Indian Journal of Science and Technology, Vol 9(20), DOI: 10.17485/ijst/2016/v9i20/88005, May 2016 ISSN (Print) : 0974-6846 ISSN (Online) : 0974-5645

Evolution of Recommender Systems from Ancient Times to Modern Era: A Survey

Richa Sharma* and Rahul Singh

University Institute of Engineering, Chandigarh University, Gharuan, Mohali - 140413, Punjab, India; richas30@yahoo.com, rahulsinghcse25@gmail.com

Abstract

Recommender systems were introduced in the mid-1990s to help people select the most suitable product for them from the plethora of options available with them. The idea that led to their development was that we people often rely on the opinions of our peers before trying something new, say it be before buying a smart phone, a laptop, before going for a movie, before going to a new restaurant and even before visiting a doctor. Till date, we have numerous recommender systems developed for various areas, using different recommendation approaches. Yet, there are still a few limitations of recommender systems that need to be worked on. In this paper, we present an overview of recommender systems, the various approaches of recommender systems, the application areas for which various recommender systems have been developed and we also present the limitations of recommender systems.

Keywords: Application Ares, Limitations, Metrics, Recommender Systems, Recommendation Appaches, Recommendation Algorithms

1. Introduction

Recommender systems might be a new concept in research area, but it has been prevailing in the society since ages. Humans are known to have evolved and grow certain traits of complex thinking, using language, making tools etc. during the last 100,000 years. The concept of recommendations could be seen in case of cavemen, ants and other creatures too. We may have seen ants running around in our house walking in a line behind the ants that went before and found food. This is because ants have genetically evolved to leave markers for other ants which further serve as a recommender to other ants, showing them the way to food. Ancient civilizations flourished between the period 4000 to 1200 BCE. If we talk about recommendations at that time, these could possibly range from what crop to cultivate and during what time, what religion to follow etc. Later, during the colonization period between the 11^{th}-18^{th} centuries, the recommendations got transformed into what territory to absorb in accordance with the fruitful aspects like land fertility, manpower (for slavery) and numerous other resources.

For example, when there were kingdoms around the globe, the king used to take suggestions from their ministers on almost every matter. There used to be a panel of ministers who were assigned the task of keeping their point of views in front of the king so as to help him take the decision easily. Suggesting what kingdom to take over, what policies to make for the betterment of the masses and other matters concerning the kingdom are the possible examples. These all were recommendations on the basis of which the king used to take decisions.

Also, when in old times families used to arrange marriages then usually someone from the family itself used to come up with some match. Those again were recommendations. Similarly people asking for opinions from others as to where to buy the best goods from or what destination is good for vacation, the shortest path to reach some destination etc. are all recommendations which have been in the society from long back. There were no computers in the old world, but still people used to have their recommendations from their peers. But soon with the Industrial Revolution, not only the choices

*Author for correspondence

in various fields increased but also the era of computer came into existence thereby revolutionizing the global market. People had numerous choices in front of them that often led to the confusion about what product might actually fulfil their requirements. So the need for having a system which could facilitate this selection criteria and eradicating the dilemma of masses, was realized and ultimately recommender systems of present day world were introduced.

Humans undoubtedly rely on recommendations from other sources for one purpose or the other, given the fact that the choices available to us are increasing manifold day by day. While buying a cell phone, we go through numerous sites to check the reviews for the various sets available that we might have interest in; we take suggestions from our peers and so on. So is the case with buying new vehicles, going for movies, reading suggestions and so on. Even matrimony sites have a lot to offer. Earlier what our aunts used to do is now replaced by the matrimony sites thereby saving the effort. So we can conclude recommender systems to be as software tools that narrow down our choices and provide us with the most suitable suggestions as per our requirements.

Recommender systems evolved as an independent research area in the mid 1970s in Duke University. Since then, a lot of work has been done in this area till date with the introduction of various approaches. The first recommender system that came into existence was Tapestry[1] which was developed at the Xerox Palo Alto Research Centre. The motivation that led to its development was the growing number of incoming emails, mostly unnecessary ones that were too annoying sometimes and difficult to maintain. So what was done to overcome this problem was that users had their mailing lists made and only those people who were in the contact lists could send the mails to them or the ones the user might want to hear from while the others were sent to the spam list, exactly how it works in present in our email accounts. For this collaborative filtering was used. Soon recommender systems got more popular and now play an important role in internet sites like Amazon, Pandora, Netflix, Matrimonial sites, Social networking sites, YouTube, Yahoo, Tripadvisor etc. RSs did not travel this whole path alone, rather included Artificial intelligence, Information retrieval and Human-computer interaction in their journey[2], and hence got more efficient and gained more popularity.

Till today, researchers have developed many recommender systems for almost every domain like entertainment, social-networking sites, content based sites (e-learning, books or articles recommendation, e-filtering etc.), e-commerce, tourism, match-making and a lot more, all dealing with real-world[3]. Thus, this paper gives an overview of various recommender systems developed so far and their application areas. We have categorised recommender systems on the basis of their application areas: Entertainment, Social-networking based, Content-based, Services-based and E-commerce. For each category, we have presented the most suitable examples along with the recommendation approach used.

