fall in pheromone weight range3 in Table2	2
2	1 or 3	Selected random number falls in pheromone weight range3 in Table2	3
2	1 or 3	Selected random number does not fall in pheromone weight range3 in Table2	1
3	3	Selected random number falls in pheromone weight range1 in Table2	3
3	3	Selected random number does not fall in pheromone weight range1 in Table2	3
3	1 or 2	Selected random number falls in pheromone weight range1 in Table2	1
3	1 or 2	Selected random number does not fall in pheromone weight range1 in Table2	2

Table 4 Cluster Judgement Record J_1 for First Ant $T=1$

Users	U_1	U_2	U_3	U_4	U_5	U_6	U_7	U_8	U_9	U_{10}	U_{11}
Cluster Judgement Record J_1	2	3	3	2	3	1	2	3	1	2	3

judgement records in the same fashion. This encourages efficient cluster grouping since it depends on the combination of fuzzy intelligent rules and random number generation that improves the quality of the generated cluster judgement records in this paper.

3.3 Best possible cluster selection based on ant colony optimization

Now that the cluster judgement records (J_1 unto J_8) for all iterations (ants) are evaluated, the next step is to optimize these records and pull out the best possible cluster judgement records among them by deploying the global as well as local optimization methods.

3.3.1 Global ant colony optimization approach

Global optimization approach presented in this paper is grounded on a primitive study [41] where quality of the judgement records is based on cluster modal function determination. This is loosely based on the Euclidean distance of the cluster centres from the corresponding individual rating matrix values. Inverse of the cluster modal function defines the fitness value of the particular cluster judgement record. This involves two more matrices—(i) attribute weight matrix and (ii) cluster centre matrix which form the basis for evaluating the cluster modal function.

Attribute weight matrix is based solely on the generated cluster judgement records. For the first ant $T=1$, attribute weight matrix values will be populated with 1 if the user is selected against its corresponding cluster based on the cluster judgement record J_1 and 0, otherwise. This matrix states clearly that any of its records will have only one column with $W_{ab} = 1$ and the rest of the columns with $W_{ab} = 0$, whereas there can be many K clusters with more than one $W_{ab} = 1$ value since there will obviously be more than one user in the same cluster, and this is given by Eq. 1 as shown below.

$$W_{ab} = \begin{cases} 1, & \text{if user 'a' is selected into cluster 'b', } a \in U, b \in K \\ 0, & \text{otherwise} \end{cases} \tag{1}$$

Cluster centre matrix M_{bc} is based on attribute weight matrix W_{ab} as well as input rating matrix X_{ac} and is given by Eq. 2.

$$M_{bc} = \frac{\sum_{a=1}^U \sum_{b=1}^K \sum_{c=1}^W W_{ab} X_{ac}}{\sum_{a=1}^U W_{ab}} \tag{2}$$

Based on these two matrices, the cluster modal function is computed as given in Eq. 3 and the cluster judgement records are sorted in the ascending order of these cluster modal function values. Lower the objective function value, higher the fitness of the corresponding cluster

judgement record. Practically, top $R=20\%$ records out of the global optimization approach are chosen to the best possible cluster partitions. So, out of the $T=8$ iterations (ants) considered in this illustration, the best possible two ($R=2$) cluster judgement records are selected.

$$f(W, M) = \sum_{a=1}^U \sum_{b=1}^K \sum_{c=1}^W W_{ab} (X_{ac} - M_{bc})^2. \quad (3)$$

3.3.2 Local ant colony optimization approach

There are several local search procedures available in the optimization field nowadays. In this paper, a novel local search procedure is proposed where fuzzy IF–THEN rule is applied on the generated random number corresponding to the cluster number to be modified in the optimal cluster judgement record, as represented in the flowchart shown in Fig. 3, where K refers to the current cluster number. It is assumed that random numbers generated this time for 11 users are 0.561322, 0.672355, 0.795321, 0.918348, 0.121878, 0.721854, 0.478254, 0.643785, 0.891213, 0.444636 and 0.005055.

Among them, only the 11th user's cluster number in the cluster judgement record R will be updated, since the generated random number only for 11th user is less than a priori probability threshold, chosen to be 0.01 for practical computations and user index $i \in U$. If modulo 3 of the iteration number (appended with 1 to avoid cluster

number = 0) is same as the current cluster number, next available cluster number is chosen in cyclic order. Otherwise, previous available cluster number is chosen in cyclic order. The entire logic is considered only when the recomputed cluster modal function value decreases compared to its original value. The pseudocode for this proposed local search procedure is represented in Fig. 4. After global and local optimization on the ant colony cluster judgements, the best possible R optimal cluster records are selected and the pheromone trail matrix is updated accordingly.

