

# Personalized Recommendation Systems (PRES): A Comprehensive Study and Research Issues

**Mr. Raghavendra C K**

Department of Computer Science & Engineering, BNM Institute of Technology, Bengaluru, 560070, Karnataka, India  
Email: raghav.ck.clk@gmail.com

**Dr. Srikantaiah K.C**

Department of Computer Science & Engineering, SJB Institute of Technology, Bengaluru, 560060, Karnataka, India  
Email: srikantaiahkc@yahoo.com

**Venugopal K. R**

Bangalore University, Bengaluru, 560056, Karnataka, India  
Email: venugopalkr@gmail.com

Received: 26 July 2018; Accepted: 05 September 2018; Published: 08 October 2018

**Abstract**—The type of information systems used to recommend items to the users are called Recommendation systems. The concept of recommendations was seen among cavemen, ants and other creatures too. Users often rely on opinion of their peers when looking for selecting something, this usual behavior of the humans, led to the development of recommendation systems. There exist various recommender systems for various areas. The existing recommendation systems use different approaches. The applications of recommendation systems are increasing with increased use of web based search for users' specific requirements. Recommendation techniques are employed by general purpose websites such as google and yahoo based on browsing history and other information like user's geographical locations, interests, behavior in the web, history of purchase and the way they entered the website.

Document recommendation systems recommend documents depending on the similar search done previously by other users. Clickstream data which provides information like user behavior and the path the users take are captured and given as input to document recommendation system. Movie recommendation systems and music recommendation systems are other areas in use and being researched to improve. Social recommendation is gaining the momentum because of huge volume of data generated and diverse requirements of the users. Current web usage trends are forcing companies to continuously research for best ways to provide the users with the suitable information as per the need depending on the search and preferences.

This paper throws light on common strategies being followed for building recommendation systems. The study compares existing techniques and highlights the opportunities available for research in this area.

**Index Terms**—Collaborative based filtering, content based filtering, hybrid filtering, personalized recommendation technique, recommendation system.

## I. INTRODUCTION

It is the need of the hour for most of the companies to initially search, map and further provide the users with the appropriate portion of information as per their requirement and tastes on World Wide Web. By studying the past behavior of the users, intelligent and smart recommendation systems are being built by Companies. The built-in algorithms of Recommendation Systems (RS) aim to deliver accurate and most relevant content to the user. This is achieved by applying filter to the huge pool of information base to get useful stuff. The engines of recommendation systems discover data patterns in the data sets. This is done by learning the choices of consumers and producing the outcomes that co-relate to their interests and needs.

Two common strategies being followed while building decision supportive system named recommendation system are content-based filtering and collaborative filtering techniques. Information like keywords, answers to queries, demographic etc., are gathered by Content based filtering to generate a profile of the user. Based on their profiles, users are matched to items. Collaborative filtering is built on the history of users' behavior. Users rating, purchasing or viewing history is used to establish association between users with similar behaviors and between items of interest to the same users. Fig. 1 and Fig. 2 depict the flow of actions that are executed in content-based and collaborative based RS.

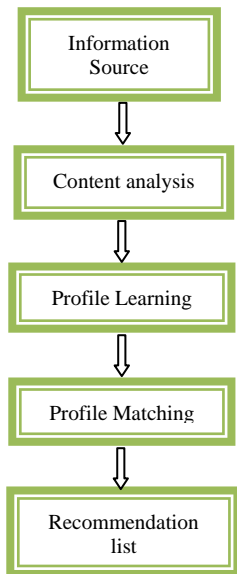


Fig.1. Content based recommendation system

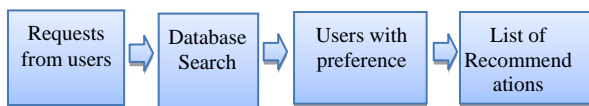


Fig.2. Collaborative based recommendation system

Collaborative recommenders have advantages over content-based recommenders considering few parameters as mention in Table I.

Table 1. Comparison of recommendation systems

Content based recommendation system	Collaborative based recommendation system
More parameters lead to increased complexity	Easy to implement
Prediction accuracy depends on number of attributes	More accuracy in predictions
Personalized	Non-personalized and personalized

The important applications of recommendation systems include product recommendations, Movies recommendations and News articles delivery to the users. Fig. 3 indicates the different recommendation systems in use.

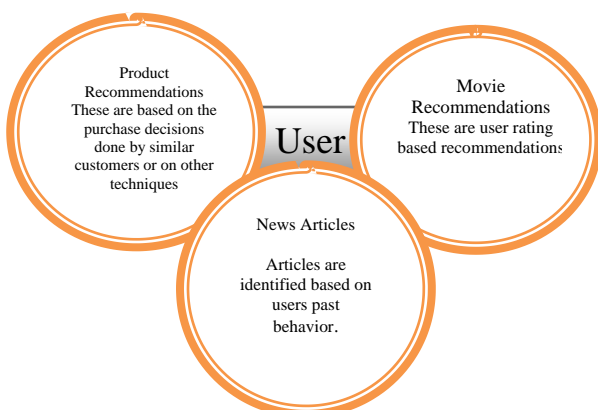


Fig.3. Recommendation systems in use.

## II. FILTERING TECHNIQUES APPLIED IN RECOMMENDATION SYSTEMS

The biggest problem encountered by the web users is the amount of information available in web for the item of interest which invariable fails to offer the required, specific and supportive facts for the user to bank upon and further to take conclusive decisions. The recommendations systems developed can able to provide solution to this problem by applying various filtering techniques such as content based and collaborative based filtering approaches to assess the available content initially and identify the right, elite and customized information further according to the needs of the of the user based on their questioned item.

