IJCSE International Journal of Computer Sciences and Engineering Open Access

Research Paper Vol.-6, Issue-5, May 2018 E-ISSN: 2347-2693

Hybrid Recommender Systems: Process, Challenges, Approaches and Metrics

K.Reka[1*], T.N.Ravi[2]

[1*]Department of Computer Science, Cauvery College For Women, Trichy, TamilNadu, India
[2] Department of Computer Science, Periyar E.V.R College, Trichy, TamilNadu, India

**Corresponding Author: rekaccw@gmail.com*

Available online at: www.ijcseonline.org

Accepted: 24/May/2018, Published: 31/May/2018

Abstract: Recommender systems plays a significant role, by providing personalized information to users over the internet. With the evolution of the internet, the recommender systems too have evolved from being based on simple demographics, user and item information, into complex hybrid models capable of providing an effective real-time recommendation on a per-user basis. This works provides an overview of traditional recommender system approaches, their taxonomy and discusses the various hybridization techniques used for creating complex models that provide hyper-personalized recommendations. A detailed discussion of the research challenges and how they impact the performance of the various recommender models have been presented as a solution to the existing issues in recommendation systems. Metrics for evaluation and the need for diversity and novelty in recommender systems have also been discussed. Future research directions concerning mobile and IoT based, context-aware recommender systems and the effectiveness of Deep Learning models and how Transfer Learning could address the major drawbacks of recommender systems have also been discussed.

Keywords: Recommender Systems; Hybrid Filtering; Collaborative Filtering; Content-Based Recommendation; Context-Aware Recommender; Demographic Filtering

I. INTRODUCTION

Recommender Systems (RSs) provide or predict the ratings that a user might provide for an item or sometimes it could even provide an ordered set of items that might be preferred by the user. It is made possible by collecting information about the user's preferences for a set of given items and also if possible about the items and users themselves. The preference information could either be collected explicitly through reviews and ratings or can be collected implicitly [1, 2, 3] by monitoring the user's behavior on a website or app. Further, data from alternate data sources like Social Media Platforms, Location Based Services, etc., are also integrated to provide better recommendations.

A Recommender must make sure that its recommendations are both accurate and at the same time novel and interesting for the users [4]. Early recommenders were modeled based on the way we humans learn about items in general, i.e. based on our own experience and that of others. With the advent of mobile applications and the widespread use of the internet, the need for recommendations on per user basis has increased exponentially [5]. Recommender Systems are currently being employed on a variety of domains including music [6,7,8], movies [9, 10] , books [11, 12], e-commerce products [13, 14], documents [15, 16], web search [17], fine dining, travel, hotel booking, and even news articles and blog posts.

Early recommender systems were based on filtering and the most common filtering methods listed by Pazzani [18] are Content-Based Filtering, Collaborative Filtering and Demographic Filtering. Breese et al. [19] proposed empirical models for evaluating the predictive capacity of the early Collaborative Filtering based Recommender Systems. The advent of the internet and the availability of more and more data further fuelled the evolution of Hybrid Recommender Systems, wherein different types of recommenders were merged to obtain better predictions. This is made possible due to the flexibility and synergistic nature of the various filtering models. Since then, several hybrid recommender models have been proposed to overcome the disadvantages in the individual techniques.

According to Resnick et al. [20], the major aim of a Recommender system is to eliminate or at the least partially overcome the information overload by projecting only the most relevant information and services out of a large pile of data, enabling the possibility of personalized services. The advent of social networking and the ability to harvest network followers and related data has enabled a new class of Social Filtering based Recommender Systems. Neighborhood-based Collaborative Filtering (CF) models were found to be the most popular and Herlocker et al. [21]

has laid down a set of guidelines for architecting such neighborhood-based models.

Every Recommender System model has its own pros and cons. Collaborative Filtering suffers from Data Sparsity, Scalability and Cold Start Problems as reported by Adomavicius et al. [22] and Schafer et al. [23]. Content-Based Models suffer from Limited Content Analysis, Overspecialisation and Incoherent Items in user profile information as stated by Boratto et al. [24]. Model-Based or Machine Learning approaches suffer from non-intrusiveness and overfitting as stated by Kunaver et al. [25]. To overcome these problems, several hybrid models based on Social Filtering [26], Context-Aware Filtering [27, 28], Knowledge-Based Filtering [29], Case-Based Reasoning [30], Computational Intelligence Based Recommendations [31, 32] and Group Recommendation Techniques [33, 34] have been proposed in the literature.

Many real-world recommender system applications have been developed based on the hybrid recommendation techniques. Recent research points out that the focus of Recommender Systems Research in the current big data era is towards application study. Improvements in accessibility and mobility have increased the dependency on e-commerce, e-library, e-learning, e-government, e-tourism and e-business services. Domain and application specific case studies are also found in the literature. Personalization of E-Government interfaces and E-government service recommendations includes government to citizen (G2C) [35, 36] and government to business (G2B) [37, 38] services. Further Recommender Systems focussing on Business to Consumer (B2C) and Business to Business (B2B) [39, 40] users were also proposed in the literature.

Rest of the paper is organized as follows, Section I contains the introduction about recommendation systems and their significance, Section II presents the taxonomy of existing Recommender Systems, Section III explains the research challenges in Recommender Systems, Section IV discusses the various approaches for building Hybrid Recommender Systems, Section V explains the various metrics used for evaluation of Recommender Systems, Section VI provides the list of public datasets available for building Recommender Systems, Section VII provides future research directions and finally Section VIII provides the conclusion.

