

Recommender Systems: From Achievements to Requirements

¹Richa Sharma, ²Sharu Vinayak, ³Rahul Singh

¹Student, ²Student, ³Assistant Professor

¹CSE Department,

¹Chandigarh University, Mohali, India

Abstract - Recommender Systems are often referred to as software tools that help making the selection process easier and time saving. The aim of Recommender Systems is to provide the user with the most suitable recommendation, from the plethora of options available. Till date, a number of recommender systems have been developed for various application areas. Although, a lot of work has been done in this particular research area, yet there are still some limitations that need attention from a researcher's point of view. In this paper, we present a brief overview of Recommender systems including their applications, limitations that still need to be worked on and we have also proposed some ideas that can be used to overcome some of those limitations.

Index Terms - Recommender systems, Shilling attacks, Click-through rate.

I. INTRODUCTION

Recommender systems can often be referred to as the software tools or techniques that help people selecting the most suitable product for them from the plethora of options available [1]. These systems use the basic concept of selection, prioritization and automation. Recommender systems are automated systems that prioritize each product on the basis of user ratings, hit ratios and user views and then select the most appropriate product for a user, based on his requirements and interests. The first Recommender system was developed in 1992 by Xerox Palo Alto Research centre. It was named Tapestry and it used Collaborative filtering approach [2]. But even before that, the concept of Recommendation was prevalent in our daily lives, yet unnoticed. It 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. That is because ants tend to leave markers for other ants which show them the way to food and act as a recommender. Then there were the ministers of kingdoms, who would recommend to their kings, what policies to make or what kingdoms to take over. The most common example of recommender systems that was prevalent in old-times and still does exist is aunts suggesting brides or grooms to the families while arranging marriages. But nothing was automated at that time. But then began the Industrial Revolution, bringing a lot of options in front of the masses to choose from. And then came the biggest invention of times, Computer, that made our lives way easier and there was Internet connecting people. And soon the need of Recommender Systems was realized.

Recommender systems proved to be a boon to the masses as soon as they were developed. The selection process got easier and saved more time. Till date, we have a number of recommender systems developed for various areas. For example: Amazon, LinkedIn, Hulu, Docear, Facebook, Pandora, Netflix, Jester, YouTube, Yahoo etc. Recommender Systems have evolved a lot over the past two decades and now also use the concept of Artificial intelligence, Information retrieval and Human-computer interaction [1], and hence have got more efficient and have gained more popularity.

TABLE 1. DIFFERENT RECOMMENDATION APPROACHES

S.no.	Approach	Description
1.	Collaborative Filtering	<ul style="list-style-type: none"> This is basically people-to-people co relation. It uses the wisdom of the crowd to recommend items to the user. The ratings can either be explicit or implicit. It assumes that people who had similar tastes in past, will also have similar tastes in future also.
2.	Content based Filtering	<ul style="list-style-type: none"> It is based on the concept "Show me more of what I have liked". It takes into account user preferences and on the basis of that, makes the recommendations. It recommends those items which are similar to the user preferences based on their past behavior.
3.	Knowledge based	<ul style="list-style-type: none"> It is based on the concept "Tell me what fits my needs". No user profile is maintained, rather user interaction is there. It uses knowledge of users & items.

4.	Hybrid Recommender Systems	<ul style="list-style-type: none"> • A hybrid system combining techniques A and B tries to use the advantages of A to fix the disadvantages of B.
----	----------------------------	--

Recommender Systems were considered important from the research point of view not more than two decades back, yet a lot has been achieved in this particular field. Jie Lu et al. [3] in their work have given detailed information of numerous Recommender Systems that have been developed so far for various application areas along with the various Recommendation approaches. J. Bobadilla et al. [4] in their work discussed briefly about the concept of recommender systems, recommendation approaches and algorithms developed so far with. They also discussed how recommender systems evolved from the very beginning to the most recently developed ones. Lalita Sharma et al. [5] in their work discussed about the most common recommendation approaches of collaborative filtering, content based filtering and hybrid recommender systems. They also gave an insight into the major limitations of these approaches and gave some idea about the possible research areas to work on. Danial Asanov et al. [6] reviewed the traditional and modern recommendation approaches. They also discussed the various challenges faced by Recommender systems. Gediminas Adomavicius et al. [7] gave an overview of Recommender systems and the state-of-the-art. The authors also discussed the most common recommendation approaches, their limitations and what can possibly be done to overcome those limitations.

This paper is organized as follows. In Section II, we have discussed some of the application areas of Recommender Systems. Section III covers the various limitations of Recommender Systems. In Section IV, we have suggested some ways to overcome the limitations of Recommender systems. And finally we have concluded the paper in Section V.

