

Application of hybrid of Fuzzy Set, Trust and Genetic Algorithm in query log mining for effective Information Retrieval

Suruchi Chawla^{*1}

Accepted : 31/12/2017 Published: 30/03/2018

Abstract: The precision of Information Retrieval (IR) System is low due to imprecise user queries as well as because of information overload on web. The Fuzzy set infers the user's information need from vague and imprecise queries and web recommender systems are used to overcome information overload problem. The performance of recommender system is still low due to data sparsity. The concept of trust is used to deal with data sparseness problem and improves the performance of recommender system. Genetic Algorithm(GA) have been applied in domain of information retrieval for optimization of web search results and effective web search. In this research hybrid of Fuzzy set, GA and Trust has been used together in query log mining for personalized web search based on using fuzzy queries for recommendation of optimal ranked set of trusted documents. Thus the use of hybrid of Fuzzy set, trust and GA together infer the user's information need from vague and imprecise user's queries and optimize the web page ranking of trusted web pages for effective personalized web search. The experimental results were analyzed statistically as well as compared with GA IR, and Fuzzy Trust based IR. Hence based on comparative analysis of results, thus hybrid of Fuzzy Set, Trust and GA shows the improvement in average precision of search results and confirms the effective personalization of web search.

Keywords: Fuzzy Set, Genetic Algorithm, Information Retrieval, Information Scent, Recommender System, Trust.

1. Introduction

Web Information is voluminous and accessed using search engines. The search engines retrieve huge set of web documents with very few are relevant to user. The user search queries are vague and imprecise due to which user's information need is difficult to infer and therefore information overload problem arises. Fuzzy set has been used to deal with vague and imprecise queries for inferring the user's information need. Recommender systems are used to overcome information overload problem but the performance of recommender system is low due to data sparseness. Trust based recommender system overcomes the data sparseness and generate the reliable recommendation. Optimization technique like Genetic Algorithm has been used to increase the effectiveness of recommender system. [1-3] In this research an approach is proposed using hybrid of Fuzzy set, Trust and Genetic Algorithm together in query log mining to generate the recommendation of relevant documents according to user's information need. An algorithm is developed for web page recommender system which uses the Fuzzy user query to represent the user's information need and cluster is selected using fuzzy query for the recommendation of optimal ranking of trusted documents for effective Information Retrieval. An algorithm uses two phases for its complete execution: Phase I (offline) and Phase II (online). In phase I, web query sessions are preprocessed and keyword vectors are generated for clustering. The W_1 (term-document matrix) is formed using tf.idf content of the clicked documents in query sessions. Fuzzy thesaurus FR is built from W_1 and W_1^T in order to create term-term correlation matrix. The genetic algorithm is executed on clicked URLs selected based on trust threshold value in a given cluster for rank optimization of

trusted web pages (Clicked URLs).

In phase II, the user input query represented by Fuzzy Set A1 is expanded with related terms based on Fuzzy thesaurus FR. The cluster is selected using fuzzy input query after expansion for the recommendation of optimal ranking of trusted documents. The user response to recommended documents is captured and user profile keyword vector is generated for expansion with related words based on Fuzzy thesaurus. The profile keyword vector after expansion selects the cluster for the recommendation of optimal ranked documents. Thus the expansion of user profile terms and recommendations of optimal ranking of trusted documents is used for effective information retrieval. Experiment was performed on the data set of clustered web user query sessions for web information retrieval and the precision of search results was improved with the use of Fuzzy, Trust and GA IR together in comparison to using only Fuzzy Trust based IR [29] and GA based IR [21].

2. Related Work

In [4, 5] modeling of user preferences had been described with fuzzy profiles. In [6] Fuzzy logics were used to summarize text for extracting the most relevant sentences. In [7] fuzzy queries on the web were processed using both document index and perception index. In [8] Fuzzy logic based personalization of newsletters was introduced. In [9] fuzzy set model was used to define fuzzy queries. In [10, 11] fuzzy relationship between query terms and documents was introduced. In [12] fuzzy IR system uses fuzzy logic to retrieve documents similar to the query document. The system was tested on Arabic documents. In [13] fuzzy logic based ranking model was introduced. In [14] trust metric was introduced in collaborative filtering and hence increased its effectiveness. In recommender systems the use of trust outperforms traditional collaborative filtering [15-18].

¹ Shaheed Rajguru College of Applied Science for Women, University of Delhi, Vasundhara Enclave, Delhi, INDIA.