3.4 Cluster pheromone density computation

Individual cluster K 's pheromone density $PD_K(t)$ is the proportion of list of users in the corresponding cluster to the total users i.e., U , as shown in Eq. 4. This completes all four steps in the offline computation model of the proposed system.

$$PD_K(t) = \frac{\text{Count of users in cluster } K \text{ after one complete run}}{\text{Total count of users } U} \quad (4)$$

3.5 Online user profile-based recommendation selections

This is the first step in the online computational model of the proposed system where the online user's U_O profile is considered. It is normalized after consolidating the

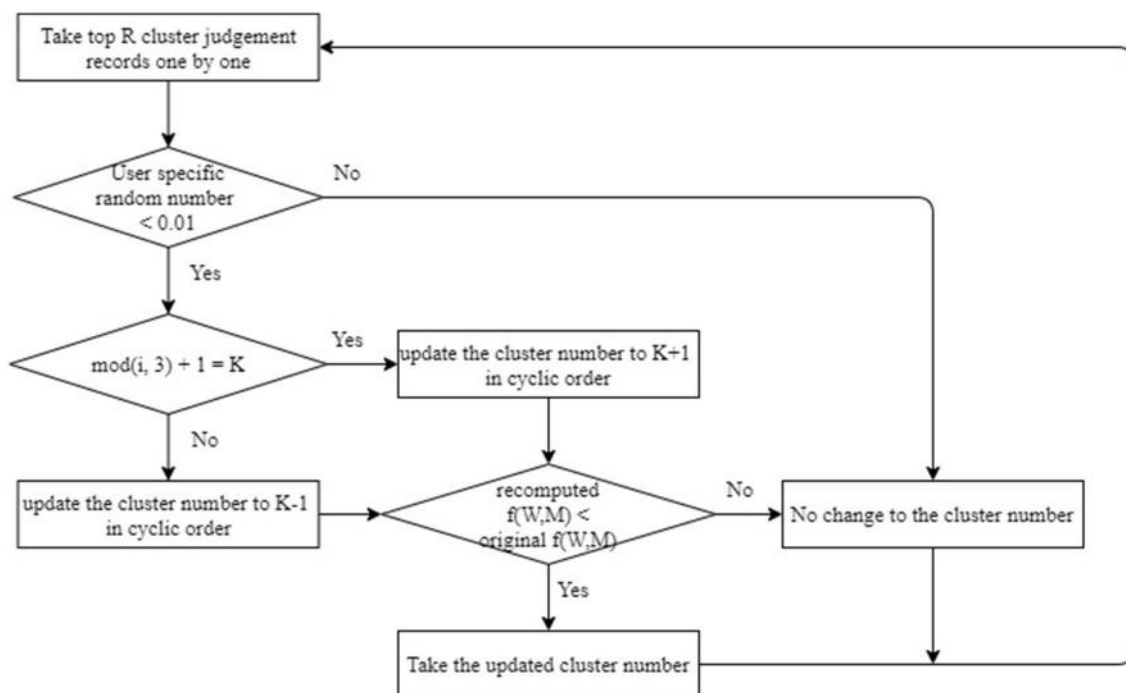


Fig. 3 Flowchart representing fuzzy local search optimization approach

Fig. 4 Local Search Optimization in pseudocode

```

for each record in top R cluster judgement records array
  if the user specific random number is less than 0.01 then
    if (current iteration number modulo 3)+1=current cluster number of judgement record then
      | update the cluster number of the judgement record from K to K+1 in cyclic order
    else
      | update the cluster number of the judgement record from K to K-1 in cyclic order
    if the recomputed cluster model function value is less than its original value then
      | consider the updated cluster number for the judgement record
    else
      | retain the same cluster number for the judgement record
  else
    | retain the same cluster number for the judgement record
return the optimal solution for top R records
    
```

individual ratings on different parameters on the social networking portals that he/she is interested to rate. These are then represented as fuzzy 2-tuples for model accuracy. These U_o ratings are then compared on the basis on cluster wise degree of similarity $DoS(U_o, M_{bc})$ against the cluster centre matrix M_{bc} where $b \in K, c \in W$ using distance measure $DM(U_o, M_{bc})$ between these two matrices as shown in Eqs. 5 and 6.

$$DM(U_o, M_{bc}) = \frac{2 * \min(U_o, M_{bc})}{n * (U_o + M_{bc})} \text{ where } b \in K, c \in W \quad (5)$$

where n refers to the total number of social networking portals rated by the user U_o .

$$DoS(U_o, M_{bc}) = \frac{1}{1 + DM(U_o, M_{bc})} \quad (6)$$

Recommendation Score $RS_K(t)$ for each cluster K_i ($i \in K$) as shown in Eq.7 is determined based on the cluster pheromone density $PD_K(t)$ and $DoS(U_o, M_{bc})$ as shown in Eqs. 4 and 6, respectively.

$$\text{Recommendation Score } RS_K(t) = \frac{PD_K^{(t)}.DoS(U_o, M_{bc})}{\sum_{i=1}^K PD_i^{(t)}.DoS(U_o, M_{bc})} \quad (7)$$

Only those recommendation scores that fall in the range, High $RS_K(t)$ —0.2 and high $RS_K(t)$ are chosen to be in the similar cluster as that of the online user profile. Finally, those social networking portals that are not already rated by the online users are recommended to him/her in the descending order of the chosen cluster centres.

3.6 Pheromone density review

This is the final step that circles the offline computation model back to the online computation model. This is the concept that differentiates general ant clustering model from other clustering algorithms since the pheromone

density is kept under review for each iteration of the model run and thereby keeping the recommendation score in every run up to the moment of the latest pheromone trail concentrations. In this step, a constant rate of evaporation 0.01 is deducted from the last pheromone density value and if the social networking portal that is considered in the current run is part of the cluster from which recommendations are made, then associated pheromone density is incremented to the original value. Pheromone density review function $PD_K(t + 1)$ is given in Eq. 8 as shown below.

$$PD_K(t + 1) = PD_K(t) * (0.09 + \Delta P) \quad (8)$$

where

$$\Delta P = \frac{P_{cc}}{P_{cc} + 1} \quad (9)$$

where P_{cc} is the pheromone density corresponding to the new cluster centre.