### A. Content-based filtering (CBF):

CBF works principally based on cognitive type filtering approach, which supports in providing information or recommends the items of interest, depends solely on a comparison note developed between the items related to content and a profile of user. The filtering system works by considering the data of each questioned item in the form of a set of descriptors or terms, which means typically the information provided in the document and the user profile, also denoted with similar set of descriptors and built-up by assessing the data of objects seen by the user. CBF works very effective in particularly situations such as extracting the text documents which are represented by a group of keywords. CBF system typically uses text documents as principle information source of the content and searching by considering the terms and the keywords of the document (selects single words from documents). Further these terms were used in model approaches i.e. vector space model design and latent semantic indexing approach to represent the content documents as a single vector in a multi-dimensional space.

PRES (Personalized Recommender System) is one of the common CBF based approach applied in the recommendation systems. PRES usually work by comparing user profile with the data of all the text based documents identified. Initially data in the form of text is extracted from the pooled documents by running through a number of continuous steps based on the set of terms, further all HTML tags and stop words are identified and removed to provide customized information [1,2].

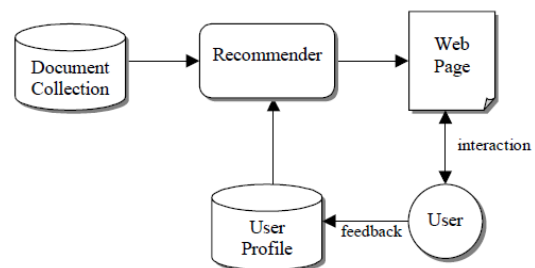


Fig.4. Flow chart depicting the workflow pattern of PRES filtering technique. Source [1]

One of the personal news systems known as News Dude provides a news stories to the users utilizing the speech content developed through TF-IDF model. In this model, short-term recommendations are determined based on the synthesized information and further these recommendations were compared with Cosine Similarity Measure and learning algorithm to provide users customized news stories [3]. One of the highly utilized and rated research paper repositories on the web IBRA today by the user populations known as CiteSeer is an involuntary automated citation indexing cite, that utilizes variety of heuristics and machine learning algorithms of CBF based recommendation system implemented through Naïve Bayes classifier to assess the web based documents and provides the recommendations depends on the needs of user profile and also develops a list of titles based on information provided by an individual user at the time of query. The developed system is competent enough to supply elucidation on any recommendations suggested for the users by cataloging the attributes that support the premier ratings and therefore, motivating the users to have self-reliance on recommendations suggested by the model [4].

CBF system faces limitation in extracting the customized information depends on the requirements of the user profile from multimedia information, because the system observation of the item content such as colors, textures, etc... differs significantly with that of the users. The donated terms for the items need to annotate based on universally accepted annotation system to avoid the miscommunication between the user profile and the system read. An automated method or an algorithm need to be designed and established to correlate the user perception with the machine language while retrieving the recommendation. Another most important setback of the CBF technique is a failure to appraise the value of an item as it cannot differentiate quality documents from the poor when both denoted with the terms. CBF do not have a way of finding opportune items that are interesting for the user but not listed in the user profile [5].

#### B. Collaborative filtering system (CF):

CF system is the most basic filtering approach of the recommendations systems, works based on the withdrawn hypothesis illustrating the psychology of the users; indicating that the people may prefer their stuff based on similarity percentage between items they basically like with that of the items available, and also based on the preferences or choices of the items made by the other who has the similar kind of taste or requirement. Significant research was carried on the applications of the CF technique in the recommendation systems and among the reported approaches the most popular and widely applied technique is model based matrix factorization which works based on low-dimensional factor models. CF techniques are broadly categorized into two different types' i.e., 1. Memory based approach and 2. Model based approach.

a) *Memory based approach:* Memory dependent CF approaches further categorized into two main classes

such as User-Item approach, which filters the content relies on the similarity ratings of the new users given on the similar items of interest with respect to the selected user profile whereas an Item-item approach selects an item first and finds the users profile liked the same item, from that identify the other items of the same users or similar user and suggest them with a recommendation. The memory based approaches works chiefly relay on the similarity between the user and item choices calculated applying by only arithmetic operations such as Cosine similarity or Pearson correlation coefficients without depending on the gradient descent and also lack of application of other optimization algorithms to optimize both User-Item and Item-Item based similarity information to withdrawn the final recommendations. Overall methodology of CF system applied through either User-Item or Item-Item based model is an easy to use approach, but the performance of the model will decrease when less data is available, in turn reduces the efficacy of the approach with respect to the most of the real world problems faced by the users.

b) *Model based approach:* Model based approaches of the CF system works based on machine learning algorithms equipped with an ability to assess the user's rating for unrated or less rated items given by the similar user profiles. The machine learning algorithms applied in the CF were classified into three sub-types based on the working principle.

i) *Matrix Factorization (MF):* Model works on the identified Embeddings known as small number of the hidden information helps in determining the preferences of the user.

ii) *Clustering based algorithm (KNN):* Similar to memory based filtering systems but, in KNN system user and item similarities considered as weights were measured based on an unverified learning model, rather than Pearson correlation or cosine similarity used in the memory based models. Approach is scalable as the algorithm is equipped with the option of limiting the number of similar users selected presented as k.

iii) *Neural Nets/ Deep Learning:* It possess further addition to matrix factorization method where learn the values by splitting the original sparse matrix information into product of low rank orthogonal matrices and assess them in the embedded matrix itself to withdraw the final recommendations [2, 6].