Numerous survey papers on Recommender systems have been published till date. For example,[4] in their paper reviewed more than 200 research articles about research-paper recommender systems and presented some descriptive statistics in the paper and a discussion about the major advancements and shortcomings and an overview of the most common recommendation concepts and approaches. [5]discussed about the three main recommendation approaches i.e. collaborative filtering, content based and hybrid recommender systems. It also throws light on the major limitations of these approaches thereby explaining the possible research areas to work on.[6] in their paper explained the utility of recommender systems in E-commerce sites. It analyses various sites using recommendation and discusses the various approaches used by them along with the possible opportunities for further work in e-commerce.[7] gives overview of recommender systems, the recommendation approaches and algorithms along with the evolution of recommender systems to the most recently developed ones. [8]reviewed the traditional and modern recommendation approaches along with the challenges faced. [9]reviewed various recommender systems developed till date in numerous domains. They discussed various aspects concerning recommender systems namely: recommender system software, real-world application domains and application platforms. Hence we can see that some of the papers are based on recommendation approaches, a particular domain of their utility or both. We on the contrary have not only discussed the recommendation approaches but also further categories of the approaches, various algorithms used in recommender systems along with some domain areas where recommender systems have been developed, their functionality and shortcomings.

This paper is structured in 6 sections. In section 2, we have described the recommendation approaches. Section 3 describes the sub-categories of approaches discussed in section 2. In section 4, we have discussed in brief the various recommendation algorithms. Section 5 includes the application areas of recommender systems in various domains. Then in section 6, we have discussed various parameters that are considered for evaluating the performance of recommender systems and in section 7 we have discussed the limitations of recommender systems. Finally we have concluded our work in section 8.

2. Recommendation Approaches

To get a better familiarisation with the concept of Recommender systems, this section first describes the most common recommendation approaches i.e. collaborative filtering, content based filtering, hybrid recommender systems and knowledge based systems, and some other recommendation approaches like demographic based systems and community based systems. Also, the subcategories of the most common approaches are discussed in brief.

2.1 Collaborative Filtering

Collaborative filtering is often referred to as people-to-people correlation[10]. The basic concept of collaborative filtering is that two or more individuals sharing some similar interests in one area tend toget inclined towards similar items or products from some other area too. The similarity between the users can be figured out on their browsing behaviour (click-through rate), browsing pattern and ratings (explicit, implicit). The concept of collaborative filtering can be clearly understood with the help of a simple example: consider Facebook. You can always see people you may know option on your home page with multiple number of people in the list. So the basic criteria behind making those suggestions are based on the concept of recommender systems only. The suggestions are filtered out on multiple parameters like: number of mutual friends you have with that individual, number of similar pages you both have liked or groups in common and also common places you have been to or you belong to. The approach used here is Collaborative filtering i.e. if a person x and you have a number of friends in common than the chances are that you two may also know each other. Hence it is called people to people co-relation.

2.2 Content-based Filtering

Content based systems are based on the concept "Show me more of what I have liked"[10]. The basic idea behind such systems is to recommend items or products to a particular user, which are similar to the ones that user has already liked in the past. The similarity between two or more items can be calculated based on their similar features. To understand better, continuing with the same example discussed above, whenever you watch any video on Facebook like the ones posted on Pet lover's page, after you finish watching it, you get links of similar videos on your home page. Also, when you like some page, you get suggestions of pages similar to them. So what actually is behind all this is Content based filtering i.e. if you like an xyz item from some particular category, it is likely that you might like any other item abc similar to it (from the same category or some category similar to it).

2.3 Knowledge-based Systems

Knowledge based systems are based on the concept "Tell me what fits my needs"[10]. Such systems make recommendations to the user based on a specific domain knowledge i.e. the system gets the user requirements, matches it with its knowledge base about that particular domain and recommends those items which according to it are the most appropriate and useful for the user, besides keeping user preferences in consideration. For better understanding, let's take up an example: consider online shopping sites. When you want to buy something, say a laptop, you will be asked to provide you requirements. Once you do that, you are recommended the most suitable product, which not only meets your requirements but is also the product that the system considers, would be the most suitable one for you.

2.4 Hybrid Recommender Systems

Hybrid systems as the name suggests are the combinational systems, the idea behind which are to combine the features of two systems (recommendation approaches) in such a way that the shortcomings of one are overcome by the other. Or you can say it gives the best of both worlds. For example consider Netflix, it is a hybrid of collaborative filtering and content-based filtering i.e. it recommends movies to the user on the basis of both his likes and his similarity with other users. Suppose a user likes The Notebook, Fated to Love you, PS I Love you etc. then the next time he visits the site, he would be recommended movies from Romantic genre (content-based). Also, if a user x and a user y has plenty of movies they have liked in common, then each of them would be recommended the next movie either of them likes (collaborative-filtering).