4 Experimental analysis and results

Fuzzy Intelligent Recommendation Model based on Ant Clustering Optimization (FIRMACA) model proposed in this paper was applied on a local dataset that contains ratings provided by 6875 students and young adults on 7 social networking portals on various rating parameters like privacy preservation, advertisements importance, content edit quality, ease of navigation, innovation, visual clarity, portal response time, etc. The raw dataset was aggregated and modified into a 6875×7 user-social networking ratings matrix as given in Sect. 3.1. The derived dataset was then normalized and represented as fuzzy 2-tuples as shown in Table 5 below.

Social networking portals represented in Table 5 are Facebook (FB), Twitter (TW), Tumblr (TM), Youtube (YT), Google + (G +), Instagram (IG) and Reddit (RD). For better understanding, we consider the same example of

pheromone trail matrix and random numbers generated in Sect. 3.2 of this paper. Based on these values and the above fuzzified user rating matrix, cluster judgement records (J_1 unto J_8) and cluster centre matrices required for the evaluation of fitness function are computed. For the cluster judgement record J_1 belonging to the first ant $T=1$, the computed cluster centre matrix M_{bc} ($b \in K, c \in W$) using Eq.2 is shown in Table 6.

Based on the attribute weight matrix and cluster centre matrix as shown above, fitness function for cluster judgement record J_1 is computed. Similarly, successive iterative ants generate their own attribute weight W_{ab} and cluster

centre M_{bc} matrices $a \in U, b \in K, c \in W$. Table 7 shows the cluster judgement records (J_1 unto J_8) for all 8 iterations (ants) in the descending order of their corresponding fitness function values (or ascending order of cluster modal function values) using Eq.3 for global optimization of the clustered ants.

User notations U_1 unto U_{11} represented in Tables 7 and 8 refer to the users U101549, U102437, U101741, U102252, U101959, U102493, U102846, U101872, U102559, U101988 and U101673, respectively. Now, for the local optimization of these cluster judgement records, top $R=2$ (J_1 and J_2) ants are chosen. Out of these two records, the same

Table 5 User-Social Networking Rating Matrix

Users	Social Networking Portals						
	FB	TW	TM	YT	G+	IG	RD
U101549	(L, 0.4)	(L, 0.1)	(D, 0.4)	(D, 0.1)	(D, 0.3)	(DD, 0.2)	(D, 0.4)
U102437	(LL, -0.4)	(L, 0.4)	(L, -0.2)	(L, -0.4)	(D, 0)	(D, 0.4)	(L, -0.2)
U101741	(L, 0.1)	(D, 0.3)	(L, -0.5)	(L, 0.3)	(D, -0.4)	(L, -0.3)	(DD, 0.4)
U102252	(L, -0.2)	(L, -0.2)	(L, 0.3)	(L, 0.4)	(D, 0.2)	(D, -0.1)	(L, 0.3)
U101959	(D, 0.2)	(L, 0.1)	(L, 0.1)	(D, 0)	(L, 0.3)	(L, 0)	(L, 0.1)
U102493	(L, 0.3)	(L, -0.5)	(L, -0.4)	(D, -0.2)	(LL, -0.5)	(LL, -0.5)	(L, -0.4)
U102846	(L, -0.4)	(L, -0.1)	(DD, 0.4)	(D, -0.1)	(L, 0.1)	(D, 0.4)	(DD, 0.3)
U101872	(D, 0.1)	(D, 0.2)	(L, 0.3)	(D, -0.4)	(L, -0.5)	(D, 0.2)	(L, 0.3)
U102559	(LL, -0.4)	(DD, 0.4)	(L, -0.1)	(L, -0.5)	(L, -0.1)	(L, 0.1)	(L, -0.1)
U101988	(L, 0.3)	(D, 0)	(LL, -0.5)	(L, -0.1)	(D, -0.2)	(DD, 0)	(LL, -0.5)
U101673	(L, 0.3)	(LL, -0.4)	(L, 0.4)	(LL, -0.5)	(DD, 0.4)	(L, -0.2)	(L, 0.4)

Table 6 Cluster Centre Matrix based on Cluster Judgement Record J_1 for First Ant $T=1$ and Weight Attribute Matrix

Clusters	Social Networking Portals						
	FB	TW	TM	YT	G+	IG	RD
K_1	(L, 0.45)	(D, -0.05)	(L, -0.25)	(D, 0.15)	(L, 0.2)	(L, 0.3)	(L, -0.25)
K_2	(L, -0.05)	(L, -0.3)	(L, -0.35)	(L, -0.425)	(D, 0.35)	(D, -0.375)	(L, -0.375)
K_3	(L, -0.14)	(L, -0.08)	(L, 0.02)	(L, -0.4)	(D, 0.16)	(L, -0.38)	(L, -0.2)

Table 7 Cluster Judgement Records (CJR) for $T=8$ Ants and 11 Users in the Descending Order of their Fitness Functions