* Corresponding Author Email: sur_chawla@rediffmail.com

In [19] genetic algorithm was used to solve flow shop scheduling problem. In [20] recommender system was proposed based on Genetic algorithm and k-NN algorithm. In [21] recommender system based on collaborative filtering was improved using genetic algorithm. In [22] Genetic algorithm was used to generate the optimal combination of terms for effective information retrieval. In [23] wireless sensor network energy issues and secure routing were solved using trust and genetic algorithm. In [24] hybrid of genetic algorithms and trust had been used in wireless sensor network. In [25] Genetic Algorithm Inspired Load Balancing Protocol was proposed for Congestion Control in Wireless Sensor Networks using Trust Based Routing Framework. In [26] Friend Recommender System for WBSN was proposed. Learning was done using Genetic algorithm and trust propagation was used for solving data sparseness. In [27] novel trust model was proposed which combines both peer profiling and anomaly detection. Genetic algorithm was used to detect the anomalous behaviour. In [28] trust oriented genetic algorithm (TOGA) was proposed for finding a near optimal service composition plan with QoS constraints. In [29] a method was proposed using both Fuzzy set and trust for web information retrieval. It is found that hybrid of Genetic algorithm and Trust had been applied in various domains and results shown are promising. Thus in this research, the hybrid of Fuzzy IR, GA and trust is used for developing an algorithm for web search personalization. The proposed method is novel since there has been no work done combining the advantage of FuzzyIR, Trust and GA together for effective web information retrieval.

3. Background

3.1. Information Scint

The Inferring User Need by Information Scint (IUNIS) algorithm infers the user's information need based on user's traversal path of web pages. Information Scint isc_{id} of the clicked web page CP_{id} in session i is given below in (1)(2) where n_1 is the number of distinct clicked page in session i and MQ is the total number of query sessions [30-33].

$$isc_{id} = PF \cdot IPF(CP_{id}) \times Time(CP_{id}) \forall i \in 1..MQ \forall d \in 1..n_1 \quad (1)$$

$$PF \cdot IPF(CP_{id}) = \frac{f_{CP_{id}}}{\max_{d \in 1..n_1} f_{CP_{id}}} \times \log \left(\frac{MQ}{m_{CP_d}} \right) \quad (2)$$

$PF \cdot IPF(CP_{id})$: PF is the normalized frequency $f_{CP_{id}}$ of clicked web page CP_{id} and IPF is the ratio of total number of query sessions MQ in the whole data set to the number of query sessions m_{CP_d} that contain the given page CP_d .

$Time(CP_{id})$: Time spent on a clicked web page (CP_{id}) to the total time of session i [34-42].

The QV_i session keyword vector is modelling the information need of session as given below in (3). The keyword vectors are clustered using k-means. The clusters quality are accessed using score function. (x_1, x_2) in (4) expresses the degree to which the meaning of x_1 is synonyms with meaning of x_2 [43-45].

$$QV_i = \sum_{d=1}^{n_1} isc_{id} * CP_{id} \forall i \in 1..MQ \quad (3)$$

3.2. Fuzzy Set Theory in Information Retrieval (IR) System

Fuzzy information retrieval methods based on fuzzy set handle the vague and imprecise uncertain information and infer the best results to a user query.

For implementing the Fuzzy Information Retrieval based on Fuzzy

Sets, two finite crisp sets T and $D1$ are defined, where $T1$ is the set of m_1 index terms, $T1 = \{x_1, x_2, \dots, x_{m_1}\}$ and $D1$ is a set of n documents, $D1 = \{d_1, d_2, \dots, d_n\}$

A fuzzy document—term relation $W1, W1: D1 \times T1 \rightarrow [0,1]$ where $W1(d_j, x_i)$ represents the grade of relevance of index terms x_i to individual document d_j and the membership value is based on TF.IDF. A fuzzy thesaurus FR is a relation from $T1$ to $T1$.

$$FR: T1 \times T1 \rightarrow [0,1]$$

$$FR(x_1, x_2) = \frac{1}{n} \sum_{i=1}^n W1(d_i, x_1) * W1(d_i, x_2) \forall x_1, x_2 \in T1 \quad (4)$$

$FR(x_1, x_2)$ in (4) expresses the degree to which the meaning of x_1 is synonyms with meaning of x_2 . The relationship FR uses synonymous relationship to retrieve documents relevant to user query.

The augmented query represented by Fuzzy set $B1$ on $T1$ is generated using Fuzzy input query $A1$ and Fuzzy thesaurus FR as given in eq(5, 6).