II. RECOMMENDER SYSTEMS: TAXONOMY

Traditional recommender systems and a taxonomy of how the classical algorithms, methods, filtering approaches and databases, etc., relate to one another was provided by Bobadilla et al. [41]. A modified and enhanced version of the taxonomy which includes recent approaches and techniques has been shown in Figure 1. Creation of a Recommender system involves several decisions and it is based on several constraints. First and foremost, it depends on the type of data available (e.g., reviews, ratings, user demography, features and content about items, social relationships and location information). Secondly, selection the filtering approach to be used: Simple or hybrid (selection of approaches that are to be hybridized). Third being the choice of model: Memory-based or model-based approach and the selection of technique(s) being employed like ANN, Bayesian, Fuzzy, Machine Learning (ML), Singular Value Decomposition (SVD), etc. Finally, the objective of the Recommender is of utmost importance: to provide predictions or top K recommendations like Hit Ratio and Normalized Discounted Cumulative Gain (NDCG) and also the choice of metrics to evaluate the same. Figure 1 clearly shows some of the choices involved in the creation of a Recommender System.

A. Content-Based Recommendation Techniques

Content-Based (CB) recommendation techniques utilizes the previous preferences of users to recommend items or services [42, 43]. CB creates a user profile based on the user's preferred items or services and recommends items that have high similarity to those in the user's profile. Though traditional CB models relies on Cosine Similarity, building CB models using Statistical methods and Machine Learning methods is also possible. Traditional Content Based Models are no longer used in isolation due to the associated research challenges like Limited Content Analysis and Overspecialization. Several hybrid models employing Content Based systems in collaboration with other models tends to outperform the standalone recommender models both in terms of accuracy and coverage.

B. Collaborative Filtering Based Recommendation Techniques

Collaborative Filtering (CF) based models recommends items to the user based on the ratings of users who share similar interests [44, 45, 46]. Collaborative Filtering can be further divided into User Based CF and Item-Based CF. In User-based CF, recommendations are based on items liked by similar users. But users are fickle as their taste and preferences do change over time. In Item Based CF, recommendations are based on the items that the user has rated or reviewed or liked in the past. Similarity between items or users could be measured based on the following measures or metrics: Cosine Similarity (COS), Pearson Correlation (CORR), Constrained Correlation (CCORR), Adjusted Cosine (ACOS), Mean Squared Differences (MSD) and Euclidean Distance (EUC).

C. Demographic Filtering Based Recommendation Techniques

Demographic Filtering (DF) [18, 47, 48] provides recommendations based on the similarity in the user's demographic information. It assumes that users with common demographic attributes like age, sex, occupation, location, country, race, etc. will have similar interests. Irrespective of the influence of demography on prediction accuracy, it is very hard to gain access to demographic information, especially on large-scale databases. Several concerns regarding the privacy and security of user demographic data leads to the fact that most public datasets have little to no user information and if it contains user demography, it will be highly anonymized.

D. Hybrid Recommendation Techniques

Burke [49] proposed seven basic hybridization mechanisms for combining existing Recommender System techniques to achieve better prediction. Hybridization could be used to achieve better performance and in some cases to overcome the drawbacks of a traditional recommendation technique. Two or more basic recommendation techniques can be combined to create a hybrid technique that surpasses the individual techniques in terms of predictions accuracy and effectiveness. Common practice is to combine Collaborative Filtering Technique with other techniques like CB, or DF to overcome the challenges faced by traditional CF models like Data Sparsity, Cold-Start, etc.