II. APPLICATION AREAS OF RECOMMENDER SYSTEMS

Till date, a number of Recommender Systems have been developed for various application areas. From e-commerce to e-library, we have plenty of Recommender Systems. Some of the most popular Recommender Systems are discussed below.

i. E-commerce

Any form of business or transaction that is carried out over the internet is what we refer to as e-commerce. Till date, a number of e-commerce recommender systems have been developed to provide online assistance to users browsing online shopping sites or carrying out any kind of transaction. User ratings and feedback are generally used for making the recommendations. For example, in the Google play store, users are often asked to give ratings to the applications downloaded by them so that these ratings can be used for making recommendations to other users. There are plenty of online shopping sites developed till date, like Amazon, EBay, Myntra, Zovi, and Shopclues and so on.

Amazon is one of the leading online shopping sites. From clothing to electronic products, one can find almost everything on Amazon. While using Amazon, we often come across features like: Most popular, people also viewed, recently viewed and featured at the bottom of the site page. These suggested products are nothing but the recommendations made on the basis of our own likes or on the basis of how our interests are similar to that of some other user. For each customer, personalization of his profile is done based on his preferences obtained by observing the click-through rate and his browsing behavior [8]. Amazon uses Item-item collaborative filtering and content-based filtering.

Similarly, e-Bay, also an online shopping site, uses collaborative and demography based recommendations. It has the feature of Feedback profile where the buyers and sellers give feedback of the users they have worked with. The feedback is in the form of a Likert scale having ratings in the form of satisfied, neutral and dissatisfied [9]. This feedback is then further used to make the recommendations.

ii. E-library

E-library has gained its recognition in educational institutes for over a decade now. Such systems provide learning material to the users based on their preferences and learning activities. E-libraries are considered the sources of e-learning where the user can find abundant knowledge sources and information. A great deal of work has been done in this area with Docear, CiteSeer and TechLens being the most common examples.

Docear is an academic literature collection that aims to search, organize and create research articles [10]. The concept of mind maps is used to manage user data and the recommendations are made using Content-based filtering approach. Similarly, TechLens uses a hybrid system of both collaborative and content-based filtering. It takes implicit feedback into consideration to find the co-relations among different users to make the recommendations [11].

iii. Entertainment

Recommender systems have not only gained huge popularity but they have also proven to be a huge success in the field of entertainment. From music and movies to IPTV, recommender systems have left an overwhelming impact. MovieLens.com, iTunes, Jester, Netflix, YouTube, Pandora, TiVo are a few examples. Such systems are usually based on collaborative and content-based filtering. For example, in case of MovieLens.com, when we like and rate a movie from one genre, we get recommendations of movies from the same genre. Similarly, two users who have shown similar interests in the past are considered to have similar preferences in future also.

YouTube is the world's most popular online video community. It uses personalized recommendations to suggest videos to the user based on their preferences and on the kind of videos they have browsed [12]. Pandora is an online radio station where for each user an individual station is built based on his musical interests [13]. Songs that are similar to the ones liked by the user are played. The user has the option of hitting either a "thumbs up" that implies he likes what he hears or a "thumbs down" means that he does not want to hear that song again.

iv. Social-networking sites

Social networking sites have gained popularity over time and replaced the concept of sending letters or even phone calls. Such sites not only aim to keep people connected, but also tell them what their friends, family members and colleagues are up to. Facebook, Twitter, LinkedIn etc. are the leading examples. People you may know, Pages you may like, Suggestions for you etc. are how the recommendations are made on such sites using click-through rate, browsing behavior and using user tags.

Twitter is the leading site for micro-blogging movies, news, entertainment etc. It allows users to post short messages and status updates that can be followed by other users. It uses Collaborative filtering to make the recommendations [14]. LinkedIn is the largest online professional social network and it uses item-to-item collaborative filtering to recommend jobs, companies or candidates for a particular job [15]. The recommendations can be of the form: "People who viewed this profile also viewed" or "Companies you may want to follow".

Wherever Times is specified, Times Roman or Times New Roman may be used. If neither is available on your word processor, please use the font closest in appearance to Times. Avoid using bit-mapped fonts. True Type 1 or Open Type fonts are required. Please embed all fonts, in particular symbol fonts, as well, for math, etc.