2.5 Demographic Systems

Demographic systems as the name says clearly are based on the demography of the user or on the region the user belongs to. The basic idea is that the recommendations made are based on the demographic region of the user. For example, consider eBay (online shopping site). Here the user has to select the region (country in this case like eBay. in, ebay.uk etc.) he or she belongs to for better use. As a result, only the products available in that particular region selected would be recommended to the user and the cost too would be given as per the relevant currency. And hence only relevant recommendations would be made.

2.6 Community-based Systems

Community-based systems follow, "Tell me who your friends are, and I will tell you who you are[2]". Such systems make recommendations to the user based on the preferences of the friends of the user i.e. unlike the case with collaborative filtering where recommendations are based upon the similarity of users (random or friends), here the recommendations are purely based on the similarity with the user's friends. Consider Facebook for example. It is quite likely for you to add someone in your profile, suggested by your friend, while the chances of you adding the one shown in the people you may know are comparatively less. And that is what community-based systems are all about, which clearly is based on the fact that we rely more on the recommendations taking our friends into consideration rather than the ones where random individuals are considered.

Table 1 provides the comparison of the pros and cons of above discussed approaches.

3. Further Classification of Recommendation Approaches

Figure 1 shows the classification of Recommendation approaches.

Table 1. Comparison of the pros and cons of above discussed approaches

TYPE	PROS	CONS
CF	Provides better results when the no. of users and the no. of ratings are available in abundance. No knowledge engineering required.	Cold-start problem both in case of new users and new items. Sparsity problems
CBF	It does not rely on other users' preferences. Comparison between items possible.	Users with thousands of purchases or varying tastes, make it a little difficult to make the valid recommendations. Cold start for new users. No surprises.
KBS	Deterministic recommendations Assured quality No cold start	Cost of knowledge acquisition from domain experts, users etc. are too high. Basically static
HS	Avoids some of the shortcomings.	Most datasets do not allow to compare different recommendation paradigmz.
DS	Recommendations valid to only a particular region are made, thus avoiding invalid ones.	Relatively less research in this particular technique.
CBS	It enables simple & comprehensive data acquisition related to users' social relations.	Recommendations are not always accurate.

Here,

CF- Collaborative Filtering
CBF- Content-Based Filtering
KBS- Knowledge Based Systems
HS- Hybrid Systems
DS- Demographic Systems
CBS- Community Based systems

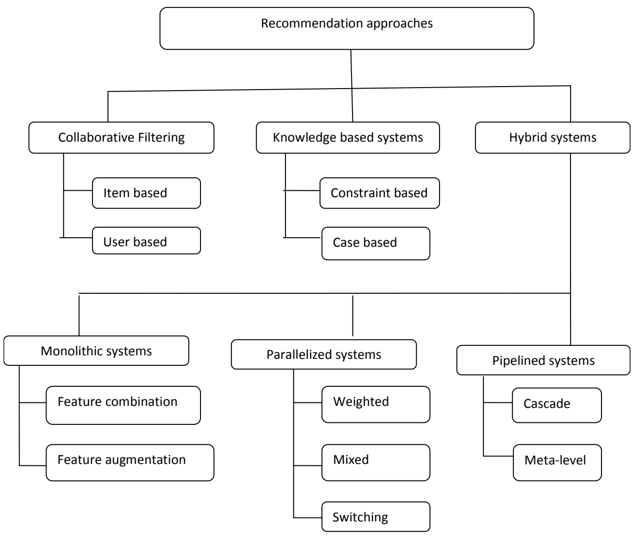

Figure 1. Classification of Recommendation approaches.

3.1 User based Collaborative Filtering

User based collaborative filtering as the name suggests is based on user-user co-relation. It is the first ever collaborative filtering method which was introduced first in the Group Lens Use net article recommender system. In this approach, for a user x (active user) whose previous ratings and user id is available, a group of similar users with similar behaviour in terms of past ratings is found out and their ratings are used to predict what might interest the active user[11].

$$s : U \times U \rightarrow R$$

3.2 Item based Collaborative Filtering

It being item-item co-relation, is one of the most widely developed collaborative filtering technique. It does not take the user similarity into consideration; rather it considers the similarity between the rating patterns of the items. For two items a and b having similar users who like and dislike them, they are considered similar and the users thus are anticipated to have similar interests for similar items. It is similar to content based filtering in terms of overall structure, except the fact that here the similarity between the items is found based on the user ratings pattern and not on the item.

$$s : I \times I \rightarrow R$$

3.3 Constraint based Knowledge based Systems

In constraint based systems, the user specifies his requirements, based on which he is recommended a product or an item with an explanation as to why that particular recommendation is made. The recommendation is governed by some explicitly defined set of recommendation rules that relate customer requirements to item features. It has a knowledge base that further includes: two sets of variables ($V = V_C \cup V_{PROD}$), one that describes potential customer requirements and the other one describes product properties and three different sets of constraints ($C = C_R \cup C_F \cup C_{PROD}$) that define what items to be recommended to a customer and in what situation[2].