CJR	Users List											Fitness F
	U_1	U_2	U_3	U_4	U_5	U_6	U_7	U_8	U_9	U_{10}	U_{11}	
J_1	2	3	3	2	3	1	2	3	1	2	3	2.518033
J_2	2	1	2	3	1	2	3	1	1	3	2	2.354087
J_3	3	1	2	1	3	3	1	2	3	2	1	2.282237
J_4	1	2	3	1	3	2	1	1	2	1	3	2.014713
J_5	3	1	3	1	2	3	1	2	1	3	2	1.96511
J_6	2	3	1	1	3	2	2	1	2	1	3	1.900884
J_7	1	3	3	2	1	3	1	3	2	1	2	1.774566
J_8	3	2	1	2	3	2	1	1	3	2	1	1.640243

Table 8 Best Possible Optimized Cluster Judgement Record R_1

CJR	Users List										
	U1	U2	U3	U4	U5	U6	U7	U8	U9	U10	U11
R_1	2	3	3	2	3	1	2	3	1	2	1

Table 9 Online User’s Normalized Fuzzy Profile

Online User	Social Networking Portals		
	Twitter	Google+	Instagram
U103569	(L, -0.45)	(D, 0.2)	(LL, -0.5)

example of random number generated in Sect. 3.3.2 is followed and cluster number in both records get qualified for a local update, only for the user U101673. But, as shown in the fuzzy local optimization flowchart in Fig. 3, only the first record’s (R_1) cluster number for the user U101673 is updated from 3 to 1, since its redefined cluster modal function value is found to be less than that of its corresponding original record J_1 as shown in Table 8. On the other hand, second record’s (R_2) cluster number for the same user U101673 is retained as 2 (and not updated to 1), since its redefined cluster modal function value is not less than that of its corresponding original record J_2 . So, redefined cluster judgement record R_1 is chosen to be the best possible optimized cluster partitioning of users, and it keeps evolving after every consecutive run of the model proposed in this paper.

So, it is inferred from cluster judgement record R_1 that first cluster K_1 groups the users U102493, U102559 and U101673. Second cluster K_2 groups the users U101549, U102252, U102846 and U101988 whereas the third cluster K_3 groups the users U102437, U101741, U101959 and U101872. Hence, based on Eq. 4 in Sect. 3.4, the cluster pheromone density for each of these clusters K_1, K_2 and K_3 are found to be 0.272, 0.363 and 0.363, respectively.

Now, let us move on to the online computation section. Online User U103569 has provided ratings on 3 social networking portals—Twitter, Google + and Instagram which are then normalized and the fuzzified ratings after consolidating the ratings on different parameters for each of those portals are shown in Table 9.

Based on Eqs. 5, 6 and 7 as given in the previous section, the recommendation score for each of the clusters K_1, K_2 and K_3 in the best optimal cluster judgement record R_1 are evaluated as shown in Table 10.

Clusters in the range $(0.4803-0.2) \leq$ recommendation score ≤ 0.4803 are chosen for user recommendations. From this Table 10, it is very evident that the clusters K_1 and K_2 will be chosen. Since the social networking portals ‘Twitter,’ ‘Google +’ and ‘Instagram’ (websites W_3, W_5 and W_6 in the user rating matrix $U \times W$) are already present in the

Table 10 Recommendation Score for each cluster against online user profile

Cluster	$PD_K(t)$	$DoS(U_O, M_{bc})$	$RS_K(t)$
K_1	0.272	0.578	0.4803
K_2	0.363	0.327	0.3623
K_3	0.363	0.142	0.1573

Table 11 Recommendation Rankings Compatible with the Online User Profile

Social Networking Portal	Recommendation Rank
Facebook (W_1)	1
Tumblr (W_2)	4
Youtube(W_4)	3
Reddit (W_7)	2

Table 12 Pheromone Density Review after completion of first run

Pheromone Density $PD_K(t)$	Clusters		
	K_1	K_2	K_3
Initial Status	0.272	0.363	0.363
Final Status after adding each recommendation portal in the ranking order	0.384	0.495	0.238

user profile, they are not listed in the online user recommendation ranks. Based on the cluster centres of the chosen clusters K_1 and K_2 , the portals ‘Facebook’ and ‘Reddit’ (websites W_1 and W_7 in the user rating matrix $U \times W$) are selected from cluster K_2 whereas the portals ‘Tumblr’ and ‘Youtube’ (websites W_2 and W_4 in the user rating matrix $U \times W$) are selected from cluster K_1 and ranked in the descending order of their cluster centres as shown in Table 11.

Finally, pheromone densities for all three clusters K_1, K_2 and K_3 are reviewed one by one after recommending social networking portals ‘Facebook,’ ‘Reddit,’ ‘Youtube’ and ‘Tumblr’ in the given ranking order and the final cluster pheromone density values are shown in Table 12.

It can be noted that the pheromone density has increased after adding the social networking portals suggested from the clusters K_1 and K_2 , whereas the same has decreased for cluster K_3 since there were no social networking portal recommendations from this cluster.