#### C. Hybrid filtering Techniques:

Different recommendation system discussed were individually have their own limitations, with the aim of achieving much improved recommendation systems solutions, a combination of different recommendation system (Hybrid filtering technique) were tested and reported with the higher efficiencies which in turn denotes the significance of Hybrid filtering technique for the much improved, informative and customized recommendation systems for the modern day users. Application of hybrid filtering system certainly improves the accuracies and the prediction abilities of the recommendation systems, as the limitation of the

algorithm can be overcome with the help of other, while working in combination. So hybrid system has clear advantages compared to the individual systems. Hybrid filtering systems can be established following any of the combination i.e., Algorithms can be applied separately and the results can be combined to make the final recommendations, incorporating content-based filtering approaches with that of collaborative approach, applying few CF techniques in CBF approach and establishing a unified combinatorial recommendation system that brings together both approaches [7].

a) *Weighted hybridization (WH)*: WH add the outcome of individual models present as a combination in hybrid approach and withdraw the recommendation by summing the final scores retrieved from all the techniques taken in hybrid system applying a linear formula. P-tango hybrid system is one of the commonly employed weighted hybridization system [2, 7, 8].

b) *Switching hybridization*: Works by swapping any of the recommendation techniques of the hybrid design, applied involving trial and error kind of recommender ability and generates a high-quality rating. This approaches overcome the algorithm limitation by switching the recommendation system present in the hybrid. This approach was found to be very sensitive with regards to the strengths and weaknesses of its component recommenders. DailyLearner system is an exemplar of a switching based hybrid recommender system [2, 7, 9].

c) *Cascade hybridization*: This technique utilizes mathematic logic approach termed iterative based enhancement procedure while developing sequence of predilection amongst dissimilar items and refines the recommendation of the one technique with the other technique in hybrid combination. Techniques are very efficient to the background noise of the algorithm and EntreeC is an example of cascade based hybridization method, developed with the combination of cascade knowledge based and CF systems [2,7,10].

d) *Mixed hybridization*: Unite the recommendation results of all the items collected from the different recommender components of the hybrid system, together at the same time. In this mixed kind of hybridization technique, the individual technique activity or efficiencies do not always influence the universal performance of a local region. PTV system is one of the suggestive examples of mixed hybridization recommender system [7, 11].

e) *Feature-combination*: In this features of one recommendation system is utilized in the other. For example, an attribute of CF i.e. ranking information of related users is applied in a case dependent reasoning approach as key feature to resolve the similarity index between the various substances. Pipper is a renowned example of this hybrid type, which applies the collaborative filter's ratings in a content-based system as a feature for recommending movies [7, 12].

f) *Feature-augmentation*: Utilizes the scoring and preference information retrieved from all the earlier recommenders. Libra system is an example of this kind hybrid approach, which provides CBF based

recommendation of different volumes of books on the information found in website Amazon.com through applying a naive bayes text classifier model. Hybrid approaches related to the Feature augmentation models are more efficient compared to feature-combination methods in that the former will add a small number of features to the primary recommender [5, 7].

g) *Meta-level*: The inner representation given by one recommendation model is considered as an input data for the next recommender system of the hybrid. Meta-level is one of the hybrid systems capable to provide solution to most common sparsity problem of CF techniques. LaboUr [13] an example of this technique, which uses instant based learning to produce content-based user profiles initially and further compared in a collaborative manner to make the final recommendations.

### III. STUDIES CONDUCTED ON VARIOUS RECOMMENDATION TECHNIQUES

Several researchers have attempted to improve the existing document recommendation systems. This section navigates through few such studies and highlight the advantages and drawback of the techniques used.

Yu Liu, Shuai Wang, M. Shahrukh Khan, and Jieyu [14] opine that the research that combines the collaborative filtering and content based recommendations with deep learning is not yet done. The deep hybrid recommender framework that is based on auto-encoders (DHA-RS) that integrates user and item side information has been proposed in this work. DHA-RS combines stacked de noising auto-encoders with neural collaborative filtering, which corresponds to the process of learning user and item features from auxiliary information to predict user preferences. From the experiments done on real-world data set, it is proved that the performance of DHA-RS is superior than the existing computational methods.

F. S. Gohari, F. S. Aliee and H. Haghghi [15] proposed Trust-aware Group Recommendation (TGR) method for improvement of performance in group recommendations. Trust based metric of a novel group optimized involving a method known as Particle Swarm Optimization (PSO) has been used in the proposed scheme. For a user group, this provides a set of neighbors. The proposed approach considers group as a whole and utilizes a metric called GTM (Group Trust Metric) for identification of most truthful neighbors in the group. The process of aggregation is done during trust computation step. The method is designed based on the usual practice that the degree of trust is computed from a group to each of the user. Using this methodology, the TGR metrics can attain a set of neighbors with respect to the user's group thus leading to improved run time performance. The results obtained by experimentation reveal that TGR improves the overall period (run time) performance and accuracy of group based RS designs.

B. K. Sunny, P. S. Janardhanan, A. B. Francis and R. Murali [16] state that the traditional RS designs cannot satisfy the needs of the user because of the limited capability of the models and analysis. Because of increased availability



of technologies for processing of big data, the implement of real time suggestions or recommendation has become very easy and effective. The velocity attribute of big data is being handled by new paradigm called stream-computing. This technology enables to develop real time big data applications. Real time RS is implemented in this work. The implementation has been done using Apache Spark which is a platform for stream computing. The system recommends TV channels to users as an existent real instance. This is a huge challenge as viewers preferences are context dependent and the set of channels are dynamically changing. The traditional approaches are not capable of handling recommendation scenarios characterized by dynamic nature, time constraints, data volume etc. The researchers implemented a TV channel RS that is optimized for data streams based on the high throughput manner (Real time) that originates from set-top boxes. The self-adaptive method for building recommendation model is implemented. Apache Spark's distributed processing capability has been effectively utilized for real time recommendation. The machine learning libraries provides several algorithms for developing recommendation system. Lambda Architecture has been used for processing of huge volume of data in the system efficiently.

Manasa N, Bavya S, Kavitha G [17] utilized the perception of just-in-time retrieval, which helps in instinctively recommending documents that are associated to users' present activities. The proposed system provides well-organized way for document recommendation system for users using the conversational data. Text file of informal data is given as input. These informal data are partitioned into  $m$  clusters. Clusters contain numerous numbers of keywords including surplus words. Important and useful topic related keywords are extracted using Word dictionary. Keywords are graded based on their number of occurrences or weights. By selecting maximum ranked keyword document recommendation method will be achieved.