That is,

$$A1 \circ FR = B1 \quad (5)$$

where \circ is the max-min composition. [29, 46]

$$B1(x_j) = \max_{x_i \in T1} \min(A1(x_i), FR(x_i, x_j)) \forall x_j \in T1 \quad (6)$$

3.3. Trust

Trust concept has been used widely by research communities in various domains like online recommender system. A trust is modeled on social phenomena based on trust relationship between people in artificial world like web. [47, 48]

In [49] the general properties of trust in e-services were surveyed and analyzed the general properties of trust listed as follows:

- Trust is relevant to specific transactions only.
- Trust is a measurable belief.
- Trust is directed.
- Trust exists in time.
- Trust evolves in time, even within the same transaction.
- Trust between collectives does not necessarily distribute to trust between their members.
- Trust is reflexive,
- Trust is a subjective belief.

In [50] "trust" is a measure of reliability of a partner profile to deliver accurate recommendations in the past. There are two models of trust called profile and item level used for generating reliable and accurate recommendations. Thus the use of trust in collaborative recommendation process generates trust-based weighting and trust-based filtering both at either profile-level or item-level trust metrics. Thus the use of trust values reduces overall prediction error rate therefore improves the prediction accuracy.

3.4. Genetic Algorithms (GA)

Genetic Algorithm is based on natural theory of evolution for reproduction. The problem solving using genetic algorithm operates on population of chromosomes for their reproduction and evolves over generations to generate the fittest chromosome as the solution to the problem. The selection of chromosomes for reproduction is based on fitness value. The problem specific fitness function is defined and therefore used for the evaluation of

chromosomes. The chromosomes with the high fitness values are selected and evolved using genetic operators like crossover and mutation. There are numerous methods for selection of chromosomes for reproduction such as roulette-wheel selection, stochastic universal selection, ranking selection, tournament selection and truncate selection. The crossovers operators used are k-point Crossover, Uniform Crossover, Uniform Order-Based Crossover, Order-Based Crossover and Partially Matched Crossover (PMX). The mutation genetic operator is used for changing the gene at the specific position in the chromosome such as bit wise/point mutation. The replacement techniques used for replacing the current generation of population with next generation during evolution are elitist replacement, generation-wise replacement, steady-state-no-duplicates and steady-state replacement methods [51-53].

4. Proposed Approach

An algorithm is designed for web page recommender system using hybrid of Fuzzy set, Trust and Genetic algorithm and implemented in two phases Phase I and Phase II. During Phase I, the web query sessions keyword vector are generated based on Information Scent and content of clicked URL. The session keyword vectors are clustered and therefore group clicked URLs with similar content in one cluster. The trust value of clicked URLs is initialized using information scent and subsequently calculated based on frequency of clicks to URLs of the total number of recommendations during online web search. Genetic Algorithm is executed on clicked URLs in clusters selected based on trust threshold value for rank optimization using fitness function. In Phase II, the fuzzy user input query after expansion is used for the selection of cluster based on similarity measure and the selected cluster generates recommendation of optimal ranked set of relevant documents. The stepwise execution of Phase I & Phase II is given below.

Phase I: Offline Preprocessing

1. Collection and preprocessing of web query sessions.
2. Global term document relation $W1$ and the fuzzy thesaurus FR are constructed using (4).
3. Information Scent Metric of clicked URL is calculated using (1).
4. The trust of a given clicked URL d in session i is initialized using Information Scent.
5. Modeling of query keyword vector using trust and TF.IDF content as given below.

$$QV_i = \sum_{d=1}^{n1} trust_{id} * CP_{id} \quad \forall i \in 1..M \quad Q$$

$n1$ is the number of distinct clicked URLs in a given session i

6. Clustering of Query keyword vectors and clusters represented by mean keyword vector (cluster_mean).
7. Initialization of Clicked count and recommended count to zero for trust computation of each clicked URL.
8. The trust value of a given cluster i is computed as follows
 $Trust1(i) = \{ClickedURL_j : Trust1(ClickedURL_j) > \epsilon \forall j \in 1..n2 \text{ where } n2 \text{ is number of distinct clicked URLs in cluster } i\}$
9. For each j th cluster, identify the L_j list of clicked URLs where $Trust1(clicked \text{ URL}) > \epsilon$.
10. Apply the Genetic Algorithm for optimal ranking on the List L_j of clicked URLs for the generation of optimal ranking s $TOR1_j$ (Trusted Optimal Ranking j).

SubAlgorithm:

Genetic Algorithm for optimal ranking of clicked URLs

Input: List L_j , cluster mean keyword vector $clust_mean_j$, Trust threshold value ϵ , Information Scent threshold value ρ .