Fig. 5 100% Stacked Histogram for 4 Social Networking Portals against 4 Ranks

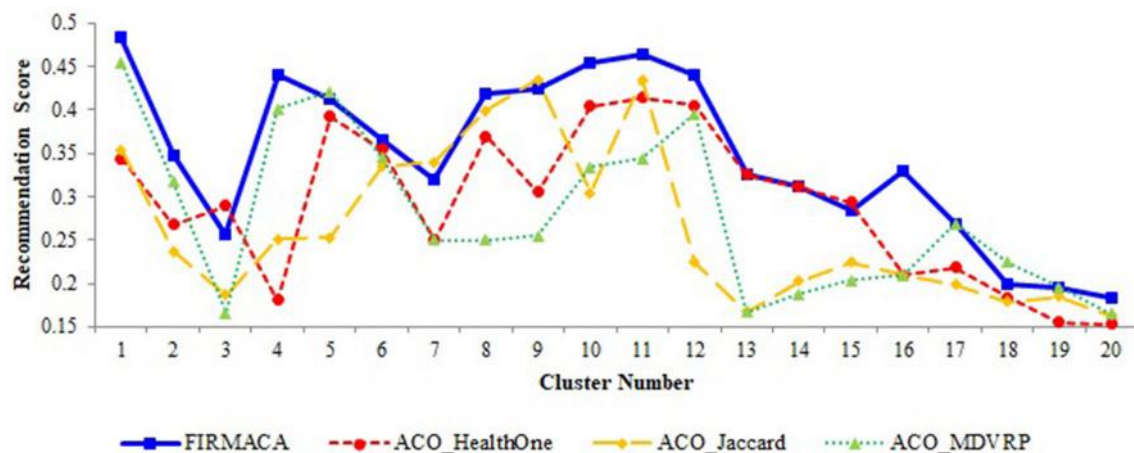
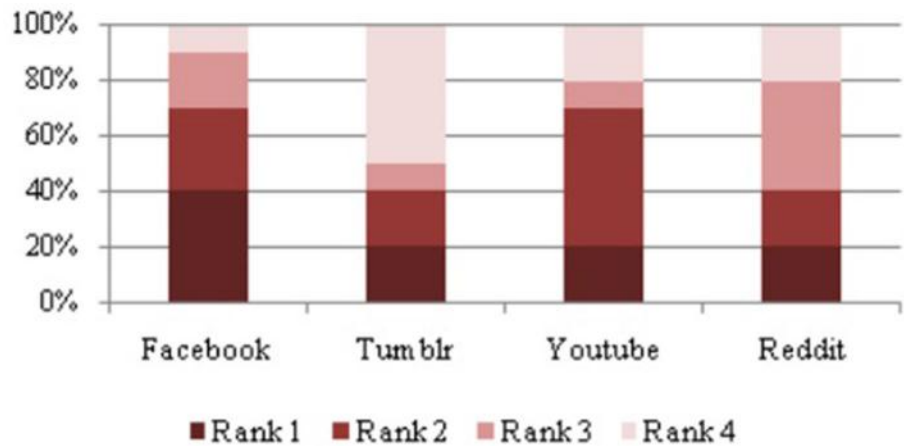


Fig. 6 Comparison of Recommendation Scores for clusters $K=1$ to 20 across 4 algorithms

FIRMACA model is repeated for 9 more runs to see the consistency of the recommendation rankings of these 4 social networking portals—'Facebook,' 'Tumblr,' 'Youtube' and 'Reddit.' The results are presented on a stacked bar chart as shown above in Fig. 5.

From the base dataset ratings and the online user profile, it is evident that 'Facebook' managed to be in Rank 1 with high proportion of 40% out of total 10 runs. 'Youtube' managed to be in Rank 2 with high proportion of 50% out of total 10 runs. 'Reddit' and 'Tumblr' were in Ranks 3 and 4 with high proportion of 40%–50%, respectively, out of total 10 runs. Regardless of the number of clusters in each run, there is no overlap among these ranks, emphasizing on the recommendation quality.

Proposed algorithm in this paper for the base dataset of 6875×11 matrix is compared with three other algorithms [33, 35, 36] based on ant colony optimization (ACO) for recommendation strategy. ACO_HealthOne [35] applied

page ranking algorithm based on ACDT (Ant Colony Decision Tree) and k -means clustering and KNN for user clusters. ACO_Jaccard [33] made use of ACO approach for TSP (Travelling Salesman Problem) and max–min ant approach with Jaccard similarity-based recommendations. ACO_MDVRP [36] hybridizes ACO with simulated annealing and deterministic local search procedure. Considering the cluster size of $K=20$, the recommendation scores for the proposed algorithm is consistently higher than those from other algorithms across all clusters as shown in Fig. 6.

Precision in this context refers to the fraction of relevant social networking portal recommendations made to the online user out of the total recommendation list as shown in Eq. 10 and recall is the ratio of retrieved relevant social networking portal recommendations over its complete recommendation list to be made, as shown in Eq. 11. Comparison of precision and recall measures of various ACO algorithms is shown in Figs. 7 and 8.