Michael J. Pazzani and Daniel Billsus [18] proposed a system that suggests a product or information to a user depends on explanation of the article and user's interests. These type of Content-based recommendation systems helps in recommending web pages, hotels & restaurants, places, institutes, news bulletin articles, TV programs list, and on various online shopping websites. Content based RS have universal means of description of items, a method for creating users profile that indicate the items liked by the user and method of comparing profile of users to detect what is recommended. The user profile is created initially and further updating is done in reply to the feedback on the interest the users shows on presented items.

RanaChamsi Abu Quba [19] investigated specific aspects of the recommendation systems like General Purpose Social or Public Networks to verify if predictions can be made depends on natural activities on social or public networks. The researchers have concerns like whether the implicit data available in GPSNs can be

transformed into recommendation engine or recommendation system engine evaluation can be done apart from traditional accuracy metrics like MAE. The researchers developed a hybrid social recommender called intelligent Social Network Transformer into Recommendation Engine (iSoNTRE). The engine transforms the richness in GPSNs into recommendation system that is robust. The system can be applied across various domains by converting user's information into recommendation based information. In this way, iSoNTRE saves users efforts as they need not enter what they like.

As pointed by S. Meng, W. Dou, X. Zhang and J. Chen [20] there is an enormous increase in the amount of customers, information and service. Because of this, problem related to the BIG data analytics has been raised for service based recommender systems. Scalability and inefficiency are the existing issues for processing and analysis of huge data in traditional recommendation systems. Currently accessible recommendation systems provide same rankings and ratings of services to users without taking into considerations of diverse user preference. Because of this, these systems fail to meet up the tailored requirements of users. The researcher group of the study proposed a KASR (Keyword Aware Service Recommendation method) to deal with these issues. The proposed method presents a service recommendation that is personalized and appropriate to the users. The preferences of the users are indicated by keywords. For generating suitable recommendations, a user-based CF algorithm is implemented. In order to adapt to the present scenario of big data analytics, KASR is executed along with Hadoop which is a widely used distributed computing framework. The experiments are done using real world data. The accuracy and scalability of the service recommenders can be enhanced by using the proposed method.

Due to large number of options available, choosing the right service is not easy. Back Sun Sim, Heeseong Kim, KwangMyung Kim and H. Y. Youn [21] opine that the recommendation system currently in use doesn't properly judge the public relation of the entities in giving the rating for the choices. In this work, a novel approach is proposed that is context-aware that explains the user type for estimation of the proximity between the social network users, collectively with the cosine based similarity measure. The computational model performed for this projected work shows that there is a significant improvement in accuracy of ranking compared to current techniques. The amount of various items suggested or recommended by this new method is comparatively much larger quantity.

V. B. Savadekar and M. E. Patil [22] opine that traditional recommendation systems undergo incompetence problems in the assessment of big data generated from large amount of information available online. Many of the existing methods provide fixed ratings and items scorings to various users devoid of considering the needs of the users resulting in failure to deliver personalized requirements. In this work, a

Table 2. Strengths and drawbacks of various RS

Authors/ Year of publication	Domain tested	Technique/ Algorithm employed	Advantages	Disadvantages/ Remarks
Yu Liu, Shuai Wang, M. Shahrukh Khan and Jieyu He [14]	Experimented using movielens dataset	Hybrid recommender system based on auto-encoders (DHA-RS) integrating user and item side information to enhance performance	Helped to conquer the setbacks such as limited learning capacities, data sparsity and the cold-start.	More processing time is required
F. S. Gohari, F. S. Aliee and H. Haghighi[15]	Experimented using movielens dataset	Proposed TGR approach for enhancing the performance of group recommendations by applying Particle Swarm Optimization (PSO) method	Improves the overall run time and accuracy of group based RS designs	Measuring diversity among group members with respect to the different data dimensions is to be addressed as different groups may have different dynamics.
B. K. Sunny, P. S.Janardhanan, A. B. Francis and R. Murali[16]	TV Program recommendation	Apache Spark which is a platform for stream computing.	This research work explains the TV channel recommendation system that is developed using the real-time Click-stream data to a particular viewer in real time.	The implementation methodology used here can be applied for various areas that need real-time demand on daily basis.
Manasa.N, Bavya S, Kavitha G[17]	Recommendation for small business meetings	Clustering method, works based on the Key words	Researchers utilized the perception of just-in-time retrieval, which helps in instinctively recommending documents that are associated to users' present activities	The researchers have achieved considerable improvement. But there is scope for improvement.
Michael J. Pazzani and Daniel Billsus[18]	Discussion on multiple domains like restaurants and news articles	Vector space Model	Proposed a system that suggests a product or information to a user depends upon explanation of the thing and user's interests.	Algorithms are only a part that provides recommendations to users. There exists a scope for research concerning user knowledge in the different possible ways of retrieving the data and the associated problems.
RanaChamsi Abu Quba[19]	Social Networks	Intelligent Social Network Transformer into Recommendation Engine (iSoNTRE).	Investigated specific aspects of the recommendation systems like General Purpose Social Networks to decide if predictions can be made based on natural activities on social networks.	The recommendations done are based only on their properties or content. This is not based on any advertising mechanism involving paid inclusion in a search result or product promotions.
S. Meng, W. Dou, X. Zhang and J. Chen[20]	Personalized service	User-based Collaborative filtering algorithm	KASR addresses the issue of considering the diverse users' preferences. Highly scalable.	Works only with frequency of keywords. The Performance measurement is done using Keyword extraction.
Back Sun Sim, Heeseong Kim, KwangMyung Kim and H. Y. Youn[21]	Social Network	Content based recommendation system	Proposed a new context-aware system for providing services to the users by consideration of social position of user and the context.	Improve the accuracy of rating by using context information.
Xiaoming Liu, Chao Shen[26]	Social Networks	Probabilistic graph model	Proposed a model for discovering groups which are similar across multiple social networks	Need to explore more advanced models by expanding feature sources for group detection which are similar.
Surong Yan, Kwei-Jay Lin, Xiaolin Zheng and Xiaoqing Feng[27]	Social Networks	Social network with matrix factorization (SNMF)	Proposed a model which utilizes user's social networks relations to generate recommendations.	Researchers attempted to ease inherent social recommendation problems by matching needs of accuracy and scalability but performance can still be improved.
Xiwang Yang et al..[28]	Online social voting in Social Networks	Collaborative filtering	Proposed a model for online social voting based on matrix factorization and nearest neighbor.	Researchers need to study content of voting information to tackle cold votings.
Ningning Yi, Chunfang Li, Xin Feng, Minyong Shi[30]	Movie Recommendation System	Collaborative filtering.	Implemented a movie RS based on user based CF using graph databases.	The scoring matrix will be pre filled to tackle sparsity problem

personalized recommendation system based on English verbal communication (language) Keyword list and domain vocabulary is proposed. The suggestive (or) recommendation list presented is customized and most appropriate items are provided effectively to the users as per user's needs. Hadoop has been used for implementation of the work which supports distributed computing for parallel processing. It is demonstrated that the system enhances the accuracy and measurability of the recommendation systems.