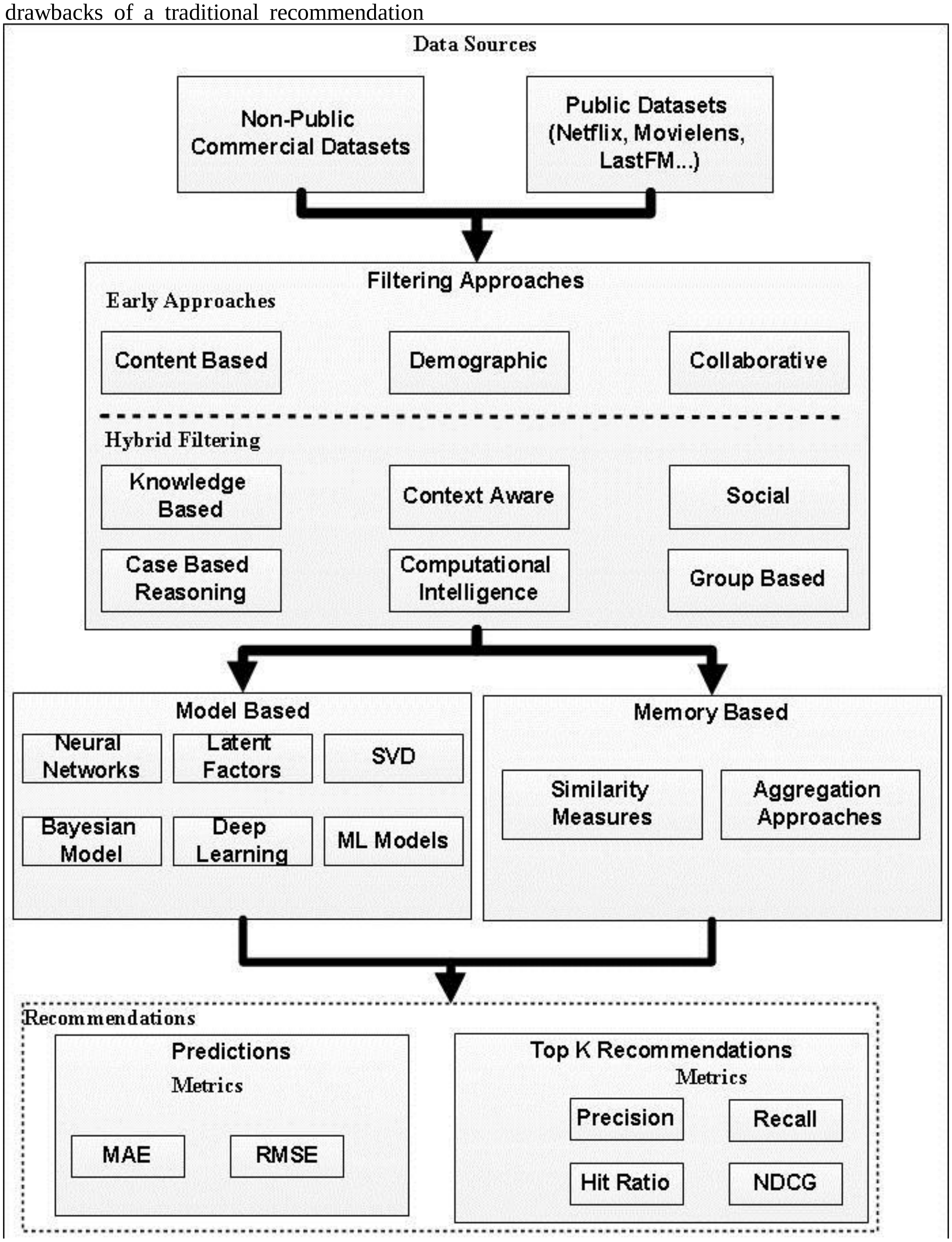

Figure 1: Taxonomy of Recommender Systems: Models and Relationships

E. Knowledge-Based Recommendation Techniques

Knowledge-Based (KB) techniques recommend items based on the inferences about the relationship between a user's need and a possible recommendation [29]. Such systems construct and maintain a knowledge base that contains vital information about how a specific item meets a user's need. Case-Based Reasoning [30], Ontology-Based Models [50], and models based on Semantic similarities of items [51] are some examples of Knowledge-Based systems. Knowledge-Based systems are often used in conjunction with Context-Aware Recommender Systems to achieve better predictions and also to improve their practical efficacy.

F. Computational Intelligence Based Recommendation Techniques

Computational Intelligence (CI) involves the use of techniques like Artificial Neural Networks, Bayesian Inference, Evolutionary Algorithms and Fuzzy Sets. Yu et al. [52] proposed a Hierarchical Bayesian Network based framework that effectively combines both Content-Based and Collaborative Filtering Based approaches. Christakou et al. [53] proposed a hybrid model that utilized a trained ANN for representing individual user preferences and was proven to be effective in combining CB and CF-based models for movie recommendations. Genetic Algorithm, Clustering and several Stochastic Search Techniques have also been successfully applied to combine existing models to produce effective predictions [54]. Fuzzy set theory offers the flexibility to handle non-stochastic uncertainty. It is found to be suitable to handle imprecise information and also accommodates gradients in user preferences.

G. Social Network Based Recommendation Techniques

The unprecedented growth of social networking platforms and the availability of social interaction profiles led to the development of recommender systems based on Social Network Analysis (SNA). Most real-time systems, provide an opportunity to interact socially with fellow users and makes use of this information to provide better recommendations. This in collaboration with Collaborative Filtering models helps overcome the data sparsity problems. Trust-based models for Collaborative Filtering on Social Networks is a widely discussed topic. Several studies prove that there is a positive correlation between trust and user similarity in online communities [55]. Social bookmarks, tags, physical context and co-authorship relationships are currently being exploited to provide better recommendations.

H. Context-Aware Recommendation Techniques

Context Awareness in terms of time, geolocation, presence of friends, families or colleagues and how that affects the interaction between the user and the application is currently being exploited to provide better recommendations. Especially for domains like fine dining, travel planning, hotel reservations, creating vacation packages and even for simple cases like suggesting combos or offers for a user ordering a pizza (whether from office or home?), the importance of context cannot be ignored. According to Adomavicius et al. [56], in technology-enhanced environments context becomes a multifaceted concept. It transforms the problem of recommendation from a 2D space to a multidimensional space. It is no longer the prediction of rating for a given user-item pair, instead, it is the prediction of rating for a user-item pair, given the context.

I. Group Recommendation Techniques

Group Recommender Systems (GRS) often involves aggregation strategies based on social choice theory and consensus decision making [33, 34]. It provides a group of user suggestions when they are unable to meet each other for negotiation or in cases where their preferences are not clear in spite of them meeting each other. Several strategies like least misery, most pleasure, and their fuzzy adaptations are most commonly used in GRS. Application areas include travel planning and accommodation for groups, ski holiday suggestions, choice of news articles and blog posts to be include in the mailer for subscribers, etc.

III. CHALLENGES IN RECOMMENDER SYSTEMS

To provide improved user experience, the recommender system should not only be accurate but also provide recommendations that are both novel and interesting. It is found that the introduction of diversification into the recommendation process offers a lot of potential for new developments. Several challenges are encountered during the creation of a recommender system that is balanced in terms of its prediction Accuracy, Coverage, Diversity, Hit Ratio and F-Measure. In this work, we discuss those challenges that are most often encountered while creating a balanced, well performing Recommender System.

A. Data Sparsity

Data sparsity is one problem that affects all Recommender System approaches both Content-Based and Collaborative Filtering. The fundamental problem is that the user-item rating matrix is sparse. But one must understand that it is impossible for every user to rate and review every item and if that happens the very idea of recommendation fails as there will be nothing left to recommend. Data sparsity makes it impossible to locate successful neighbors and thus results in the creation of a weak recommender. As the user base and the number of items on sale grows, the sparsity of the user-item matrix grows exponentially. Data Sparsity affects the *coverage* of a recommender system, which refers to the percentage of items in the system for which a recommendation can be made. It is therefore mandatory for a recommender system to inherently handle the data sparsity problem. Data from alternate sources like social media, third-party datasets and even questionnaires, etc., can be effectively put into use to address the data sparsity issue, but the reliability of such data sources is still a question of importance.

B. Cold Start

Cold start refers to a scenario where it is not possible to provide reliable recommendations due to lack of ratings or user profile information. Though the new user problem is heavily discussed, the new community and new item problems are also found to be equally important, when it comes to the creation of a successful recommender system. New user problem represents the scenario in which the user would not have provided enough ratings to get personalized recommendations. New users often tend to feel that they are being ignored and may even leave the service. The new item problem is much more difficult to address as that leads to a cycle where new items will not be recommended and hence be ignored by a vast group of users and this in turns prevents those items from getting the required number of ratings to get noticed. This has less impact in domains where there are more than one way to discover new items (movies), but not so in other domains like e-commerce, blog posts, etc.