Output: Ranked list of clicked URLs, $TOR1_j$

Begin

1. For a given List L_j , there is a set $M1_j, M1_j \subseteq L_j$, $Length(L_j) = n3$, $Length(M1_j) = m1$ and $m1 < n3$.
2. Generation of Population P of $m1!$ chromosomes where each chromosome represent possible ranking of clicked URLs present in the set $M1_j$, $length(P) = m1!$.
3. Evaluate the Fitness function for all chromosomes in P .
 - a. For each $t1^{th}$ chromosome of size $m1$ in Population P do the following

begin

For every clicked URL k of $t1^{th}$ chromosome starting from the left most position do the following

begin

Calculate the content similarity of the clicked URL with the mean keyword vector $clust_mean_j$ and store it in matrix $L2[t1][k]$.

end

end
4. Identify those rows of matrix $L2$ where row values are arranged in descending order starting from the first column.
5. Initialize $HS1$ array to zero. Harmonic means for each of selected rows of $L2$ is calculated and stored at corresponding position in $HS1$.
6. Chromosomes with the highest value of harmonic value in $HS1$ array are selected using Tournament selection and also followed Elitism to copy the best chromosome (or a few best chromosomes) to new population without mutation and crossover.
7. Uniform order based crossover and single point mutation are applied on chromosomes obtained in step 6 but not selected using Elitism.
8. Steady-state-no-duplicates replacement policy replaced the population of parent chromosome with the child chromosomes created in step 6 and 7 in order to generate the next generation of population P .
9. Check for the required number of $n2$ iterations or terminating conditions where the difference between the optimal Fitness values of last 50 generation is less than the threshold value τ and goto step 10 otherwise. Goto step 3.
10. On terminating condition, the chromosome with the maximum fitness value is selected and optimal ranked list of $m1$ selected clicked URLs is stored in the ordered List $TOR1_j$ associated with the cluster j .

endFor

Phase II: Online Processing

1. User query Q is represented by Fuzzy set $A1$ on $T1$ using $tf.idf$.
2. Fuzzy thesaurus FR restricted to the support of $A1$ and non zero columns is used for the expansion of set $A1$. The support of $A1$ is the set of terms belonging to set $A1$ and used in expressing query Q .
3. Compute expanded input query $B1 \leftarrow A1 \circ FR$ using (6)
4. For each cluster i , the similarity set $B1$ with $cluster_mean$ is measured using the formulae
 $MatchScore1_i(B1, cluster_i) = \{ 2(\text{sim}(B1, cluster_i_mean) * Trust1(i)) / Trust1(i) + \text{sim}(cluster_i_mean, B1) \}$ when $Trust1(i) \neq 0$
 $\text{sim}(B1, cluster_i_mean)$ when $Trust1(i) = 0$
5. Identify the most matching cluster j .
6. if $Trust1(j) = 0$ and $MatchScore1_i(\text{input query}, cluster_j) < 0.5$

Msg “No Trusted Recommendations. Terminate the current session, Initiate new search session”.

Goto step 1.

Else if (Trust1(j)≠0 or MatchScore1(input query , clusterj)≥0.5)

The Ordered ranked list of clicked URLs TOR1j is selected.

7. Recommendation of TOR1j to user.
8. For each recommended URL,the Recommendedcount1 is incremented by 1.
9. Capture the user response to the recommendations and generate the user profile.
10. For each clicked recommendations, the ClickedCount1 is incremented by 1.
11. Update the trust value of the recommended URLs d in a selected cluster as given below.
 $Trust1(URL_d) = \{1 - Distrust1(URL_d) \mid Distrust1(URL_d) \neq 0\}$
 $Distrust1(URL_d) = \{ (Recommendedcount1 (URL_d) - ClickedCount1 (URL_d)) / Recommendedcount1 (URL_d) \}$
12. Update the trust value of the selected cluster j as defined below.
 $Trust1(j) = \{ |CorrectSet1(j)| / RecSet1(j) \}$
 $CorrectSet1(j) = \{ |ClickedURL_{jd} \mid Trust1(ClickedURL_{jd}) > \epsilon \}$
where Recommendedcount1(ClickedURL_{jd})≠0
RecSet1(j) is the total number of recommendations made using cluster j
 $RecSet1(j) = \{ |ClickedURL_{jd} \mid Recommendedcount1(ClickedURL_{jd}) \neq 0 \}$
13. if request for the next result page
 - a. Generate the User session profile keyword vector FUP1 using the trust and TF.IDF content of the clicked URLs.
 - b. Compute expanded Fuzzy User Profile FB1 ← FUP1 o FR where the part of fuzzy thesaurus FR, restricted to the support of FUP1, and non zero columns, is used for expansion.
 - c. Compute the Match Score1 similarity measure of user profile with each jth cluster using the formulae
 $MatchScore1i(clusterj_mean, FB1) = 2 * (sim(FB1, clusterj_mean) * Trust1(j) / Trust1(j) + sim(clusterj_mean, FB1))$ when Trust1(j)≠0
 $MatchScore1i(clusterj_mean) = sim(FB1, clusterj_mean)$ when Trust1(j)=0
 - d. Goto step 5.