III. LIMITATIONS OF RECOMMENDER SYSTEMS

Despite the work done in the field of Recommender systems, there are still some drawbacks which need to be removed. The most common of such drawbacks are discussed briefly in this section.

i. Cold start problem

Cold start problem arises when it is quite tedious to make any recommendations to the user in case either the user is new to the system or a given item or product is added to the system for the first time. There are basically two types of cold-start problems: new user cold start problem and new item cold start problem. When a new user starts using the system and very less information about him is available, we have New user cold start problem and making recommendations to him becomes difficult. Such a situation can be seen in case of content based filtering. While on the contrary, when a new item is added to the system and no user ratings or reviews are available new item cold start problem occurs, resulting in difficulty in making recommendations. This problem arises in case of Collaborative filtering.

ii. Privacy protection

Recommender systems take as much information as possible from the users to provide better recommendations but it often leads to threats to the user privacy. The user is always apprehensive that his data might be used by some malicious users being easily accessible. An example of such a situation is: a father got to know about the pregnancy of his teenage daughter through the use of targeted ads and based on the products she purchased the company accurately predicted the due date of his daughter [16].

iii. Over-specialization

Over-specialization becomes an issue when the user only gets one kind of recommendations based on his past behavior and as a result, there is no surprise element left for the user (content-based systems). There is no variety in the recommendations pattern and the likelihood that the user might discover something new, that might truly interest him and prove to be beneficial, is nearly negligible.

iv. Gray sheep problem

Gray sheep problem arises when the user shows quite inconsistent behavior i.e. the user has no definite preferences and can like one thing at one moment while liking the exact dissimilar at the other. For example, a user might like both romantic and horror movies, and a user having similar likes of horror movies to those similar to him, would also be recommended romantic movies, which in no way would interest him and form irrelevant recommendations. Therefore, Gray sheep users decrease the efficiency of recommender systems.

v. Sparsity

Sparsity originated from the word "Sparse" which means "scattered". In case of recommender systems, sparsity refers to the irregular, insufficient or highly unstable user ratings. The major cause of sparsity is that most of the users do not offer their ratings and the ones accessible are usually too scattered or sparse.

vi. Shilling attacks

Shilling attacks are classified as: push attacks and nuke attacks, with the objective of promoting or demoting the product rating predictions respectively. In case of Push attacks, the attacker creates a fake profile and gives fake positive ratings to increase the reputation of the products he is biased to and in case of Nuke attacks; he gives fake negative ratings to the products of his competitors to drop off their reputation.

IV. SCOPE OF IMPROVEMENT

i. To improve the efficiency of Click-through rate.

Click-through rate is the ratio of users who click on a specific link to the number of total users who view a page, email, or advertisement. It aims to measure the success of an online advertising campaign for a particular website [17]. In case of Recommender Systems also, click through rate plays a major role in making the recommendations i.e. higher the click-through rate, higher would be the chances of recommending that product or item. Consider YouTube that provides features like search engine, front page highlight, and related videos recommendation [18]. The video with the highest views is recommended to the users i.e. the number of times a video is streamed or even clicked on, the number of views for that particular video increases. People who want to promote their videos often ask their peers to just click on the video link and close it as soon as the video starts. This way they need not view the whole video but still their views would be counted, and hence more number of views but somehow these are false ratings.

We can overcome this issue by setting a validation here i.e. what can be done is, the views of any video should only be considered and counted when the whole video is viewed and not just a part of it. This way we will get the exact statistics that whether the number of likes of the viewers is consistent with those of the number of views. As a result only the videos that were actually viewed by a user the most number of times would be recommended rather than the ones which were just clicked on but not viewed.

ii. To avoid Shilling attacks

Shilling attacks, whether it is nuke attacks or push attacks, have been a major drawback of recommender systems. Xiang-Liang Zhang et al. [19] in their work proposed how to reduce the impact of shilling attacks using social clustering. Zhihai Yang et al. [20] proposed a novel attacks detection approach to make the system defiant to the impact of grey shilling attacks by taking into account the rating deviation of an item so as to discriminate between the unadulterated and fake ratings. A lot of work similar to this has been done, each of the proposed approach aims to either trace the shilling attacks or to reduce the impact of such attacks, but none aims to remove the mere possibility of shilling attacks.

We propose a very simple yet potentially effective idea as to how we can utterly bring the risk of shilling attacks to an end. Each citizen of any country has a unique identification number, Aadhar Card in case of India. What can be done is if all the e-mail ids are validated using the unique identification number of the user and for each identification number, a user can have not more than two email-ids i.e. one for personal use and one for professional use, then no user can create any fake profile. Thus ultimately the possibility of shilling attacks or profile injection attacks can be brought to a permanent halt.