C. Over-fitting

Once the Recommender System overcomes the above-mentioned problems like cold start and data sparsity and starts generating consistent and reliable recommendations for each user, a new problem arises. The problem is that the recommender sticks on to a very narrow spectrum of items strongly determined by the users' preferences. But users will have multiple interests and most of them will not be known to the system apriori due to lack of ratings or review. This often results in what is called overfitting or overspecialization where the system will not recommend anything out of the users' interests. This lack of novelty and diversity may even push the user to quit the service or platform altogether.

D. Magic Barrier Problem

It could be observed from the literature that a part or subset of ratings given by the user might be considered as anomalies or outliers, as the same user may provide different ratings to a given item under different contexts. Though the context is mostly temporal, it could also be geolocation or others. This problem is referred to as the *magic barrier* problem by Boratto et al. [24]. It states that, the recommender system will reach a point beyond which its accuracy cannot be actually improved. I.e. any improvement after reaching the magic barrier is caused due to overfitting and will not result in enhanced system performance. This is caused due to the noise in the data and is often ignored. The impact might be high in Content-Based systems due to the way in which item and user profiles are constructed.

E. Limited Content Analysis

Limited Content Analysis is found to be one of the most challenging problems faced by Content-Based Recommender Systems in reality. It can be considered as an aspect or feature extraction problem and is often caused by the difficulty in extracting reliable information from a variety of sources (images, audio, video and text) using automated mechanisms. The issues surrounding classic information retrieval techniques also affects the performance of the Recommender Systems. For certain domains like music, blogs and videos, the task of generating or extracting attributes is very complex and may result in the creation of weak recommenders.

F. Presence of Incoherent Items in User Profile

The assumption about user preferences being unchanged over time affects the performance of the Recommender Systems by introducing incoherent and misleading elements in the user profile. Temporal dynamics is often ignored both during the creation and maintenance of user profile information. Another major factor is the temporary use of a user's account by other people. This adds fuel to the fire and further degrades the performance of the recommender. In order to handle both these issues, the system should periodically unlearn some of the weaker elements or entities and this will further help in increasing the diversity of the recommender as a whole.

G. Scalability

Even though data sparsity is considered one of the most important challenges for any recommender system, one cannot ignore the opposite side of the spectrum. Since most of these systems handle millions of users and items, it is crucial to have an effective and efficient a storage structure and an equally effective processing model. This makes it mandatory to apply recommendation techniques that are both scalable in nature and parallelizable to some extent to meet the real-time requirements. Even though dimensionality reduction techniques like Singular Value Decomposition could be used to reduce the complexity, it is also found to affect the interpretability of the models to a greater extent.

H. Other Challenges

In Content-Based approaches, it is almost impossible to acquire feedback about the performance of the recommender as the user will not provide any explicit feedback by means of rating or review. In such cases dependency on implicit feedback mechanisms and other alternate data sources becomes the only option and the reliability and usability constraints associated with those methods further affects the performance of the recommender. Synonymy is another important factor that affects the performance of Content-Based approaches. It is defined as the tendency of very similar items to have different names. Automated systems often find it difficult to clearly delineate or distinguish the closely related items that are labeled or branded differently. It is further complicated because, in most domains, branding makes all the difference.

IV. HYBRID RECOMMENDER SYSTEM APPROACHES

As discussed in Section II, hybrid recommender systems involves the combination of one or more pure recommender models to overcome the disadvantages of those individual techniques and the goal is to attain synergy. The hybridization can be achieved through several means: separate implementation of individual approaches followed by a combination of results, combining Content-Based Filtering and Collaborative Filtering Approaches in some form to bring the best of each, or could even create a unified hybrid recommendation model that brings several approaches together to surpass the effectiveness of individual models. This work analyses seven commonly used hybridization approaches from the perspective of implementation and usability.

A. Weighted Hybridization

This Methodology refers to a form of hybridization where the results of several different recommenders implemented separately are combined numerically. Initially, each of the individual methods will be assigned an equal weight but later they are adjusted based on their prediction capabilities. The major advantage of a weighted hybridization is that the strength of all the recommenders is utilized during the process in a straightforward manner. P-tango [57] is an example of a weighted Recommendation system.

B. Switching Hybridization

From an implementation standpoint it is similar to the weighted hybridization, where we implement the individual models separately and the difference is that, instead of giving each model a weight, we switch the models that is used for providing predictions dynamically based on a heuristic that reflects the recommender's ability to produce good rating. This kind of hybridization is sensitive to both the strengths and weaknesses of the contributing models. The major disadvantage of the model is that the heuristic switching mechanism introduces unnecessary complexity. DailyLearner [58] is an example of a switching hybrid that employs CB and CF-based approaches.

C. Cascade Hybridization

Cascade hybridization is one where the recommenders are given a strict order of preference. It applies an iterative process and earlier or lower level models provide coarse recommendations followed by higher or later models that provide finer recommendations. Low priority model breaks ties in the scoring of high priority ones. This technique is found to be tolerant to noise in the data due to the coarser-to-finer iterative nature of the recommendation process. EntreeC [59] is a knowledge-based collaborative recommender that makes use of Cascade Hybridization.

D. Mixed Hybridization

From an implementation standpoint it mimics the weighted model, but instead of assigning weights for each of the models and providing a single recommendation, it takes into consideration the recommendations of all those individual models and presents the results together. Here the general performance of the model is not affected by the performances of the individual models and hence achieves better coverage and diversity. PTV [60] is a TV listing and scheduling service that uses mixed hybridization.

E. Feature Combination

In feature combination, features extracted from several heterogeneous sources are combined together to act as input to a single recommender algorithm that provides the final recommendations. Sometimes the features are extracted by using other simple recommendation models. Piper [61] is an example of a recommender created using feature combination.

F. Feature Augmentation

Libra [62] is an example of a feature augmentation based recommendation model. It makes use of a Content Based recommender to recommend books on Amazon.com by employing a Naïve Bayes text classifier for creating the input features. It makes use of the ratings and other information produced by previous recommenders as input and are always found to be superior to feature combination based models.