Else if request of new search session arises
Goto step 1.
else
a) Update the optimal ranking of cluster using subalgorithm **Genetic Algorithm for optimal ranking of clicked URLs.**
b) Goto step 1.

End

5. Results & Discussion

The experiment was executed on web user query sessions. Architecture was developed on i3 processor, Windows 8 with 120 GB RAM and provide the GUI interface for the web search. During data set generation, the search query was used for the retrieval of Google search results marked with check boxes .The user clicks were stored in database as web query sessions.

The tf.idf of web pages was crawled using web sphinx and processed for session keyword vector generation. The clustering and initialization of trust for both clusters as well as clicked URLs was done using agent implemented in JADE. The parameters used for execution of genetic algorithm were MAXGeneration, Length(P) length of Population, crossover rate [0.6-0.8], mutation rate [.1-.3]. GA was executed for 100 generations on clusters using different mutation rate and crossover rate. The graph show the maximum fitness value of population at different mutation rate in Academics, Entertainment and Sports domain since the mutation operator induces exploration of search space. The results in Fig 1 shows that fitness value converges to its maximum value early as mutation rate increases and become stable with no change as the number of generation

increases. The results were optimal at [0.8, 0.25] (crossover rate, mutation rate), threshold value of Information Scent (ρ) at 0.5 and the threshold value of Trust ε was set to 0.5.

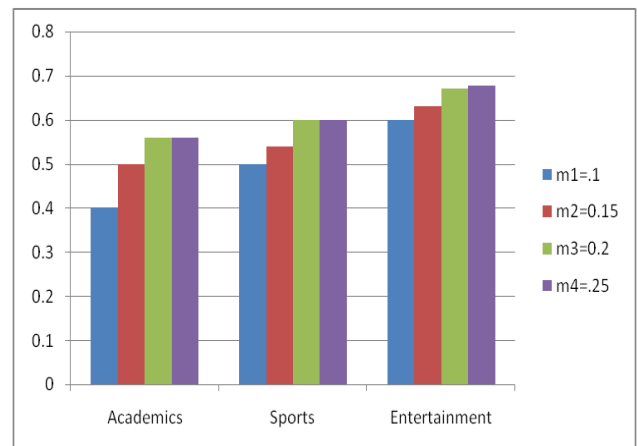


Fig 1. Shows the maximum fitness value at different mutation rate in Academics, Sports and Entertainment domain.

The performance of proposed method was evaluated using average precision of recommended search results. The average precision in a given domain was computed using average of precision of 25 test queries where precision is computed using the fraction of recommended documents that are relevant. The test queries were chosen stochastically for broader coverage of web queries in experimental analysis. The significance of the documents to test queries was resolved by the domain experts (Academics, Entertainment & Sports). The results of the Fuzzy Trust GA IR, FuzzyTrustIR [29], GA IR [21] were compared as shown in Table 1 & Fig 2 below.

Table 1. Shows the average precision of test queries using GA IR, Fuzzy Trust IR, Fuzzy Trust GA IR and percentage improvement over classic IR(Google search engine)

25 test Queries	Average Precision using Information Retrieval Techniques				% improvement over Classic IR		
	Classic IR	GA IR	FuzzyTrust IR	FuzzyTrust GAIR	GA IR	Fuzzy Trust IR	Fuzzy Trust GA IR
Academics	0.51	0.64	0.76	.77	25.49	49.01	50.98
Entertainment	0.5	0.63	0.78	.79	22.44	59.18	61.22
Sports	0.49	0.6	0.78	.79	26	56	58

The results show the increase in average precision using hybrid of Fuzzy, Trust and GA IR for web search personalization. The percentage enhancement of Fuzzy Trust GA IR, GA based IR and Fuzzy Trust based IR over classic IR is shown in Table 1 and Fig 2. The experimental results obtained using Fuzzy Trust GA IR, Fuzzy Trust IR, GA IR were statistically evaluated using paired t-test. The t value for average precision of Fuzzy Trust GAIR versus (FuzzyTrust) for a combined sample(74 degree of freedom) was 16.65, 12.24 in Academics, 8.66 in Entertainment and 8.66 in Sports with 24 degree of freedom in each of domain . The t value for average precision of Fuzzy Trust GA IR versus GA IR for a combined sample was 39.77, 21.56 in Academics, 31.37 in Entertainment and 54.10 in Sports.