V. CONCLUSION

Recommender Systems were developed around two decades back with the purpose of making the selection process easier. These systems have gained popularity over the time and we now have one or more Recommender Systems for almost every application area possible. But still there are a few limitations of Recommender systems that cannot be overlooked and need to be resolved to make the recommendations more reliable. In this paper, we gave a brief overview of Recommender systems, discussed some of the application areas and the limitations that still need to be worked on. We have given some ideas as to how to overcome some of such limitations i.e. to improve the efficiency of click-through rate and to avoid the shilling attacks. We hope that these ideas can be applied into practical application so that we can overcome some of the limitations of Recommender Systems in order to improve the quality of the recommendations.

REFERENCES

- [1] Ricci, Francesco, Lior Rokach, and Bracha Shapira. *Introduction to recommender systems handbook*. Springer US, 2011.
- [2] Huttner, Joseph. "From Tapestry to SVD: A Survey of the Algorithms That Power Recommender Systems." (2009).
- [3] Liu, Jie, Dianshuang Wu, Mingsong Mao, Wei Wang, and Guangquan Zhang. "Recommender system application developments: a survey." *Decision Support Systems* 74 (2015): 12-32.
- [4] Bobadilla, Jesús, Fernando Ortega, Antonio Hernando, and Abraham Gutiérrez. "Recommender systems survey." *Knowledge-Based Systems* 46 (2013): 109-132.
- [5] Sharma, Meenakshi, and Sandeen Mann. "A survey of recommender systems: approaches and limitations." *Int J Innov Eng Technol. ICAECE-2013, ISSN (2013)* (2013): 2319-1058.
- [6] Asanov, Daniar. "Algorithms and methods in recommender systems." *Berlin Institute of Technology, Berlin, Germany* (2011).
- [7] Adomavicius, Gediminas, and Alexander Tuzhilin. "Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions." *Knowledge and Data Engineering, IEEE Transactions on* 17, no. 6 (2005): 734-749.
- [8] Linden, Greg, Brent Smith, and Jeremy York. "Amazon.com recommendations: Item-to-item collaborative filtering." *Internet Computing, IEEE* 7, no. 1 (2003): 76-80.
- [9] Schafer, J. Ben, Joseph Konstan, and John Riedl. "Recommender systems in e-commerce." In *Proceedings of the 1st ACM conference on Electronic commerce*, pp. 158-166. ACM, 1999.
- [10] Beel, Joeran, Stefan Langer, Marcel Genzmehr, and Andreas Nürnberger. "Introducing Docear's research paper recommender system." In *Proceedings of the 13th ACM/IEEE-CS joint conference on Digital libraries*, pp. 459-460. ACM, 2013.
- [11] Kanoo, Nishikant, Jilin Chen, John T. Butler, Garv C. Foutv, James A. Stemmer, John Riedl, and Joseph A. Konstan. "Techlens: a researcher's desktop." In *Proceedings of the 2007 ACM conference on Recommender systems*, pp. 183-184. ACM, 2007.

- [12] Davidson, James, Benjamin Liebald, Junning Liu, Palash Nandy, Taylor Van Vleet, Ullas Gargi, Suiyov Gupta et al. "The YouTube video recommendation system." In *Proceedings of the fourth ACM conference on Recommender systems*, pp. 293-296. ACM, 2010.
- [13] Howe, Michael. "Pandora's Music Recommender." *A Case Study, I* (2009): 1-6.
- [14] Pankong, Nichakorn, and Somchai Prakanchaen. "Combining algorithms for Recommendation system on Twitter." In *Advanced Materials Research*, vol. 403, pp. 3688-3692. Trans Tech Publications, 2011.
- [15] Wu, Lili, Sam Shah, Sean Choi, Mitul Tiwari, and Christian Posse. "The Browsemaps: Collaborative Filtering at LinkedIn." In *RSWeb@ RecSys*. 2014.
- [16] Jones T. Recommender systems, Part 1: Introduction to approaches and algorithms. IBM DeveloperWorks. 2013 Dec12
- [17] Wikiedia contributors. "Click-through rate." *Wikiedia. The Free Encyclopedia*, https://en.wikipedia.org/w/index.php?title=Click-through_rate&oldid=713662891 (accessed April 28, 2016).
- [18] Zhou, Renjie, Samamon Khemmarat, and Lixin Gao. "The impact of YouTube recommendation system on video views." In *Proceedings of the 10th ACM SIGCOMM conference on Internet measurement*, pp. 404-410. ACM, 2010.
- [19] Zhang, Xiang-Liang, Tak Man Desmond Lee, and Georgios Pitsilis. "Securing recommender systems against shilling attacks using social-based clustering." *Journal of Computer Science and Technology* 28, no. 4 (2013): 616-624.
- [20] Yang, Zhihai. "Defending Grev Attacks by Exploiting Wavelet Analysis in Collaborative Filtering Recommender Systems." *arXiv preprint arXiv:1506.05247* (2015).