Here, V_C – customer requirements
V_{PROD} – product properties
C_R - constraints
C_F – filter conditions
C_{PROD} – products

3.4 Case based Knowledge based Systems

Case based systems rely on different similarity measures to fetch items which are similar to those of the requirements specified by the customer and to what extent[11]. The similarity between the customer requirements and the properties of an item can be calculated by:

$$\text{Similarity}\,(p,REQ) = \frac{\sum r \lessdot REQwr * sim(p,r)}{\sum r \lessdot REQwr}$$

Here, w_r – importance weight for requirement

3.5 Monolithic Hybrid Systems

Monolithic systems as the name says, consists of only a single recommender component that combines numerous approaches by pre-processing and integrating various knowledge sources. Thus a built-in modification of the algorithm for handling and pre-processing the input data results in a hybrid system. It is further classified as: Feature combination and Feature augmentation hybrid systems.

A Feature combination hybrid system combines multiple input data to a single recommender algorithm. While a Feature augmentation hybrid system is the one where output of one algorithm is combined with that of the input for the other.

3.6 Parallelized Hybrid Systems

Parallelized systems consist of multiple recommender systems employed simultaneously and a specific hybridization mechanism that combines their outputs. They are further categorised as: Weighted, Mixed and Switching hybrid systems.

Weighted hybrid systems make the recommendations by integrating the weighted sums of the scores of two or more recommendation systems into a single recommendation list. Mixed hybrid systems are similar to that of weighted systems, where the results of various recommender systems are combined together and presented to the user. Switching hybrid systems, based on the quality of recommendation results and user requirements, switches between recommendation algorithms to provide the user with the best results.

3.7 Pipelined Hybrid Systems

Pipelined systems make the recommendations by following a staged process where multiple recommendation techniques are built on each other. They are further classified as: Cascade systems and Meta-level systems.

Cascade hybrid systems use the concept of successor and predecessor where the output of the predecessor on being refined by the successor results into the final recommendations. Meta-level hybrid systems build a model that is based on one algorithm and further used as an input for another algorithm.

4. Recommendation Algorithms

4.1 For Computing User Similarity

4.1.1 Pearson Co-relation

This algorithm is used to find the similarity between two users based on their ratings. The values lie between -1 (strong negative co-relation) to +1 (strong positive co-relation). It is given by:

$$\mathrm{sim}(a,b)=\frac{\sum_{p\in P}\left(\mathrm{r}_{a,p}-\bar{\mathrm{r}}_a\right)\left(\mathrm{r}_{b,p}-\bar{\mathrm{r}}_b\right)}{\sqrt{\sum_{p\in P}\left(\mathrm{r}_{a,p}-\bar{\mathrm{r}}_a\right)^2}\sqrt{\sum_{p\in P}\left(\mathrm{r}_{b,p}-\bar{\mathrm{r}}_b\right)^2}}$$

Here,

a,b - users

$r_{a,p}$ - rating of user a for product p

$\bar{r}_a$, $\bar{r}_b$ – average user rating

P – set of products rated by both a and b

4.1.2 Constrained Pearson Co-relation

This algorithm is used when the co-relation coefficient between two users increases as a result of either positive or negative ratings from the user. If the rating score is below the midpoint of the scaling scheme, it is considered as negative rating while a score above the midpoint is considered as a positive rating. It is given by:

$$\mathrm{sim}(a,b)=\frac{\sum\left(p_{ai}-mp\right)\left(p_{bi}-mp\right)}{\sqrt{\sum\left(p_{ai}-mp\right)^2}\sqrt{\sum\left(p_{bi}-mp\right)^2}}$$

Here,

mp – midpoint of rating scale

4.1.3 Spearman Co-relation

It is a non-parametric method that computes a measure of co-relation between ranks. It is similar to that of Pearson co-relation except that it uses ranks instead of ratings. It is given by:

$$\mathrm{sim}(a,b)=\frac{\sum\left(rank_{ai}-\overline{rank_a}\right)\left(rank_{bi}-\overline{rank_b}\right)}{\sqrt{\sum\left(rank_{ai}-\overline{rank_a}\right)^2}\sqrt{\sum\left(rank_{bi}-\overline{rank_b}\right)^2}}$$

4.1.4 Cosine Similarity

It is a vector based approach based on linear algebra. Users are considered as vectors and similarity between them is measured by the cosine distance or dot product of the angles between the rating vectors. It is given by:

$$\mathrm{sim}(a,b)=\frac{r_a.r_b}{\|r_a\|_2\|r_b\|_2}$$

4.2 For computing Item similarity

4.2.1 Cosine Similarity

It is the most popularly used metric for calculating similarity between items. It is the dot product or cosine product of the angles between the rating vectors, where the vectors are the items itself. It is given by:

$$\mathrm{sim}(i,j)=\frac{r_i.r_j}{\|r_i\|_2\|r_j\|_2}$$

4.2.2 Conditional Probability

It is generally used in online shopping sites where the purchase histories of the customer are available. It is given as:

$$\mathrm{sim}(i,j)=\mathrm{P}_{\,\mathrm{rB}}\left[\mathrm{j}\in\mathrm{B}\,|\in\mathrm{B}\right]$$