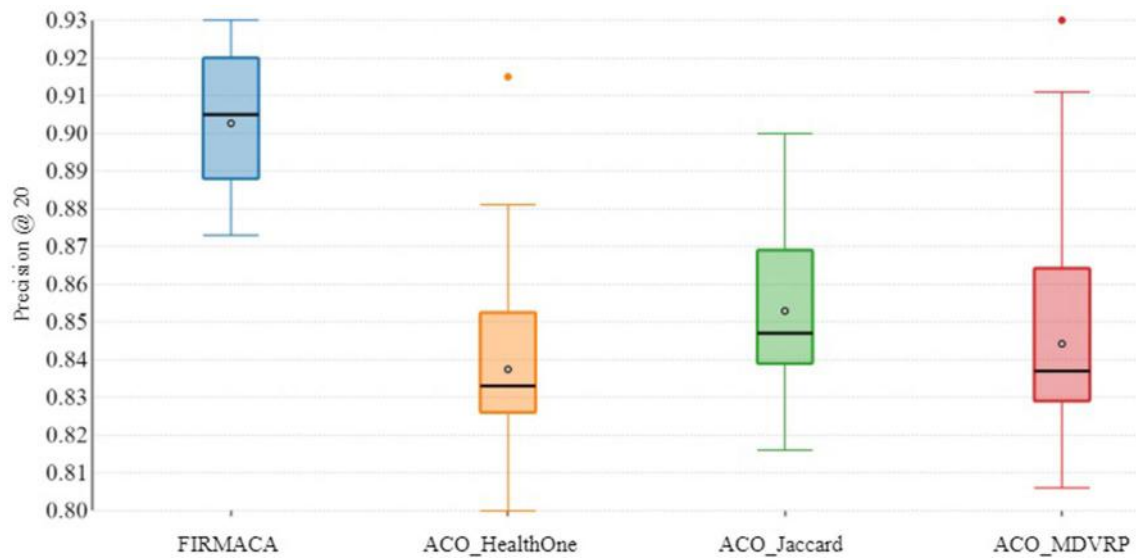


Fig. 7 Precision of 4 Ant Clustering Algorithms on Social Networking Portal Ratings Dataset

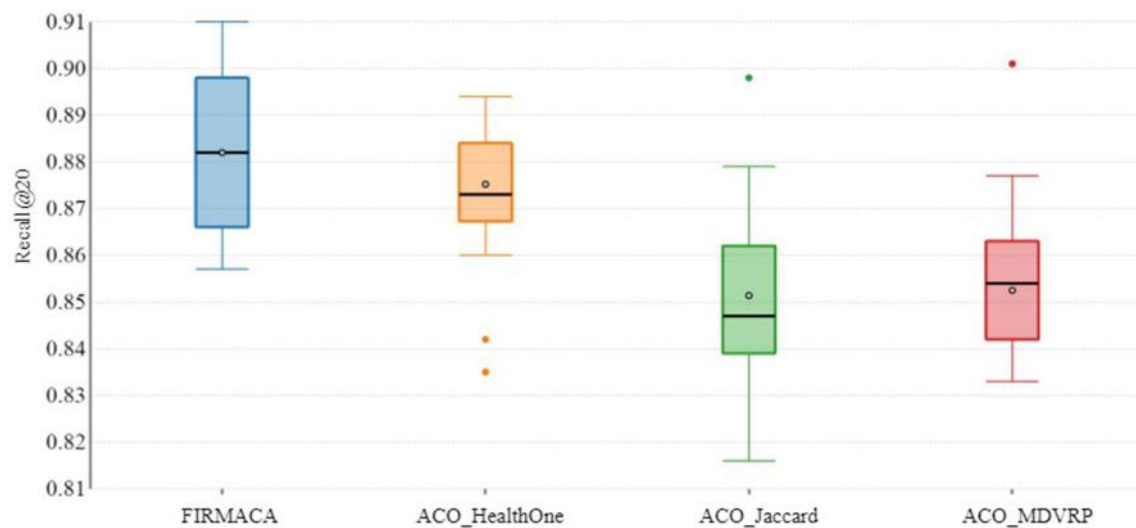


Fig. 8 Recall of 4 Ant Clustering Algorithms on Social Networking Portal Ratings Dataset

$$\text{Precision} = \frac{\text{Relevant Social Networking Portal Recommendations}}{\text{Total Social Networking Portal Recommendations made to the user}} \tag{10}$$

$$\text{Recall} = \frac{\text{Relevant Social Networking Portal Recommendations}}{\text{Total Social Networking Portal Recommendations to be made}} \tag{11}$$

Upon these recommendations thrown to the online user, two real-time metrics are evaluated to measure the prediction accuracy. They are misclassification error rate

and normalized discounted cumulative gain. Since the users partitioned into clusters stay for considerable duration of time due to enhanced global optimization technique proposed in this paper, it reduces the possibility of

Table 13 Comparison of NDCG and MER of 4 ACO algorithms on Social Networking Portal Ratings Dataset

Model	Normalized Discounted Cumulative Gain (NDCG)				Misclassification Error Rate (MER %)			
	K=10	K=13	K=17	K=20	K=10	K=13	K=17	K=20
FIRMACA	0.961	0.878	0.862	0.835	3.25	3.55	3.42	4.15
ACO_HealthOne	0.675	0.807	0.815	0.733	4.22	4.70	5.15	5.74
ACO_Jaccard	0.836	0.798	0.763	0.792	4.87	5.65	5.29	5.43
ACO_MDVRP	0.696	0.774	0.728	0.804	4.49	5.84	4.97	5.68

higher misclassification error rate to a considerable extent. Misclassification Error rate is given by Eq. 12.

$$MER = \frac{1}{U} \sum_{x=1}^U e_x \quad (12)$$

where e_x refers to the ratio of user that are not mapped to the corresponding cluster that ranges between 0 and 1. Normalized discounted cumulative gain is given by Eq. 13.

$$NDCG = \frac{DCG_u}{IDCG_u} \quad (13)$$

where the discounted cumulative gain at a specific rank position u , DCG_u is given by Eq. 14.

$$DCG_u = \sum_{i=1}^u \frac{2^{rel_i} - 1}{\log_2^{(i+1)}} \quad (14)$$

where rel_i refers to the graded relevance of the user cluster number at the rank position i and ideal discounted cumulative gain $IDCG_u$ is given by Eq. 15.

$$IDCG_u = \sum_{i=1}^{|REL|} \frac{2^{rel_i} - 1}{\log_2^{(i+1)}} \quad (15)$$

Misclassification error rate (MER) and normalized discounted cumulative gain (NDCG) of the proposed model is compared against those in other algorithms as shown in Table 13 above.