Thus the t value for paired difference of average precision was outside the 95% confidence interval and web Information retrieval based on Fuzzy trust GA shows the high t value both over Fuzzy Trust IR and GA IR. Thus Fuzzy Trust GA IR provides effective

method for information retrieval based on web search personalization.

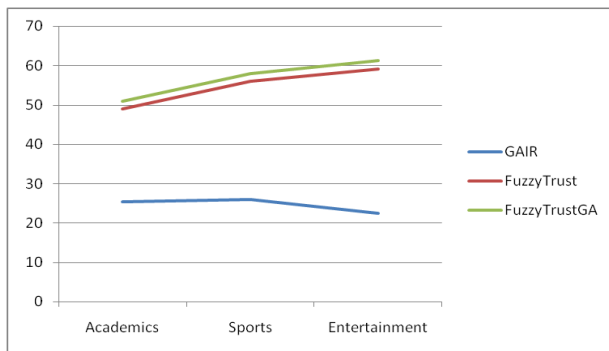


Fig.2. Plots the percentage improvement of hybrid of Fuzzy Trust GA IR/Fuzzy Trust IR/GA based IR over classic IR in Academics, Sports and Entertainment.

6. Conclusion and Future Scope

In this research an algorithm is designed using hybrid of Fuzzy set, Trust and Genetic algorithm for web search personalization. The dimensionality of data set used for experimental evaluation was high and influence the performance of the system. The research work proposed in this paper can be further improved using the feature reduction method.