G. Meta-Level Hybridization

It is based on the assumption that the output of a recommender model contains or captures more information than the actual input data and thus utilizes the output of several such models as input to another meta-level recommender that yields the final recommendations. This helps address the data sparsity problem effectively. LaboUr [63] is an example of a meta-level recommender system.

V. METRICS FOR EVALUATION OF RECOMMENDER SYSTEMS

Recommender Systems could either predict how much a user might prefer an item or it could provide a set of Top-K items that the user may be interested in. Metrics for evaluating the effectiveness of a Recommender System is usually classified into three broad categories namely, Prediction Metrics, Set Recommendation Metrics and Rank Recommendation Metrics. These metrics facilitate comparisons of several solutions for the same problem and allows selection of the most promising solutions. Due to these metrics, Recommender Systems have been continuously tested and improved.

A. *Ratings Prediction Metrics:*

Rating Prediction metrics are oriented towards the accuracy of the Recommender System. It helps us to measure how close the predicted ratings are to the actual ratings. *Mean Absolute Error (MAE)* and *Root Mean Square Error (RMSE)* are the most commonly used metrics. Low values of RMSE and MAE indicates the high predictive power of the models. RMSE tends to penalize the large errors heavily as it squares the error values, whereas MAE takes a linear approach towards the same. Additionally, *Normalized Mean Average Error (NMAE)* and *Coverage* which is a representation of the percentage of items in a system for which recommendations can be made, are also used.

B. *Set Recommendation Metrics:*

These metrics are based on the proportion between recommended items and expended items. Set Recommendation Metrics includes *Precision or True Positive Rate, Recall or Sensitivity, Specificity or True Negative Rate, F-measure* which is a family of metrics that combines *precision and recall, Area Under the Curve (AUC)* which is a more robust metric that effectively considers the variations between recall and specificity, and *Benefit Ratio* which refers to the ratio of users who got improved prediction to users who got deteriorated prediction.

C. *Rank Recommendation Metrics:*

Rank recommendation metrics are utilized when the recommender system return an ordered list of items which the user might prefer. The position of the item in the ordered list is considered to be proportional to the item's utility and helps us to measure the same. Common metrics include *Half-Life* which assumes that the interest of users decreases exponentially as they move away from recommendations at the top of the list, *Discounted Cumulative Gain (DCG) and Normalized Discounted Cumulative Gain (NDCG)* which considers that highly ranked items gives more satisfaction in comparison to poorly ranked items, *Hit Ratio* which presents a measure of the user's choice being in the top-K recommendations, *Mean Reciprocal Rank (MRR)* which measures the ranking position of the user's choice in the top-K recommendations, and *the Mean Average Precision (MAP)* which measures the precision of the first k recommended ranked items.

VI. DATASETS FOR RECOMMENDER SYSTEMS

Any data-driven system including Recommender Systems could be only as good as the data that it uses. Over a period of several years, many research communities and commercial organizations have gathered user ratings and review data for developing and testing recommender systems. Continued research and publications has further enhanced the data and some even pre-processed it for recommender system use cases. Several studies on gathering related information for building enhanced user/item profiles based on alternate data sources has also been reported in the literature. Table 1 shows 18 such datasets that are available in the public domain along with their domain of interest, a brief description of the features and the public URL in which it is available for download.

Table I. Datasets for Recommender Systems

Dataset Name	Domain(s)	Brief Description	URL
Movielens 100K	Movies	Movies, User, Ratings and User Demography	http://grouplens.org/datasets/movielens/100k/
Movielens 20M	Movies	Movies, Ratings, User, No Demographic Data	http://grouplens.org/datasets/movielens/20m/
Netflix Movie	Movies	Movies, Rating, ReviewDate	https://www.kaggle.com/netflix-inc/netflix-prize-data/data
Movie pilot	Movies	Context-Aware Data, Movies	http://camra2010.dai-labor.de/index.html%3Fp=49.html
Filmtipset	Movies	Context-Aware Data, Movies	http://camra2010.dai-labor.de/index.html%3Fp=49.html
IMDB Dataset	Movies	Movies, Ratings, Votes	https://www.imdb.com/interfaces/
Yahoo Music	Music	Music, Ratings, Genres, Taxonomy	http://www.kdd.org/kdd2011/kddcup.shtml
LDOS	Movies	Movies, Users, Ratings	http://www.lucami.org/index.php/research/ldos-comoda-dataset/?lang=en
Last.fm	Music	Music, Artist, User	http://www.dtic.upf.edu/~ocelma/MusicRecommendationDataset/
Douban	Movies, Books, Music	Users, Social Connections	http://socialcomputing.asu.edu/datasets/Douban
Amazon Reviews	Movies, books, FineFood, etc.,	Reviews, Ratings, Timestamp	https://snap.stanford.edu/data/web-Amazon.html
BeerAdvocate	Beers	Beers, Ratings for 5 aspects, Review Text	https://snap.stanford.edu/data/web-BeerAdvocate.html
RateBeer	Beers	Beers, Ratings for 5 aspects, Review Text	https://snap.stanford.edu/data/web-RateBeer.html
CellarTracker	Wines	Wines, Users, Ratings,	https://snap.stanford.edu/data/web-CellarTracker.html

		Review Text	
TripAdvisor Dataset	Travel	Text, Aspects, Users	http://times.cs.uiuc.edu/~wang296/Data/
Jester Dataset	Jokes	Users, Ratings, Jokes	http://goldberg.berkeley.edu/jester-data/
Delicious Dataset	Cognitive Authority	Users, Authority, Tags, Network	http://www.din.uem.br/~gsii/delicious-dataset/
Epinions Dataset	Product Reviews	Trust-based, Network, Ratings, Reviews	http://www.trustlet.org/epinions.html

VII. Research Directions

It could be found from the literature that the classic recommendation approaches like Content-Based Filtering, Collaborative filtering and Demographic filtering plays a dominant role in all kinds of application domains, but at the same time hybrid systems are found to be more popular than the individual approaches due to their versatility and effectiveness. Several applications domains including B2C, B2B, G2C and G2B applications are analysed and it could be seen that most of these e-services are now turning towards mobile-based approaches. Mobile-based recommendation has advantages in terms of being more context-aware by means of location services and more personal by being more restrictively used. But such mobile-based recommendation also has the drawbacks that it requires the system to have spatial-temporal auto-correlation and the options to handle heterogeneity and noise.