Here,

B- User's purchase history

4.3 For Neighbourhood Generation

4.3.1 K Nearest Neighbour Approach

K nearest neighbour approach is a simple algorithm used as a non-parametric technique to calculate the similarity measure or distance between the neighbours (users, items or products). The neighbour with the smallest distance measure will be the k-nearest neighbour and the recommendations would be made based on that. Nearest neighbour can be found out by calculating the distance between all the neighbours by using:

- Euclidean distance

$$\sqrt{\sum_{i=1}^{k}\left(x_i - y_i\right)^2}$$

- Manhattan distance

$$\sum_{i=1}^{k}\left|x_i - y_i\right|$$

4.3.2 Association Rule Mining

It is a common data-mining technique used to identify rule-like relationship patterns, co-relations and associations from the datasets. It is generally a combination of if/then statements that helps in finding relationships between varying products or users. Simply put, it finds the rules which support the relationship of a particular item with others, to make future predictions. For example, a rule can be "If a user buys bread, then he or she will buy butter or jam also in most of the cases". The measures of association rule are: support and confidence.

5. Application Areas

In this section, we will review and discuss the various developments and applications of Recommender systems in various application areas: Entertainment, Social-networking based, Content-based, Services-based and E-commerce. For each category, we have presented some examples.

5.1 Entertainment

In the field of entertainment, recommender systems have not only gained huge popularity but they have also proven to be a huge success. From music and movies to IPTV, recommender systems have left an everlasting impact. MovieLens.com, ITunes, Netflix, Pandora, TiVo are a few examples of recommender systems in entertainment. Such systems usually deploy collaborative filtering and content-based filtering. For example, in case of MovieLens.com, when we like and rate a movie from some specific genre, we get recommendations of other movies from the same genre. Also, two users having similar past preferences are considered to possible have similar interests in future also.

[12]in their work have proposed a smart and social TV program recommender system that combines Internet and web 2.0 features with television set-top boxes. The proposed work also includes explanations of the recommendations so as to gain the user trust and to enhance the user-satisfaction. [13]introduced RecomMetz, a context-aware mobile recommender system that recommends movies, movie theatre and the show-time to the users. This system is also effective in case of cold-start problem.[14] developed GroupRem, a movies recommender system for members of a group. Instead of using user ratings for making the recommendations, this system considers the tags for capturing the movie content and users' interests.[15] presented a hybrid music recommender system that uses both content and tags for making the recommendations. This system proved to be quite effective as it overcame the data sparsity problem.

5.2 Social-networking based

Social networking sites are too popular now-a-days. Such sites not only keep people connected, but also tell the users what their friends, family members and colleagues are up to. Facebook, Twitter, LinkedIn etc. are the leading examples. Recommendations in such sites can be seen in the form of People you may know, Pages you may like, Suggestions for you etc. Recommendations here are made using click-through rate, browsing behaviour and using user tags.

[16]in their work proposed a neighbourhood based recommender system that makes recommendations to the Twitter users in the form of URLs. For this, hash tags are used as the representatives of URL relevancy and the similarity between two hash tags is calculated using Euclidean, Cosine, Jaccard and Dice coefficient. [17]developed an EmotionChat- a web chartroom having emotion regulation. It is an interactive system that keeps a track

of varying emotional states of the user and makes the recommendations accordingly. The system recommends music, cartoons to e-learners on detecting negative emotional states.

[18]proposed a scheme to spot leaders in a social network site taking trust into consideration. The effects of removing the leader, were also studied. [19]studied the effects of tag forgery in real-world and the relationship between the degree of privacy and the degrading quality of recommendation. The work showed that the greater the extent of user privacy, the lesser would be the utility of the recommendations made.

5.3 Content-based

Content based systems can be personalized newspapers, recommendation for documents, recommendations of Web pages, e-learning applications or e-mail filters. E-learning has gained its popularity in educational institutes for over a decade now. Such systems aim to provide learning material to the users based on their preferences and their learning activities. E-libraries or digital libraries are the sources of e-learning where the user can find numerous knowledge sources and information. A lot of work has been done in this domain too.

[20]in their work have proposed a content-based recommender system for scientific libraries. The system deploys a keyword extraction system so as to generate relevant metadata for the documents using the Recommender and Explanation system. [21]introduced an effective hybrid approach for book recommendation, using the concept of Ontology for user profiling, thereby making the system more efficient. The system deploys Slope one Algorithm, LSH Algorithm and hybrid approach to make the recommendations. Also using demographic attributes of the user, cold-start problem was overcome. [22]for improving the quality of recommendations, proposed a hybrid recommendation approach (content based and collaborative filtering) , based on modeling of e-learning materials in multi-dimensional space taking the attributes of the materials into consideration.

5.4 Services-based

Services based systems can be recommendations of travel services, recommendation of experts forconsultation, recommendation of houses to rent or matchmaking services. E-tourism for example, provides suggestions to the tourists about places worth paying a visit when travelling to a particular destination. For more convenience and ease, such systems are now developed as mobile apps and can be easily accessed by the user.