References

- [1] V. Agarwal, and K. K. Bharadwaj. "Trust-enhanced recommendation of friends in web based social networks using genetic algorithms to learn user preferences". In Trends in Computer Science, Engineering and Information Technology, pp. 476-485. Springer Berlin Heidelberg, 2011.
- [2] K. W. Nafi, A. Hossain, and M. M. Hashem. An advanced certain trust model using fuzzy logic and probabilistic logic theory, 2013. arXiv preprint arXiv:1303.0459.
- [3] R. Hosseini, F. Latifi, and M. Mazinani. "A Fuzzy-GA Approach for Parameter Optimization of A Fuzzy Expert System for Diagnosis of Acute Lymphocytic Leukemia in Children". Journal of Advances in Computer Engineering and Technology, Vol 2, No 2, pp. 33-42, 2016.
- [4] C. Mencar, M. Torsello, D. Dell'Agnello, G. Castellano, and C. Castiello. (2009) Modeling user preferences through adaptive fuzzy profiles. In 9th International Conference on Intelligent Systems Design and Applications, ISDA 2009, pp 1031–1036.
- [5] G. Castellano, D. Dell'Agnello, A. M. Fanelli, C. Mencar, and M. A. Torsello (2010). A competitive learning strategy for adapting fuzzy user profiles. In 10th International Conference on Intelligent Systems Design and Applications, ISDA 2010, pp 959–964.
- [6] F. Kyoormarsi, H. Khosravi, E. Eslami, and M. Davoudi. "Extraction-based text summarization using fuzzy analysis". Iranian Journal of Fuzzy Systems, Vol 7, No 3, pp 15-32, 2010.
- [7] D. Choi. Integration of document index with perception index and its application to fuzzy query on the Internet. In Proceedings of the BISC International. Workshop on Fuzzy Logic and the Internet, pp. 68-72, 2001.
- [8] G. Presser. "Fuzzy personalization". In: M Nikravesh, B Azvine (eds), FLINT 2001, New Directions in Enhancing the Power of the Internet, UC Berkeley Electronics Research Laboratory, Memorandum No. UCB/ERL M01/28, 2001.
- [9] G. Bordogna, G. Pasi (1995). Handling vagueness in information retrieval systems. In: Proceedings of the Second New Zealand International Two-Stream Conference on Artificial Neural Networks and Expert Systems, Nov. 20-23, pp.110-114.
- [10] S. Miyamoto (2012). Fuzzy sets in Information Retrieval and Cluster Analysis. Springer Science & Business Media.
- [11] Y. Ogawa, T. Morita, and K. Kobayashi. A fuzzy document retrieval system using the keyword connection matrix and a learning method. Fuzzy Sets and Systems, 39(2), pp. 163-179, 1991.
- [12] Salha Mohammed Alzahrani and Naomie Salim. (2009). On the Use of Fuzzy Information Retrieval for Gauging similarity of Arabic Documents, In Second International Conference on Applications of Digital Information and Web Technologies (ICADIWT'09), pp. 539-544, IEEE.
- [13] N. O. Rubens. "The application of fuzzy logic to the construction of the ranking function of information retrieval systems.", Computer Modelling and New Technologies, Vol 10 No 1, pp. 20–27, 2006.
- [14] Jianshu Weng, Miao Chunyan and Angela Goh. (2006). Improving Collaborative Filtering with Trust-based Metrics, SAC'06, April, 2327, Dijon, France, ACM 1595931082/06/0004.
- [15] P. Massa, P. Avesani (2007). Trust-aware Recommender Systems. Proceedings of the ACM Conference on Recommender Systems, pp.17-24.
- [16] N. Lathia, S. Hailes, and L. Capra (2008). Trust-based collaborative filtering". Proceedings of the joint iTrust and PST Conference on Privacy, Trust Management and Security. Springer, pp. 119-134.
- [17] C. Hwang, and Y. Chen (2007). Using trust in collaborative filtering recommendation. Lecture Notes in Computer Science, 4570: 1052-1060.
- [18] T. Peng, and T. Seng-cho (2009). iTrustU: A blog recommender system based on multifaceted trust and collaborative filtering. Proceedings of the ACM Symposium on Applied Computing. New York, NY. 1278-1285.
- [19] N. Tyagi, and R. G. Varshney. "A Model To Study Genetic Algorithm For The Flowshop Scheduling Problem". Journal of Information and Operations Management, Vol 3 No 1, pp. 38-42, 2012.
- [20] Amit Verma, and Harpreet Virk Kaur. "A Hybrid Recommender System using Genetic Algorithm and kNN Approach", International Journal of Computer Science And Technology, Vol 6 No 3, pp 131-134, 2015.
- [21] J. Bobadilla, F. Ortega, A. Hernando, J. Alcalá. "Improving collaborative filtering recommender system results and performance using genetic algorithms". Knowledge-based systems, Vol 24 No 8, pp 1310-1316, 2011.
- [22] S. S. Sathya, and P. Simon. "A document retrieval system with combination terms using genetic algorithm". International Journal of Computer and Electrical Engineering, Vol 2, No 1, 1, 2010.
- [23] Nidhi Aley, Shruti Kolte. "Energy Efficient Trust Mechanism using Genetic Algorithm in WSN". International Journal of Computer Science and Mobile Computing, Vol 4, No 6, pp. 146 – 156, 2015.
- [24] N. B. Nimbalkar, S. S. Das, S. J. Wagh. "Trust based Energy Efficient Clustering using Genetic Algorithm in Wireless Sensor Networks" (TEECGA). International Journal of Computer Applications, Vol 112 No 9, pp. 30-33, 2015.
- [25] A. Raha, M. K. Naskar, P. Avishek, A. Chakraborty, and A. Karmakar. "A genetic algorithm inspired load balancing protocol for congestion control in wireless sensor networks using trust based routing framework (GACCTR)". International Journal of Computer Network and Information Security, vol 5, No 9, pp 9-20, 2013.
- [26] V. Agarwal, and K. K. Bharadwaj. "Trust-enhanced recommendation of friends in web based social networks using genetic algorithms to learn user preferences". In Trends in Computer Science, Engineering and Information Technology, pp 476-485, 2011, Springer Berlin Heidelberg.
- [27] C. Selvaraj, and S. Anand. "Peer profile based trust model for P2P