Temporal and context-sensitive dynamism still remains to be challenge that is yet to be addressed. Addressing the concept drift will enable effective context-aware e-shopping and e-learning platforms that can accommodate and improve the performance of recommenders in fast-changing environments. With the advent of big data and deep neural network based learning models, the problem of data sparsity could effectively be addressed using Transfer Learning mechanisms. The impact that smart wearable devices and the Internet of Things (IoT) will bring into the field of recommendations is enormous and it could effectively enable better healthcare recommendation services. The use of fuzzy systems for understanding and effectively modeling the granularity of the user preferences, could further enable effective fine-grained hyper-personal recommendations in several domains.

VIII. Conclusion

Recommendation systems helps us out of the information overload that is bestowed upon us by the information era. In this work we have presented an overview of the recommendation system process, its taxonomy and the various traditional and hybrid techniques for creating recommender systems. It also discusses the various research challenges associated with the creation of a well-balanced Recommender system. Several approaches for creation of hybrid recommender systems has been discussed from an implementation standpoint. This work also presents a huge set of publicly available datasets for use with recommendation systems. The development of frameworks for automated analysis of heterogeneous data, establishing standards for evaluation metrics, creation of attack based evaluation models that helps us measure and understand the effectiveness and behaviour of the recommender systems under real-time spam scenarios are some of the challenges that are yet to be addressed.

IX. References

[1] S.K. Lee, Y.H. Cho and S.H. Kim, “Collaborative filtering with ordinal scale-based implicit ratings for mobile music recommendations,” Information Sciences. Vol. 180 (11) ,2010, pp. 2142–2155.

[2] K. Choi, D. Yoo, G. Kim and Y. Suh, “A hybrid online-product recommendation system: combining implicit rating-based collaborative filtering and sequential pattern analysis,” Electronic Commerce Research and Applications, in press, doi: 10.1016/j.elerap.2012.02.004.

[3] E.R. Núñez-Valdéz, J.M. Cueva-Lovelle, O. Sanjuán-Martı´nez, V. Garcı ´ a-Dı´az, P. Ordoñez and C.E. Montenegro-Marı´ n, “Implicit feedback techniques on recommender systems applied to electronic books,” Computers in Human Behavior. Vol. 28 (4) ,2012, pp. 1186–1193.

[4] D. Kotkov, S. Wang and J.Veijalainen, “A survey of serendipity in recommender systems,” *Knowledge-Based Systems*, *111*, 2016, pp.180-192.

[5] D.H. Park, H.K. Kim, I.Y. Choi and J.K. Kim, “A literature review and classification of recommender Systems research,” Expert Systems with Applications 39 ,2012, pp. 10059–10072.

[6] S.K. Lee, Y.H. Cho and S.H. Kim, “Collaborative filtering with ordinal scale-based implicit ratings for mobile music recommendations,” Information Sciences. Vol. 180 (11) ,2010, pp. 2142–2155.

[7] A. Nanolopoulus, D. Rafailidis, P. Symeonidis and Y. Manolopoulus, “Music Box: personalizad music recommendation based on cubic analysis of social tags,” IEEE Transactions on Audio, Speech and Language Processing. Vol. 18 (2) ,2010, pp. 407–412.

[8] S. Tan, J. Bu, CH. Chen and X. He, “Using rich social media information for music recommendation via hypergraph model,” ACM Transactions on Multimedia Computing, Communications, and Applications. Vol. 7 (1) 2011. Article 7.

[9] W. Carrer-Neto, M.L. Hernández-Alcaraz, R. Valencia-Garcı´ a and F. Garcı´a- Sánchez, “Social knowledge-based recommender system,” Application to the movies domain. Expert Systems with Applications. Vol. 39 (12) ,2012, pp. 10990–11000.

[10] P. Winoto and T.Y. Tang, “The role of user mood in movie recommendations,” Expert Systems with Applications. Vol. 37 (8), 2010, pp. 6086–6092.

[11] E.R. Núñez-Valdéz, J.M. Cueva-Lovelle, O. Sanjuán-Martı´nez, V. Garcı ´ a-Dı´az, P. Ordoñez and C.E. Montenegro-Marı´ n, “Implicit feedback techniques on recommender systems applied to electronic books,” Computers in Human Behavior. Vol. 28 (4) ,2012, pp. 1186–1193.