H. L. Nguyen and J. J. Jung [23] states that the users trust keep changing because of changes due to various conditions. It is the need of the hour to research about change in trust as this is applicable to trust based recommendation systems which are diverse. The authors in this work proposed a general definition of trust, defined spatial and temporal patterns to explain how the trust changes which can be the output of trust-based recommendation systems. The researchers have proved the effectiveness of the method using case study between a restaurant and user of the Twitter. The researchers have opined that the research can be applied to various fields like e-learning systems.

Zhang and Xiaoying [24] designed model for real rating system for capturing and minimizing past distortions in each of the rating. The model is called Historical Influence Aware Latent Factor Model (HIALF) that allows examining the historical ratings influence patterns from the perspective of a modeling. The model perfectly matches the contrast effects and assimilation observed previously. In predicting subsequent ratings, HIALF accomplishes considerable improvements. The prediction of relationships exposed in empirical measurements of the past on real ratings. It is concluded that the HIALF provides improved recommendations. This is achieved by disintegrating user real time preferences that of ratings which are distorted. This enables the wiser consumer obtain decisions by illuminating the inherent product quality.

EvangeliaChristakopoulou, George Karypis [25] state that Sparse Linear Methods (SLIM) that are item-based approaches have proved to be good for top-N recommendation. The drawback of these approaches is that, they approximate only a particular model for all the users. In order to overcome this, the researchers worked on the model depend on the perception that the entire users does not act in a similar way - instead subsets of similar minded users exists. The variations in the preferences of the users can be captured by utilizing different item-item models for the subsets of users. This results in better performance for suggesting top-N recommendations. In the proposed work, an extended SLIM is presented resulted by combining local and global SLIM models. The researchers have presented a technique in which prediction scores are computed as a user-specific combination of the assumptions. These predictions are derived by global and local representing model based on the item to item basis. In this approach, to better the performance of finding top-N recommendations, the local model, the global model,

user-specific combination, user's assignments to the local models are optimized jointly. Experimentation revealed that the method performs better than the typical SLIM model and other existing top-N recommendation methods.

Xiaoming Liu, Chao shen [26] proposed a method consisting of three steps, first finding group structure using random walks, second similar features extraction and discovering group similarity by applying graph probabilistic model. The model is evaluated on online social networks and results shows significance improvements over existing techniques available. The summary of the study has been tabulated in the Table II.

#### IV. EVALUATION METRICS OF RECOMMENDATION SYSTEMS

Put into practice, there are some wide spread metrics available for assessing the performance of suggested recommendation systems.

*User rating:* Recommendation systems that work by calculating user's rating for new-fangled items, it's feasible to compare the predictable rating to the original rating in the assessed test data. Further the applied RS judges the possible recommendations depends on the value obtained through the standard difference from the right value.

*Precision:* It is a calculation of exactness; it measures the portion of applicable items data retrieved out of the information specifying all the items. For example, it is the percentage of recommended cinemas that are truly good. Precision tells how many of the selected objects were correct.

$$precision = \frac{tp}{tp + fp} \quad (1)$$

*Recall:* It is a measure of totality and identifies the portion of applicable items retrieved out of all significant items. It is the percentage of all good quality movies recommended. Recall tells how many of the objects that should have been selected were actually selected.

$$recall = \frac{tp}{tp + fn} \quad (2)$$

*Mean Absolute Error:* A commonly used analytical accuracy metric is referred to as a mean absolute error, or MAE. MAE considers the sum of the divergence between the user's rating and the predicted rating and categorizes with that of the number of items considered.

$$MAE = \frac{1}{|B_i|} \sum_{b_k \in B_i} |r_i(b_k) - p_i(b_k)| \quad (3)$$

Where,  $|B_i|$  represents the items ranked by user  $a_i$ .

*Mean Squared Error (MSE):* It is a measure that is depends on the (squared) variation between predicted values with that of actual value. In recommenders based

on explicit (or rating) data, you use the rating data to generate a deposit of predictive rankings. You can then use individual's predictions to rank items. The error you measure with MSE is based on the difference between the rating that you predicted for an item and the rating that a user gave that item. MSE does not work for recommenders based on implicit feedback.

$$MSE(t) = \frac{1}{n} \sum_{i=1}^k f_i(x_i - t)^2 = \sum_{i=1}^k p_i(x_i - t)^2 \quad (4)$$

**Root Mean Squared Error (RMSE):** It is an evaluation method that works based on the (squared) difference between a prediction and an actual value.

It usually gives how distant the model is from giving the right answer.

$$RMSE = \sqrt{\frac{1}{n} \sum_{j=1}^n (y_i - y'_j)^2} \quad (5)$$

**F1-Measure:** It finds the weighted average of both precision and recall.

$$f = 2 \cdot \frac{\text{precision} \cdot \text{recall}}{\text{precision} + \text{recall}} \quad (6)$$

## V. KEY FACTORS TO BE CONSIDERED FOR RECOMMENDATION SYSTEMS

1) **Data Sources and methods:** The range of data extracted includes user's current behavior, demographic data, stated preferences, usage history. The recommendation can be tuned for best performance for

range of methods and variety of data. The maximum control on recommenders is possible through this. It is important to know that whether the solution supports all these methods and data sources.