The Talking Museum Project of MaschioAngioino castle, developed by [23]exploits the concept of IoT, to enable the objects of the museum talk about their history using multimedia facilities. The stories of the museum objects are sent to the user mobile devices during his visit, based upon his interests and further visit is planned and recommended as per the user profile captured. [24]developed a personalized recommender system for cosmetic business, which combines content-based, collaborative filtering with data-mining techniques. Customer preferences are determined using a scoring approach. Clustering algorithms and Association rule algorithms are also used.

[25]introduced a tourist guide recommender system that provides recommendations using the Mahout library. The system is collaborated with social networks and is available both as mobile and web application that provides consultation, publication and recommendation of tourists destinations[26]. introduced a travel package recommendation system that made the recommendations based on the social network profile of the user.

5.5 E-commerce

In past few decades, a number of recommender systems have been developed in the field of e-commerce to provide online assistance to consumers browsing online shopping sites. User feedback and ratings form the common criteria for making the recommendations. For example, in the Google play store, users give ratings to the applications downloaded by them and these ratings are used for making recommendations to other users. We have numerous of online shopping sites developed till date, like Amazon, EBay, Myntra, Zovi, and Shopclues and so on. On such sites, we can always see recommendations in the form of: people also purchased, most viewed, similar products, top-selling etc.

[27]in their work introduced e-Zoco, a prototype of e-commerce portal that provides catalogue services and product selection services. A rule-based knowledge learning service is included to explain the existing relationships among the product attributes representing a specific product category. [28]introduced personalization recommender system, its uses and the various technologies used. [29]presented a framework of e-commerce recommendation system which is flexible, and provides dynamic and flexible recommendations to the user.

[30]presented an explanation of the relationship of recommender systems and database analysis techniques. The author discussed how the sales of e-commerce sites are increased using recommender systems and the concerned privacy issues. For case study, the author discussed the most commonly used e-commerce sites like Amazon, CDNOW, eBay, MovieFinder.com, Reel.com etc.[31]studied the human resource issues in e-business in China. The study showed that with the rapid growth in e-commerce, the demand for human resource has faced major downfall.

Table 2 provides examples of Recommender systems developed in various application domains.

6. Metrics for Evaluating Recommender Systems

In this section, we will discuss the various parameters on the basis of which, we can measure the performance of a recommender system.

6.1 Accuracy

To find out how accurate a given recommender system is, we need to calculate Root Mean Square Error (RMSE) and Mean Absolute Error (MAE). RMSE and MAE give a measure of how closely the predicted ratings are related to the actual ratings.

$$\text{RMSE}=\sqrt{\frac{1}{n}\sum_{u,i}\left(p_{u,i}-r_{u,i}\right)^2}$$

$$\text{MAE}=\frac{1}{n}\sum_{u,i}\left|p_{u,i}-r_{u,i}\right|$$

Here,

n – no. of ratings

$p_{u,i}$ – predicted user ratings for item i

$r_{u,i}$ – actual ratings

Lower the value of these two, higher is the accuracy.

Table 2. Examples of Recommender systems developed in various application domains

Application Area	Recommender system	Approach
Social Network	Facebook	Collaborative filtering Content based
	Twitter	Collaborative filtering
	LinkedIn	Collaborative filtering
	MySpace	Personalized recommendation
Movies	Netflix	Hybrid recommender system
	MovieLens	Collaborative filtering
	Reel.com	Collaborative filtering
Music	Pandora	Content-based recommendation
	Last.fm	Collaborative filtering
	MyStrands	Behaviour based recommendation
E-commerce	Amazon	Item to item collaborative filtering, Content-based recommendation
	eBay	Collaborative filtering Demographic recommendation
E-library	CiteSeer	Content-based Collaborative filtering
	Docear	Content-based
	Techlens	Content-based Collaborative filtering
Videos	Youtube	Personalized recommendation
	Hulu	Content-based recommendation Collaborative filtering
Entertainment	Jester	Collaborative filtering
	TiVo	Collaborative filtering
	TasteKid	Content-based recommendation

6.2 Coverage

Coverage for a recommender system must always be high, i.e. higher value of coverage indicates that the maximum portion of items is being recommended[32]. It can be item space coverage or user space coverage.

$$\text{Coverage}=\frac{N_d}{N}$$

Here,

N_d - no. of distinct objects

N – Total no. of objects

6.3 Precision

Precision is a measure of goodness that determines the proportion of recommended items from the available ones. It is given by:

$$\text{Precision}=\frac{t_p}{t_p+f_p}$$

Here,

t_p –true positive

f_p -false positive

6.4 Recall

Recall is a measure of completeness that gives the proportion of relevant recommendations made out of all the relevant ones.

$$\text{Recall}=\frac{t_p}{t_p+f_n}$$

Here,

f_n - False negative

7. Limitations of Recommender Systems

Although a lot of work has been done in Recommender systems, but still there are some limitations that need more attention of the researchers. Most common of those limitations are discussed below in brief:

7.1 Cold-start Problem

Cold start problem basically is a situation where it is quite difficult to make any correct recommendations to the user as either the user is new to the system or a given item or product is added to the system for the first time. So we can say that there are two types of cold-start problems: new user cold start problem and new item cold start problems. New user cold start problem is where a new user starts using the system and very less information about him is available. As a result, making recommendations to him becomes quite difficult. Such s situation occurs in case of content based filtering. While in new item cold start problem, a new item is added to the system and no user ratings or reviews are available, as a result again difficulty in making recommendations. This problem normally arises in case of Collaborative filtering.