Table 13 shows that misclassification error rate of FIRMACA model is least compared to other algorithms for any number of clusters chosen after ant colony optimization. Also, NDCG is determined to be higher compared to other models, thereby proving the fact that, as and when the gain of each result gets discounted at lower ranks, cumulative gain is in upward trend in the higher ranks.

5 Conclusion

In this paper, a new fuzzy intelligent recommendation model (FIRMACA) is proposed that is based on novel global as well as local optimization techniques on ant

clusters. In this model, the recommendation scores are calculated based on pheromone density matrix and the derived cluster centre matrix for obtaining best possible cluster partitioning groups. Moreover, pheromone density is dynamically updated for every new iteration that is up to date accounting all pheromone concentrations accumulated till the latest ant agent.

From the experiments based on local ratings dataset, it is observed that the proposed FIRMACA model provides better performance on the recommendation scores compared to other ACO algorithms. The merits of FIRMACA model are:

- Better trade-off between diversity & convergence of the optimization procedure.
- Enhanced Pheromone Density Review process
- Precision and Recall metrics are higher than contemporary algorithms.
- Reduced Misclassification error rate (MER).
- Improved Normalized Discounted Cumulative Gain (NDCG)

Further works in this direction would be the betterment of global as well as local optimization methodologies with increase in cluster number for large datasets (in the lines of distributed computing) and steep decrease in cluster number for sparse datasets and different similarity measures to map the online user profile against the derived cluster centres.

Acknowledgements The authors like to acknowledge the support extended by the staffs and students of the department of Information Science and Technology, Anna University, Chennai, India on this proposed study.

Author contributions Dr. Sankar—His contribution on this proposed study was on the analysis of the existing Ant Colony Optimization based algorithms and design of the proposed system architecture and its implementation. Dr. Ganapathy—His contribution was extensively on the application of the proposal on the university-internal dataset and gauge its metrics against those on the existing algorithms. Dr. Kannan—His contribution was on the detailed review of the proposed system architecture, refining local optimization procedure and dataset curation.

Availability of data and materials The dataset used for this proposed study contain the ratings provided by 6875 students and young adults on 7 social networking portals on various rating parameters that are internal to the university campus.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