The research in recommendation system is gaining momentum. The research community is expanding and hence, there are innumerable resources available online for learning and building recommendation systems.

2) **Interface for managing recommendations:** It is necessary to consider whether the interface is easy to see where the attention is needed and changes can be easily done, and see results. The solution must provide a single interface for managing recommendations across touch points, marketing programs and segments. Framework for designing these user interfaces include the features to collect the metric that has following characteristics.

a) **Should be Predictive:** Predictive features will map to interfaces that clearly capture a user's affinity for an item.

b) **Should not be Unambiguous:** Content actions are just confusingly named so different people interpret it differently and thus use it differently.

c) **Dense:** User-item interactions should be lightweight as possible. Otherwise dense data leads to errors in model building.

Usually, best metrics are implicit that are not entered by user. For example, time spent on watching particular page/video.

Recommender systems management interface can be developed using any programming language. Following are some of the tools that are available to manage the recommenders.

Table 3. Typical datasets in recommendation system.

SL	Datasets	Description
1	MovieLens dataset	MovieLens is a collection of movie ratings and comes in various sizes. Datasets are named as 1M, 10M, and 20M because they contain 1, 10, and 20 million ratings. The largest set uses data from about 140,000 users and covers 27,000 movies. MovieLens dataset is featured in Kaggle competitions. This dataset is used to build and benchmark recommenders and is publicly available. The dataset is unique because it is organized in easily understandable way and provides right volume. ( <a href="https://grouplens.org/datasets/movielens/">https://grouplens.org/datasets/movielens/</a> )
2	BookCrossing	This dataset contains 278,858 users that provide 1,149,780 ratings (explicit / implicit) about 271,379 books. ( <a href="http://www.informatik.uni-freiburg.de/~cziegler/BX/">http://www.informatik.uni-freiburg.de/~cziegler/BX/</a> )
3	Jester	This data is generated from jester. It is Online Joke RS ( <a href="http://eigentaste.berkeley.edu/dataset/">http://eigentaste.berkeley.edu/dataset/</a> )
4	Libimseti	This dataset is generated from the network of ratings given by users to others. In this directed network edges represent ratings given on a scale of 1-10. ( <a href="http://konect.uni-koblenz.de/networks/libimseti">http://konect.uni-koblenz.de/networks/libimseti</a> )
5	Last.fm	This datasets contain music recommendations. For each user in the dataset it contains a list of their top most listened to artists including the number of times those artists were played. ( <a href="http://www.dtic.upf.edu/~ocelma/MusicRecommendationDataset/index.html">http://www.dtic.upf.edu/~ocelma/MusicRecommendationDataset/index.html</a> )
6	OpenStreetMap	OpenStreetMap's data is provided by their users and a full dump of the entire edit history is available. The dataset include roads, buildings, points-of-interest etc., ( <a href="http://openstreetmapdata.com/">http://openstreetmapdata.com/</a> )
7	Python Git Repositories	This dataset is based on Python code contained in Git repositories. ( <a href="https://github.com/datasets">https://github.com/datasets</a> )
8	Epinions	This dataset contains 664,824 user ratings, 487,183 user relations and 8,196,077 users to item ratings. ( <a href="https://github.com/dokooh/epinions_annotreview_dataset">https://github.com/dokooh/epinions_annotreview_dataset</a> )
9	Flixster	This dataset contains 7,058,819 user relations and 16,830,839 ratings. ( <a href="http://socialcomputing.asu.edu/datasets/Flixster">http://socialcomputing.asu.edu/datasets/Flixster</a> )
10	Douban	This dataset contains 1,692,952 user relations and 293,299 ratings. ( <a href="http://socialcomputing.asu.edu/datasets/Douban">http://socialcomputing.asu.edu/datasets/Douban</a> )
11	Netflix	This dataset contains 10,280,185 user relations and 964,034 item relations. ( <a href="https://www.kaggle.com/netflix-inc/netflix-prize-data">https://www.kaggle.com/netflix-inc/netflix-prize-data</a> )
12	Meetup	This is a social networking website which allows its members to find and join similar groups. As of 2017, there are 32000000 users, 280000 groups available across 182 countries. Data can be collected by using Meetup API. ( <a href="https://www.kaggle.com/sirpunch/meetups-data-from-meetupcom">https://www.kaggle.com/sirpunch/meetups-data-from-meetupcom</a> )



Table 4. Tools for managing recommenders

Tool	Purpose
In-tag	Content discovery that relies on user intent signals to recommend the right content for each visitor.
NewsCred	It helps to manage content recommendations to get more people to stay on particular site longer.
Curata CCS	It is a content-curation tool but can be used as content recommendation.

1) *Automation and self-optimization*: It is necessary to consider what the extent of self-optimization and automation is built into the system. Whether the system is refinable at the individual level or not is another point to be considered. The solution is to be automated to try different algorithms versions for finding optimum results without user intervention. Almost every big company uses recommenders with Automation and self-optimization features especially in the e-commerce area. Following are some of the examples.

Netflix	– Movie recommendation
Amazon	– Product recommendation
Best Buy	– Product recommendation
Alibaba	– Product recommendation
YouTube	– Video recommendation

2) *APIs*: Recommendations personalize experience of the visitors. But, ever changing technology strategies lead to continuous tuning of recommendation systems in the way they use data, the way they interact with other systems like CRM and content management systems. Provision of APIs is a factor to be considered.

Recommender API gives an arrangement of simple to-utilize, completely recorded APIs and computational devices to compose proposal sub-modules. Some of them are listed here.

a) *Recombee*: It is a real-time Recommendation-as-a-Service (SaaS). It provides intuitive graphical user interface (GUI) which allows to monitor KPIs (such as CTR/CR) in near real-time.

b) *Biblio API*: It provides a recommender system designed specifically for nonfiction publishing, subscription media and educational platforms.

c) *ShepHertz*: It provides APIs for recommendation.

d) *Drupal*: It provides APIs for recommendation with modules such as browsing history recommenders, Commerce history recommenders.

e) *Meetup API*: It is used to collect data from meetup website.