7.2 Sparsity

Sparsity is derived from the word “Sparse” which means “scattered”. In terms of recommender systems, sparsity implies to the irregular, insufficient or highly varying user ratings. It is one of the major problems encountered by recommender systems. The major reason behind sparsity is that most of the users do not provide ratings and the ones available are usually too scattered or sparse. Consider rating any mobile app for example. We often ignore rating the app when asked and if we do then some of us either rate it positively or some of us are unsatisfied and rate it negatively, as a result of which the ratings available are quite scattered. This is a major issue for Collaborative filtering where the recommendations are based on the user ratings.

7.3 Scalability

Scalability means extensibility i.e. how well a system works when the volume of data increases. Recommender systems work quite well in case of small dataset, but to deal with real-world dataset that grows extensively, is quite a cumbersome task. Though there are algorithms that deal with massive and dynamic data-sets, but the results are not always accurate.

7.4 Privacy Protection

Privacy in terms of Recommender systems has turned out to be an issue as recommender systems take as much information as possible from the users to provide better recommendations. But it often hinders the user privacy in the way that sometimes the system might know too much about the user then required. Also the user data can be used by some malicious users as it is easily accessible. An example can be stated here where a father got to know about the pregnancy of his teenage daughter through the use of targeted ads, in fact the company accurate the due date of his daughter based on the products she purchased[34].

7.5 Over-specialisation

Over-specialisation becomes a problem when the user gets similar kinds of recommendations only based on his past behaviour and there is no surprise element (content-based systems). As a result, there is no diversity in the recommendations pattern and the chances of the user discovering something new, that might actually interest him and prove to be beneficial, are nearly negligible.

7.6 GraySheep Problem

Gray sheep problem refers to the one where a user shows quite inconsistent behaviour. These users have no defined preferences and can like one thing at one moment while liking the exact contrary at the other. For example, we might like both romantic and action movies, and a user having similar likes of action movies to those similar to us, would also be recommended romantic movies, which in no way interest him and hence are irrelevant recommendations. As a result, they decrease the efficiency of recommender systems.

7.7 Shilling Attacks

There are mainly two types of attacks: push attacks and nuke attacks, with the objective of promoting or demoting the product rating predictions respectively[35]. For this purpose, the attacker creates a fake profile and give fake positive ratings to increase the popularity of the products they are biased to in case of push attacks and fake negative ratings to the products of their competitors to decrease their popularity in case of nuke attacks. Depending upon the way in which the attacker rates a product, the attacks can be probe attacks, random attacks, average attacks, popular attack, probe attack, segment attack, love/hate attack, sampling attack and perfect knowledge attack[36].

8. Conclusion

Recommender systems were developed with the rapid increase of choices in front of the masses due to the development in each and every area and when the need to get the most suitable or best out of the rest products was realised. Though a relatively newer concept in terms of research, we can see that a lot has been achieved in this area. There are various recommendation approaches, the main being collaborative filtering, content based systems, knowledge based and hybrid systems, which are further classified into plenty of more approaches. For each approach, there are multiple algorithms, to find user similarity, to form cluster of users, to find neighbourhood and so on. From e-commerce to e-services, there are many areas where recommender systems have proven their worth. But still there are challenges that need to be overcome. So the need of the hour is that the researchers to focus on these challenges and thus make the recommendations more reliable.

9. References

1. Huttner J. From Tapestry to SVD: A survey of the algorithms that power recommender systems;2009.
2. Ricci F, Rokach L, Shapira B. Introduction to recommender systems handbook. Springer US; 2011. p.1–35.
3. Lu J, Wu D, Mao M, Wang W, Zhang G. Recommender system application developments: a survey. Decision Support Systems. 2015 Jun 30; 74:12–32.
4. Beel J, Gipp B, Langer S, Breitinger C. Research-paper recommender systems: a literature survey. International Journal on Digital Libraries. 2015:1–34.
5. Sharma L, Gera A. A survey of recommendation system: research challenges. International Journal of Engineering Trends and Technology. 2013 May; 4(5):1989–92.
6. Schafer JB, Konstan J, Riedl J. Recommender systems in e-commerce. Proceedings of the 1st ACM conference on Electronic commerce; 1999 Nov 1. p. 158–66.
7. Bobadilla J, Ortega F, Hernando A, Gutiérrez A. Recommender systems survey. Knowledge-based systems. 2013 Jul 31;46:109–32.
8. Asanov D. Algorithms and methods in recommender systems. Berlin Institute of Technology, Berlin: Germany; 2011. p.1–7.
9. Jannach D, Friedrich G. Tutorial: recommender systems. Proceedings of the International Joint Conference on Artificial Intelligence, Barcelona; 2011 Jul 17. p.1–26.
10. Ekstrand MD, Riedl JT, Konstan JA. Collaborative filtering recommender systems. Foundations and Trends in Human-Computer Interaction. 2011 Feb 1;4(2):81–173.
11. Jannach D, Zanker M, Felfernig A, Friedrich G. Recommender systems: an introduction. Cambridge University Press; 2010 Sep 30.
12. Chang N, Irvan M, Terano T. A TV program recommender framework. Procedia Computer Science. 2013 Dec 31;22:561–70.