References

1. Stan J, Muhlenbach F, Llargeron C (2014) Recommender systems using social network analysis : challenges and future trends. In: Alhaji R, Rokne J (eds) Encyclopedia of social network analysis and mining. Springer, New York, NY
2. Perumal, S. P., Arputharaj, K., & Sannasi, G. (2017). Fuzzy family tree similarity based effective e-learning recommender system. In eighth international IEEE conference on advanced computing (ICoAC). pp.(146–150).
3. Washid M, Ali R (2017) Use of soft computing techniques for recommender systems: an overview. In: Ali R, Beg MMS (eds) Applications of soft computing for the web. Springer, Singapore
4. Gohary A, Hanzae KH (2014) Personality traits as predictors of shopping motivations and behaviors: a canonical correlation analysis. Arab Econ Bus J 9(2):166–174
5. Omari, A., Conrad, S., & Alci, S. (2007). Designing a well-structured E-shop using association rule mining. In 4th international IEEE conference on innovations in information technology, pp.(6–10).
6. Surendren, D., & Bhuvanewari, V. (2014). A framework for analysis of purchase dissonance in recommender system using association rule mining. In International IEEE conference on intelligent computing applications, pp.(153–157).
7. Le-wei, H., & Yu-sheng, L. (2011). Exploration in commercial websites on the association rules base on the consumers' behavior. In 6th IEEE joint international on information technology and artificial intelligence conference, pp.(42–44).
8. Koroleva, K., & Röhrler, A.B. (2012). Reducing information overload: design and evaluation of filtering & ranking algorithms for social networking sites. European conference on information systems (ECIS) Proceedings. AIS electronic library (AISeL). Paper 12.
9. Budiselic, I., Sribljic, S., & Popović, M. (2007). RegExpert: a tool for visualization of regular expressions. EUROCON The international conference on "Computer as a Tool". pp.(2387–2389).
10. Urena R, Kou G, Dong Y, Chiclana F, Herrera-Viedma E (2019) A review on trust propagation and opinion dynamics in social networks and group decision making frameworks. Inf Sci 478:461–475. <https://doi.org/10.1016/j.ins.2018.11.037>
11. Ganapathy S, Sethukkarasi R, Yogesh P, Vijayakumar P, Kannan A (2014) An intelligent temporal pattern classification system using fuzzy temporal rules and particle swarm optimization. Sadhana 39(2):283–302
12. Bedi P, Sharma R, Kaur H (2009) Recommender system based on collaborative behavior of ants. J Artif Intell 2(2):40–55
13. Y Kao K Cheng 2006 An ACO-based clustering algorithm In International workshop on ant colony optimization and swarm intelligence, Springer, Berlin
14. Mavrovouniotis M, Yang S (2013) Ant colony optimization with immigrants schemes for the dynamic travelling salesman problem with traffic factors. Appl Soft Computing 13(10):4023–4037
15. Sobacki, J., & Tomczak, J. M. (2010). Student courses recommendation using ant colony optimization. In Asian conference on intelligent information and database systems. Springer, Berlin
16. Loyola P, Roman PE, Velasquez JD (2012) Predicting web user behavior using learning-based ant colony optimization. Eng Appl Artif Intell 25(5):889–897
17. Valcarce D, Parapar J, Barreiro A (2015) A distributed recommendation platform for big data. J Univers Computer Sci 21(13):1810–1829
18. Nadi S, Saraee M, Bagheri A, Davarpanh Jazi M (2011) FARS: Fuzzy ant based recommender system for web users. Int J Computer Sci Issues 8(1):203–209
19. Mocholi JA, Martinez V, Jaen J, Catala A (2012) A multicriteria ant colony algorithm for generating music playlists. Expert Syst Appl 39(3):2270–2278
20. Sanadhya S, Singh S (2015) Trust calculation with ant colony optimization in online social networks. Procedia Computer Sci 54:186–195
21. Wang YL, Pang JW (2014) Ant colony optimization for feature selection in software product lines. J Shanghai Jiaotong Univ (Science) 19(1):50–58
22. Jiang H, Li J, Yi S, Wang X, Hu X (2011) A new hybrid method based on partitioning-based DBSCAN and ant clustering. Expert Syst Appl 38(8):9373–9381
23. Xiao J, Li L (2011) A hybrid ant colony optimization for continuous domains. Expert Syst Appl 38(9):11072–11077
24. Bedi P, Sharma R (2012) Trust based recommender system using ant colony for trust computation. Expert Syst Appl 39(1):1183–1190
25. Chen Y, Wong ML, Li H (2014) Applying ant colony optimization to configuring stacking ensembles for data mining. Expert Syst Appl 41(6):2688–2702
26. Hamlich M, Ramdani M (2013) Ant colony algorithms for data learning. Int J Appl Evolut Comput (IJAEC) 4(3):1–10
27. Abdelbar, A. M., El-Nabarawy, I., Wunch, D. C., & Salama, K. M. (2016). Ant colony optimization applied to the training of a high order neural network with adaptable exponential weights - Chapter 14. Applied artificial higher order neural networks for control and recognition. C.N.U (Ed.), Hershey, PA: pp.(362–374).
28. Karthika S, Bose S, Kannan A (2014) SpyNetMiner: an outlier analysis to tag elites in clandestine social networks. Int J Data Warehous Min (IJDWM) 10(1):32–54
29. Schaefer G (2016) Gene expression analysis based on ant colony optimisation classification. Int J Rough Sets Data Anal (IJRSDA) 3(3):51–59
30. Hudedagaddi, D. P., & Tripathy, B. K. (2018). Fuzziness in ant colony optimization and their applications. In handbook of research on modeling, analysis, and application of nature-inspired metaheuristic algorithms, Hershey, PA: IGI Global Press, pp. (363–376)
31. Kumar TV, Singh R, Kumar A (2015) Distributed query plan generation using ant colony optimization. Int J Appl Metaheuristic Computing (IJAMC) 6(1):1–22
32. Cabanas-Abascal A, García-Machicado E, Prieto-González L, de Amescua Seco A (2013) An item based geo-recommender system inspired by artificial immune algorithms. J Univers Computer Sci 19(13):2013–2033
33. Zhang D, You X, Liu S, Yang K (2019) Multi-colony ant colony optimization based on generalized jaccard similarity recommendation strategy. IEEE Access 7:157303–157317
34. Osche, Pierre-Edouard., Castagnos, Sylvain., & Boyer, Anne. (2019). From music to museum: applications of multi-objective ant colony systems to real world problems. Adaptive and learning agents workshop at AAMAS (ALA 2019), Montréal, Canada. (hal-02472009)

35. Shin M, Hwang S, Lee S, Jeong A, Kim B (2020) Healthone: personalized healthcare recommendation system. *Int J Sci Technol Res* 9(3):103–107
36. Stodola P (2020) Hybrid ant colony optimization algorithm applied to the multi-depot vehicle routing problem. *Nat Computing* 19:463–475
37. Xu M, Zhou J (2020) Elite immune ant colony optimization-based task allocation for maximizing task execution efficiency in agricultural wireless sensor networks. *J Sens*. <https://doi.org/10.1155/2020/3231864>
38. Saju Sankar S, Vinod Chandra SS (2020) An ant colony optimization algorithm based automated generation of software test cases. In: Ying T, Yuhui S, Milan T (eds) *Advances in swarm intelligence ICSI 2020*. Lecture notes in computer science, vol. 12145. Springer, Cham. https://doi.org/10.1007/978-3-030-53956-6_21
39. J, Dhalia Sweetlin., H, Khanna Nehemiah., & Arputharaj, Kannan. (2018) Computer aided diagnosis of pulmonary hamartoma from CT scan images using ant colony optimization based feature selection. *Alex Eng J* 57(3):1557–1567
40. J, Dhalia Sweetlin., H, Khanna Nehemiah., & Arputharaj, Kannan. (2017) Feature selection using ant colony optimization with tandem-run recruitment to diagnose bronchitis from CT scan images. *Computer Methods Progr Biomed* 145:115–125
41. Shelokar PS, Jayaraman VK, Kulkarni BD (2004) An ant colony approach for clustering. *Anal Chim Acta* 509(2):187–195

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.