3) *Potential Issues*: A comprehensive study [29] on RS which highlights the potential issues are.

a) *Domain Deficiency*: Most of the articles show research is mainly related to entertainment kind of domains like movies, music and so on. Articles related to domains like social reviews are less explored and there is a scope of research on these domains.

b) *Academic Digital Library*: There is deficiency in academic libraries datasets which are significant for future exploration.

c) *Recommendations in E-commerce websites*: Users social media data can be incorporated with existing techniques for better recommendations.

d) *Data Sparsity*: Several methods exist to handle data sparsity problem but still in many domains data sparsity problem persists.

e) *Security*: security is another major concern exist in social media is security due to IoT devices so there is a need of RS which verifies manages the data in IoT devices.

## VI. ANALYSIS OF STUDIES AND CONCLUSION

The existing recommendation systems are designed depends up on CBF and CF techniques. Some of the solutions are hybrid of these two methods. But, considering the information explosion happening because of general purpose social networking sites and domain specific networking sites, the requirement to provide more accurate and improved systems is increasing day by day. As the current systems do not guarantee the user satisfaction, there is a scope for improvement in the recommendation systems to cater the need of present and future.

The recommendation systems suffer from the problems such as changing data, lack of data, unpredictable items and changing user preferences. A good recommender system needs item data for analysis. The data are then fed to algorithms. The accuracy of the recommendations depends on volume of item and user data. One of the parameters used in building recommendation systems is users past behavior. This is not a good parameter because the trends keep changing. Another challenge to recommendation system is changing user preferences. User response to some items is diverse and unpredictable. Hence, it is difficult to make recommendations on such items.

The future of recommender systems lies in integrating self-actualization to do justice. There is a need to build recommendation systems which are not extremely personalized because it may be felt intrusive to the user. Also the recommendation cannot be very generic that it does not consider users' unique requirements. Striking a balance between these two is to be achieved. Today's complex systems have space for each and every technology. Future recommendation systems are to be built by exploiting the power of the tools and ever-growing computing capabilities.

## REFERENCES

- [1] Robin van Meteren and Maarten van Someren, "Using content based filtering for recommendation," *Conference Proceedings, semantics scholar*, 2000.
- [2] J. Bobadilla, F. Ortega, A. Hernando, A. Gutiérrez, "Recommender systems survey," *Knowledge-Based Systems*, Vol. 46, pp.109–132, 2013.