13. Mendoza LOC, García RV, González AR, Hernández GA, Zapater JJS. RecomMetz: A context-aware knowledge-based mobile recommender system for movie showtimes. Expert Systems with Applications. 2015 Feb 15;42(3):1202–22.
14. Pera MS, Ng YK. A group recommender for movies based on content similarity and popularity. Information Processing & Management. 2013 May 31;49(3):673–87.
15. Horsburgh B, Craw S, Massie S. Learning pseudo-tags to augment sparse tagging in hybrid music recommender systems. Artificial Intelligence. 2015 Feb 28;219:25–39.
16. Yazdanfar N, Thomo A. Link recommender: collaborative-filtering for recommending URLS to Twitter users. Procedia Computer Science. 2013 Dec 31;19:412–9.
17. Zheng D, Tian F, Liu J, Zheng Q, Qin J. Emotion chat: a web chatroom with emotion regulation for e-learners. Physics Procedia. 2012 Dec 31;25:763–70.
18. Dutta P, Kumaravel A. A novel approach to trust based identification of leaders in social networks. Indian Journal of Science and Technology. 2016;9(10):1–9.DOI: 10.17485/ijst/2016/v9i10/85317.
19. Puglisi S, Parra-Arnau J, Forné J, Rebollo-Monedero D. On content-based recommendation and user privacy in social-tagging systems. Computer Standards & Interfaces. 2015 Sep 30;41:17–27.
20. De Nart D, Tasso C. A personalized concept-driven recommender system for scientific libraries. Procedia Computer Science. 2014 Dec 31;38:84–91.
21. Chandak M, Girase S, Mukhopadhyay D. Introducing hybrid technique for optimization of book recommender system. Procedia Computer Science. 2015 Dec 31;45:23–31.
22. Salehi M, Pourzaferani M, Razavi SA. Hybrid attribute-based recommender system for learning material using genetic algorithm and a multidimensional information model. Egyptian Informatics Journal. 2013 Mar 31;14(1):67–78.
23. Amato F, Chianese A, Mazzeo A, Moscato V, Picariello A, Piccialli F. The talking museum project. Procedia Computer Science. 2013 Dec 31; 21:114–21.
24. Wang YF, Chuang YL, Hsu MH, Keh HC. A personalized recommender system for the cosmetic business. Expert Systems with Applications. 2004 Apr 30;26(3):427–34.
25. Umanets A, Ferreira A, Leite N. GuideMe–a tourist guide with a recommender system and social interaction. Procedia Technology. 2014 Dec 31;17:407–14.
26. Reddy CA, Subramaniyaswamy V. An enhanced travel package recommendation system based on location dependent social data. Indian Journal of Science and Technology. 2015;8(16):1–7.DOI: 10.17485/ijst/2015/v8i16/63571.
27. Castro-Schez JJ, Miguel R, Vallejo D, López-López LM. A highly adaptive recommender system based on fuzzy logic for B2C e-commerce portals. Expert Systems with Applications. 2011 Mar 31;38(3):2441–54.
28. Ya L. The comparison of personalization recommendation for e-commerce. Physics Procedia. 2012 Dec 31;25:475–8.
29. Ronca D, Calvanese A, Birtolo C. A flexible framework for context-aware recommendations in the Social Commerce domain. Proceedings of the Joint EDBT/ICDT 2013 Workshops; 2013 Mar 18. p. 105–10.
30. Schafer JB, Konstan JA, Riedl J. E-commerce recommendation applications. Applications of Data Mining to Electronic Commerce,Springer, US; 2001. p. 115–53.
31. Lang T, Gupta B. The human resource issues and their impact on firm growth in small e-businesses in China. Indian Journal of Science and Technology. 2015;8(S4):135–44.
32. Shani G, Gunawardana A. Evaluating recommendation systems. Recommender systems handbook, Springer, US;2011. p. 257–97.
33. Sharma M, Mann S. A survey of recommender systems: approaches and limitations. International Journal of Engineering and Innovative Technology. 2013:1–9.
34. Jones T. Recommender systems, Part 1: introduction to approaches and algorithms. IBM DeveloperWorks. 2013 Dec12.
35. O'Mahony MP, Hurley NJ, Silvestre GC. Recommender systems: Attack types and strategies. Association for the Advancement of Artificial Intelligence. 2005 Jul 9:334–39.
36. Kumar KK, Geethakumari G. Analysis of semantic attacks in online social networks. Recent Trends in Computer Networks and Distributed Systems Security, Springer Berlin Heidelberg; 2014 Mar 13. p. 45–56.