[12] R. González-Crespo, O. Sanjuán-Martı´nez, J. Manuel-Cueva, B. Cristina-Pelayo, J.E. Labra-Gayo and P. Ordoñez, “Recommendation system based on user interaction data applied to intelligent electronic books,” Computers in Human Behavior. Vol. 27 (4), 2011, pp.1445–1449.
[13] Z. Huang, D. Zeng and H. Chen, “A comparison of collaborative filtering recommendation algorithms for e-commerce,” IEEE Intelligent Systems. Vol. 22 (5), 2007, pp. 68–78.
[14] J.J. Castro-Sanchez, R. Miguel, D. Vallejo and L.M. López-López, “A highly adaptive recommender system based on fuzzy logic for B2C e-commerce portals,” Expert Systems with Applications. Vol. 38 (3),2011, pp. 2441–2454.
[15] J. Serrano-Guerrero, E. Herrera-Viedma, J.A. Olivas, A. Cerezo and F.P. Romero, “A google wave-based fuzzy recommender system to disseminate information in University Digital Libraries 2.0,” Information Sciences. vol. 181 (9), 2011, pp. 1503–1516.
[16] C. Porcel and E. Herrera-Viedma, “Dealing with incomplete information in a fuzzy linguistic recommender system to disseminate information in university digital libraries,” Knowledge-Based Systems 23 (1), 2010, pp. 32–39.
[17] K. Mcnally, M.P. O’mahony, M. Coyle, P. Briggs and B. Smyth, “A case study of collaboration and reputation in social web search,” ACM Transactions on Intelligent Systems and Technology. Vol. 3 (1) ,2011. Article 4.
[18] M. Pazzani, “A framework for collaborative, content-based, and demographic filtering,” Artificial Intelligence Review-Special Issue on Data Mining on the Internet. Vol.13 (5-6), 1999, pp. 393–408.
[19] J.S. Breese, D. Heckerman and C. Kadie, “Empirical analysis of predictive algorithms for collaborative filtering,” in: 14th Conference on Uncertainty in Artificial Intelligence, 1998, pp. 43–52.
[20] P. Resnick and H.R. Varian, “Recommender systems,” Communications of the ACM 40 ,1997, pp. 56–58.
[21] J.L. Herlocker, J.A. Konstan and J.T. Riedl, “An empirical analysis of design choices in neighborhood-based collaborative filtering algorithms,” Information Retrieval 5 , 2002, pp. 287–310.
[22] G. Adomavicius and A. Tuzhilin, “Toward the next generation of recommender systems: a survey of the state-of-the-art and possible extensions,” IEEE Transactions on Knowledge and Data Engineering 17 ,2005, pp. 734–749.
[23] J.B. Schafer, D. Frankowski, J. Herlocker and S. Sen, “Collaborative filtering recommender systems,” in: P. Brusilovsky, A. Kobsa,W. Nejdl (Eds.), The Adaptive Web, Springer, Berlin Heidelberg 2007, pp. 291–324.
[24] L. Boratto, S. Carta, G. Fenu and R. Saia, “Semantics-aware content-based recommender systems: Design and architecture guidelines,” Neurocomputing, Vol. 254, pp. 79-85, September 6, 2017.
[25] “Diversity in recommender systems – A survey Matevž Kunaver, Tomaž Požrl, Knowledge-Based Systems,” Vol. 123, pp. 154-162, 1 May 2017.
[26] D. Ben-Shimon, A. Tsikinovsky, L. Rokach, A. Meisles, G. Shani and L. Naamani, “Recommender system from personal social networks,” Advances in Intelligent Web Mastering, Springer, 2007, pp. 47–55.
[27] W. Woerndl,M. Brocco and R. Eigner, “Context-aware recommender systems inmobile scenarios,” International Journal of Information Technology andWeb Engineering 4 ,2009, pp. 67–85.
[28] K. Verbert, N. Manouselis, X. Ochoa, M. Wolpers, H. Drachsler, I. Bosnic and E. Duval, “Context-aware recommender systems for learning: a survey and future challenges,” IEEE Transactions on Learning Technologies 5 ,2012, pp. 318–335.
[29] S. Middleton, D. Roure and N. Shadbolt, “Ontology-based recommender systems,” in: S. Staab, R. Studer (Eds.), Handbook on Ontologies, Springer, Berlin Heidelberg 2009, pp. 779–796.
[30] B. Smyth, “Case-based recommendation,” in: P. Brusilovsky, A. Kobsa, W. Nejdl (Eds.), The Adaptive Web, Springer, Berlin Heidelberg 2007, pp. 342–376.
[31] X. Amatriain, A. Jaimes, N. Oliver and J. Pujol, “Data mining methods for recommender systems,” in: F. Ricci, L. Rokach, B. Shapira, P.B. Kantor (Eds.), Recommender Systems Handbook, Springer, US 2011, pp. 39–71.
[32] S. Hsu, M.-H.Wen, H.-C. Lin, C.-C. Lee and C.-H. Lee, “AIMED— a personalized TV recommendation system,” in: P. Cesar, K. Chorianopoulos, J. Jensen (Eds.), Interactive TV: a Shared Experience, Springer, Berlin Heidelberg 2007, pp. 166–174.
[33] A. Jameson, B. Smyth, “Recommendation to groups,” in: P. Brusilovsky, A. Kobsa, W. Nejdl (Eds.), The Adaptive Web, Methods and Strategies of Web Personalization, Springer, Berlin Heidelberg 2007, pp. 596–627.
[34] I. Garcia and L. Sebastia, “A negotiation framework for heterogeneous group recommendation,” Expert Systems with Applications 41 ,2014, pp. 1245–1261.
[35] P. De Meo, G. Quattrone and D. Ursino, “A decision support system for designing new services tailored to citizen profiles in a complex and distributed e-government scenario,” Data and Knowledge Engineering 67 ,2008, pp. 161–184.
[36] B. Esteban, Á. Tejeda-Lorente, C. Porcel, M. Arroyo and E. Herrera-Viedma, “TPLUFIBWEB: a fuzzy linguistic Web system to help in the treatment of low back pain problems,” Knowledge-Based Systems 2014, http://dx.doi.org/10.1016/j.knosys.2014.03.004.
[37] C. Cornelis, X. Guo, J. Lu and G. Zhang, “A fuzzy relational approach to event recommendation, Proceedings of the Second Indian International Conference on Artificial Intelligence (IICAI-05),” Pune, INDIA 2005, pp. 2231–2242.
[38] J. Lu, Q. Shambour, Y. Xu, Q. Lin and G. Zhang, “BizSeeker: a hybrid semantic recommendation system for personalized government-to-business e-services,” Internet Research 20 ,2010, pp. 342–365.
[39] T. Lee, J. Chun, J. Shim and S.-g. Lee, “An ontology-based product recommender system for B2B marketplaces,” International Journal of Electronic Commerce 11 ,2006, pp. 125–155.
[40] J.-C. Wang and C.-C. Chiu, “Recommending trusted online auction sellers using social network analysis,” Expert Systems with Applications 34, 2008, pp. 1666–1679.
[41] J. Bobadilla, F. Ortega, A. Hernando and A. Gutiérrez, “Recommender systems survey,” Knowledge-Based Systems 46 ,2013, pp. 109–132.
[42] R. Meteren and M. Someren, “Using content-based filtering for recommendation,” in: Proceedings of ECML 2000 Workshop: Maching Learning in Information Age, 2000, pp. 47–56.
[43] J. Mooney and L. Roy, “Content-based book recommending using learning for text categorization,” in: Proceedings of the Fifth ACM Conference on Digital Libraries, 2000, pp. 195–204.
[44] J.L. Herlocker, J.A. Konstan, A.L. Borchers and J.T. Riedl, “An algorithmic framework for performing collaborative filtering,” in: Proceedings of the 22nd Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, 1999, pp. 230–237.
[45] J.L. Herlocker, J.A. Konstan, J.T. Riedl and L.G. Terveen, “Evaluating collaborative filtering recommender systems,” ACM Transactions on Information Systems. Vol. 22 (1) ,2004 pp. 5–53.
[46] J. Bobadilla, F. Ortega and A. Hernando, “A collaborative filtering similarity measure based on singularities,” Information Processing and Management. Vol. 48 (2) ,2012, pp. 204–217.
[47] B. Krulwich, “Lifestyle finder: intelligent user profiling using large-scale demographic data,” Artificial Intelligence Magazine. Vol. 18 (2) ,1997, pp. 37–45.