- systems using genetic algorithm". Peer-to-Peer Networking and Applications, vol 5, No 1, pp. 92-103,2012.
- [28] H. Gao, J. Yan, and Yi. Mu. "Trust-oriented QoS aware composite service selection based on genetic algorithms". Concurrency and Computation Practice and Experience, Vol 26, No 2, pp 500-515, 2014.
- [29] S. Chawla. "Intelligent Information Retrieval Using Hybrid of Fuzzy Set and Trust". Oriental Journal of Computer Science. and Technology; Vol 10, No 2, pp 311-325,2017.
- [30] E. H. Chi, P. Pirolli, K. Chen, J. Pitkow. (2001). Using Information Scent to model User Information Needs and Actions on the Web, International Conference on Human Factors in Computing Systems, New York, NY, USA, pp. 490-497.
- [31] J. Heer, and E.H. Chi (2002). Separating the Swarm: Categorization method for user sessions on the web, International Conference on Human Factor in Computing Systems, pp. 243-250.
- [32] P. Pirolli (1997). Computational models of information scent-following in a very large browsable text collection, Conference on Human Factors in Computing Systems, pp. 3-10.
- [33] P. Pirolli (2004). The use of proximal information scent to forage for distal content on the world wide web, Working with Technology, Mind: Brunswikian. Resources for Cognitive Science and Engineering, Oxford University Press.
- [34] S. Chawla, and P. Bedi (2007). Personalized Web Search using Information Scent, International Joint Conferences on Computer, Information and Systems Sciences, and Engineering, Technically Co-Sponsored by: Institute of Electrical & Electronics Engineers (IEEE), University of Bridgeport, published in LNCS (Springer), pp. 483-488.
- [35] S. Chawla, and P. Bedi. (2008). Improving information retrieval precision by finding related queries with similar information need using information scent. In First International Conference on Emerging Trends in Engineering and Technology, ICETET'08, pp. 486-491, IEEE.
- [36] S. Chawla. "Trust in Personalized Web Search based on Clustered Query Sessions". International Journal of Computer Applications, 1 Vol 59, No 7, pp. 36-44, 2012a.
- [37] S. Chawla. "Semantic Query Expansion using Cluster Based Domain Ontologies". International Journal of Information Retrieval Research (IJIRR), Vol 2, No 2, pp. 13-28, 2012b.
- [38] S. Chawla. "Personalized web search using ACO with information scent". International Journal of Knowledge and Web Intelligence, Vol 4, No 2, pp. 238-259, 2013.
- [39] S. Chawla. "Personalised Web Search using Trust based Hubs and Authorities". International Journal of Engineering Research and Applications, Vol 4, No 7, pp. 157-170, 2014a.
- [40] S. Chawla. "Novel Approach to Query Expansion using Genetic Algorithm on Clustered Query Sessions for Effective Personalized Web Search". International Journal of Advanced Research in Computer Science and Software Engineering, Vol 4, No 11, pp. 73-81, 2014b.
- [41] S. Chawla. "Domainwise Web Page Optimization Based On Clustered Query Sessions Using Hybrid Of Trust And ACO For Effective Information Retrieval", International Journal of Scientific and Technology Research, Vol 4, No 11, pp. 196-204, 2015.
- [42] S. Chawla. "A novel approach of cluster based optimal ranking of clicked URLs using genetic algorithm for effective personalized web search". Applied Soft Computing, Vol 46, pp. 90-103, 2016.
- [43] R. J. Wen, Y. J. Nie, and J. H. Zhang. "Query Clustering Using User Logs", Journal ACM Transactions on Information Systems, Vol 20, No 1, pp. 59-81, 2002.
- [44] Y. Zhao, G. Karypis. "Comparison of agglomerative and partitional document clustering algorithms", SIAM Workshop on Clustering High-dimensional Data and its Applications, 2002a.
- [45] Y. Zhao, Y. Karypis. "Criterion functions for document clustering: Experiments and Analysis". Technical report, University of Minnesota, Minneapolis, MN, 2002b.
- [46] G. Klir, and B. Yuan. Fuzzy sets and fuzzy logic, 4, New Jersey: Prentice hall, 1995.
- [47] A. Abdul-Rahman, and S. Hailes. Supporting trust in virtual communities. In Proceedings of the 33rd Annual Hawaii International Conference on System Sciences, (pp. 9-pp). IEEE, 2000.
- [48] D. Harrison McKnight, and Norman L. Chervany. "What trust means in e-commerce customer relationships: an interdisciplinary conceptual typology". International journal of electronic commerce, Vol 6, No 2, pp. 35-59, 2002.
- [49] T. Dimitrakos. "A service-oriented trust management framework". In Workshop on Deception, Fraud and Trust in Agent Societies, pp 53-72, Springer Berlin Heidelberg, 2003.
- [50] J. O'Donovan, and B. Smyth. (2005). Trust in recommender systems. In Proceedings of the 10th international conference on Intelligent user interfaces, pp. 167-174, ACM.
- [51] H.J. Bremermann. 'The evolution of intelligence: The nervous system as a model of its environment'. Technical Report no 1, University of Washington, Department of Mathematics, Seattle, WA, 1958.
- [52] S.K. Pal, V. Talwar, and P. Mitra. 'Web mining in soft computing framework: relevance, state of the art and future directions'. IEEE Transactions on Neural Networks, Vol 13, No 5, pp. 1163-1177, 2002.
- [53] D.E. Goldberg. Genetic algorithms in search, optimization, and machine learning. Addison Wesley Longman Publishing Co., Boston, MA, USA, 1989.