- [3] Billsus D, Pazzani MJ, "User modeling for adaptive news access," *User modeling and User-adapted Interaction*, Vol. 10(2-3), pp.147-180, 2000.
- [4] Raymond J. Mooney and Roy L, "Content-based book recommending using learning for text categorization," *Proceedings of the fifth ACM conference on digital libraries, ACM*, pp. 195-204, 2000.
- [5] Cacheda F, Carneiro V, Fernández D, Formoso V, "Comparison of collaborative filtering algorithms: Limitations of current techniques and proposals for scalable, high-performance recommender systems," *ACM Transactions on the Web (TWEB)*, Article 2, Vol. 5(1), February 2011. "doi: 10.1145/1921591.1921593".
- [6] George A Sielis, Aimilia Tzanavari, George A Papadopoulos, "Recommender Systems Review: Types, Techniques and Applications," *In 3rd edition Encyclopedia of Information Science and Technology*, pp.7260-7270, 2015.
- [7] F.O. Isinkaye, Y.O. Folajimi, B.A. Ojokoh, "Recommendation systems: Principles, methods and Evaluation," *Egyptian Informatics Journal*, Vol. 16(3), pp. 261-273, November 2015,
- [8] Claypool M, Gokhale A, Miranda T, Murnikov P, Netes D, Sartin M, "Combining content- based and collaborative filters in an online newspaper," *In Proceedings of ACM SIGIR workshop on recommender systems: algorithms and evaluation, Berkeley, California*, 1999.
- [9] Billsus D, Pazzani MJ, "A hybrid user model for news story classification," *In Proceedings of the seventh international conference on user modeling, Banff, Canada, Springer-Verlag, New York*, pp. 99-108, 1999.
- [10] Burke R, "Hybrid recommender systems: survey and experiments", *User modeling and User-adapted Interaction*, vol. 12(4), pp.331-370, 2002. "doi:10.1023/A:1021240730564".
- [11] Smyth B, Cotter P, "A personalized TV listings service for the digital TV age," *Knowledge Based Systems*, Vol. 13(2-3), pp.53-59, April 2000. "https://doi.org/10.1016/S0950-7051(00)00046-0"
- [12] Pazzani MJ, "A framework for collaborative, content-based and demographic filtering," *Artificial Intelligence Review*, Vol. 13(5-6), pp. 393-408, December 1999.
- [13] Basu C, Hirsh H, Cohen W, "Recommendation as classification: using social and content-based information in recommendation," *In Proceedings of the 15th national conference on artificial intelligence*, Madison, WI, pp. 714-20, 1998.
- [14] Yu Liu, Shuai Wang, M. Shahrukh Khan, and Jieyu He, "A Novel Deep Hybrid Recommender System Based on Auto-encoder with Neural Collaborative Filtering," *Big Data Mining and Analytics*, Vol. 1(3), pp. 211-221, September 2018."doi: 10.26599/BDMA.2018.9020019"
- [15] F. S. Gohari, F. S. Aliee and H. Haghighi, "A trust-aware group recommender system using particle swarm optimization," *International Symposium on Computer Science and Software Engineering Conference (CSSE)*, Shiraz, pp. 80-85, 2017.
- [16] B. K. Sunny, P. S. Janardhanan, A. B. Francis and R. Murali, "Implementation of a self-adaptive real time recommendation system using spark machine learning libraries," *IEEE International Conference on Signal Processing, Informatics, Communication and Energy Systems (SPICES)*, Kollam, pp. 1-7, 2017.
- [17] Manasa.N, Bavya S, Kavitha G, "To Recommend Documents In Small Business Meetings by Extracting Keywords & Clustering them", *International Journal of Advance Research in Engineering, Science & Technology*, Vol. 3(7), July-2016,
- [18] Michael J. Pazzani and Daniel Billsus, "Content Based Recommendation System", *The Adaptive Web, Springer, LNCS 4321*, pp. 325 - 341, 2007.
- [19] RanaChamsi Abu Quba, "On Enhancing Recommender Systems by Utilizing general Social Networks Combined with User's Goals and Contextual Awareness," *Networking and Internet Architecture, University Claude Bernard - Lyon I*, 2015
- [20] S. Meng, W. Dou, X. Zhang and J. Chen, "KASR: A Keyword-Aware Service Recommendation Method on MapReduce for Big Data Applications," *IEEE Transactions on Parallel and Distributed Systems*, Vol. 25(12), pp. 3221-3231, December 2014.
- [21] Back Sun Sim, Heeseong Kim, KwangMyung Kim and H. Y. Youn, "Type-based context-aware service Recommender System for social network," *International Conference on Computer, Information and Telecommunication Systems (CITS)*, Amman, pp. 1-5, 2012.
- [22] V. B. Savadekar and M. E. Patil, "Improved recommendation system with review analysis," *International Conference on Global Trends in Signal Processing, Information Computing and Communication (ICGTSPICC)*, Jalgaon, pp. 79-82, 2016.
- [23] H. L. Nguyen and J. J. Jung, "Utilizing Dynamics Patterns of Trust for Recommendation System," *International Conference on Intelligent Environments (IE)*, Seoul, pp. 108-113, 2017.
- [24] Zhang and Xiaoying, "Modeling the Assimilation-Contrast Effects in Online Product Rating Systems: Debiasing and Recommendations," *Proceedings of the Eleventh ACM Conference on Recommender Systems*, pp. 98-106, 2017.
- [25] EvangeliaChristakopoulou and George Karypis, "Local Item-Item Models for Top-N Recommendation," *Proceedings of the 10th ACM Conference on Recommender Systems*, pp. 67-74, 2016.
- [26] Xiaoming Liu, Chao Shen, "We know who you are: Discovering similar groups across multiple social networks," accepted for inclusion in *IEEE Transactions on Systems, Man and Cybernetics Systems*, 2018.
- [27] Surong Yan, Kwei-Jay Lin, Xiaolin Zheng and Xiaoqing Feng, "An approach for building efficient and accurate social recommender systems using individual relationship networks", *IEEE Transactions on Knowledge and Data Engineering*, Vol. 29(10), October 2017.
- [28] Xiwang Yang, Chao Liang, Miao Zhao, Hongwei Wang, Hao Ding, Yong Liu, Yang Li and Junlin Zhang, "Collaborative filtering based recommendation of online social voting", *IEEE Transactions on Computational Social Systems*, Vol. 4(1), March 2017.
- [29] Anitha Anandhan, Liyana shuib, Maizatul Akmar Ismail, and Ghulam Mujtaba, "Social Media Recommender Systems: Review and Open Research Issues," *In IEEE Access*, April 2018."doi: 10.1109/ACCESS.2018.2810062"
- [30] Ningning Yi, Chunfang Li, Xin Feng, Minyong Shi, "Design and Implementation of Movie Recommender System Based on Graph Database," *14th Web Information Systems and Applications Conference*, 2017. "DOI 10.1109/WISA.2017.34"

## Authors' Profiles



**Raghavendra C K** is currently working as an Assistant Professor in the Department of Computer Science and Engineering at B N M Institute of Technology, Bangalore and pursuing the Ph.D degree in the Department of Computer Science and Engineering at S J B Institute of Technology, Bangalore under Visvesvaraya Technological University, Belagavi, India. He obtained his B.E degree in 2009 and M.Tech degree in 2011 from Visvesvaraya Technological University, Belagavi, India. His current research interests are in the area of Data Mining, Web Mining, Big Data Analytics, Recommendation systems Social Networks Analysis and Image Processing.



**Dr. Srikantaiah K C** is currently a Professor in the Department of Computer Science and Engineering at S J B Institute of Technology, Bangalore, India. He obtained his B.E from Bangalore Institute of Technology, M.E from University Visvesvaraya College of Engineering, Bangalore in 2002 and Ph. D degree in Computer Science and Engineering from Bangalore University, Bangalore. He has authored a book on Web Mining Algorithms. His research interest is in the area of Data Mining, Web Mining, Big Data Analytics, Cloud Analytics and Semantic Web.



**Venugopal K R** is currently working as Vice-chancellor for Bangalore University. He pursued Bachelor of Engineering from University Visvesvaraya College of Engineering, Bangalore University, India and Master of Engineering from Indian Institute of Science, Bangalore. He pursued PhD in Computer Science from Indian Institute of Technology, Madras and PhD in Economics from Bangalore University, Bangalore. Formerly he was working as Principal, University Visvesvaraya College of Engineering, Bangalore.

He is a member of IEEE, Member of AICTE, Bangalore. He has published more than 600 research papers in reputed international journals including Thomson Reuters (SCI & Web of Science) and conferences including IEEE and it's also available online. He has filed 101 patents and authored and edited 62 books. He has received IEEE Fellow for contributions in Computer Science and Electrical Engineering Education and ACM Distinguished Educator Award for his contributions in education. His main research work focuses on Data Mining, Optical Networks, Ad Hoc Networks, Sensor Networks, Digital Signal Processing, and IoT. He has 36 years of teaching experience and Research Experience.

**How to cite this paper:** Raghavendra C K, Srikantaiah K.C, Venugopal K. R, " Personalized Recommendation Systems (PRES): A Comprehensive Study and Research Issues", International Journal of Modern Education and Computer Science(IJMECS), Vol.10, No.10, pp. 11-21, 2018.DOI: 10.5815/ijmeecs.2018.10.02