[48] C. Porcel, A. Tejeda-Lorente, M.A. Martı´ nez and E. Herrera-Viedma, "A hybrid recommender system for the selective dissemination of research resources in a technology transfer office," Information Sciences. Vol. 184 (1) ,2012, pp. 1–19.
[49] R. Burke, "Hybrid web recommender systems," in: P. Brusilovsky, A. Kobsa,W. Nejdl (Eds.), The Adaptive Web, Springer-Verlag, Berlin Heidelberg 2007, pp. 377–408.
[50] S. Middleton, D. Roure and N. Shadbolt, "Ontology-based recommender systems," in: S. Staab, R. Studer (Eds.), Handbook on Ontologies, Springer, Berlin Heidelberg 2009, pp. 779–796.
[51] I. Cantador, A. Bellogín and P. Castells, "A multilayer ontology-based hybrid recommendation model," AI Communications 21 ,2008 pp. 203–210.
[52] K. Yu, V. Tresp and S. Yu, "A nonparametric hierarchical bayesian framework for information filtering," Proceedings of the 27th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, ACM, Sheffield, United Kingdom 2004, pp. 353–360.
[53] C. Christakou, S. Vrettos and A. Stafylopatis, "A hybrid movie recommender system based on neural networks," International Journal on Artificial Intelligence Tools 16 ,2007 pp. 771–792.
[54] K.-j. Kim and H. Ahn, "A recommender systemusing GA K-means clustering in an online shopping market," Expert Systems with Applications 34 ,2008, pp. 1200–1209.
[55] C.-N. Ziegler and G. Lausen, "Analyzing correlation between trust and user similarity in online communities," Trust Management, Springer, 2004 ,pp. 251–265.
[56] G. Adomavicius and A. Tuzhilin, "Context-aware recommender systems," in: F. Ricci, L. Rokach, B. Shapira, P.B. Kantor (Eds.), Recommender Systems Handbook, Springer, US 2011, pp. 217–253.
[57] M.Claypool, A. Gokhale, T. Miranda, P. Murnikov, D. Netes and M. Sartin, "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.
[58] D. Billsus and M.J. Pazzani, " A hybrid user model for news story classification," In: Kay J, editor. In: Proceedings of the seventh international conference on user modeling, Banff, Canada. Springer-Verlag, New York; 1999, pp. 99–108.
[59] R. Burke, "Hybrid recommender systems: survey and experiments," User Model User-adapted Interact, Vol.12(4), 2002, pp.331–70.
[60] B. Smyth and P.A. Cotter, "personalized TV listings service for the digital TV age," J Knowl-Based Syst Vol. 13(2–3), 2000, PP.53–9.
[61] C. Basu, H. Hirsh and W. Cohen, "Recommendation as classification: using social and content-based information in recommendation," In: Proceedings of the 15th national conference on artificial intelligence, Madison, WI; 1998. pp. 714–20.
[62] R.J. Mooney and L. Roy, "Content-based book recommending using learning for text categorization," In: Proceedings of the fifth ACM conference on digital libraries. ACM; 2000, pp. 195–204.
[63] I. Schwab, A. Kobsa and I. Koychev, "Learning user interests through positive examples using content analysis and collaborative filtering,"s Draft from Fraunhofer Institute for Applied Information Technology, Germany; 2001.

Authors Profile

Ms.K.Reka is an Assistant Professor in Department of Computer Science at Cauvery College For Women, TamilNadu,India.Currently, she is pursuing her Ph.D at Periyar E.V.R College, TamilNadu, India. Her research interests include Machine Learning and Bigdata Analytics.

Dr. T.N. Ravi, born on 01.06.1964, has started his teaching career in the year 1985. Currently, he is working as Assistant Professor of Computer Science, Periyar E.V.R. College (Autonomous), an institution run by the Director of Collegiate Education, Government of Tamil Nadu, India. He is basically a Mathematics Graduate, completed both Under Graduation and Post Graduation.

He had completed B.Ed., in Mathematics, P.G.D.C.A., and then completed M.Sc. & M.Phil. in Computer Science. He has successfully completed his Ph.D. in Computer Science in the year 2010.

He pursued his research in the area of Parallel Computing and also did real time testing at National Aerospace Laboratories, Bangalore, India. His area of research interests are Parallel Processing, Network Computing, Genetic Algorithms, Artificial Intelligence and Data Mining. He has guided about 40 M.Phil Scholars and currently guiding 8 Ph.D.Scholars.

He has published 39 research articles in International / National Journals. He has delivered about 30 invited talks at State Level and National Level Conferences. He acted as Chair for several Conferences at State and National Level. He has rendered his academic services to various Universities,Institutions within and outside Tamil Nadu, Tamil Nadu Public Service Commission Board Examinations etc.,

