An Optimal Enhancement of the Dynamic Features of Recommender Systems

R. Lavanya, Rithika Lahari, Palak Gupta

Abstract: Recommendation systems come under the domain of Data mining and Big Data analytics. It is useful tool that is used to predict the ratings or preferences of a user from a pool of resources. The preferences of user are dynamic in nature. The immeasurable usage of internet is having a great impact on the way we deal our lives and communicate with each other. As a result, the requirements of user browsing the internet are changing radically. Recommender Systems (RSs) provide a technology that helps users in finding relevant or preferential information among the pool of information using internet. This paper puts forward not only the issues related to the dynamic nature of user' requirements but also the changes in the systems' contents. The Recommendation Systems which involves the above stated issues are termed as Dynamic Recommender Systems (DRSs). This paper first defines the concept of DRS and then explores the various parameters that is taken into account in developing a DRS. This paper also discusses the scope of contributions in this field and concludes citing in possible extensions that can improve the dynamic qualities of recommendation systems in future.

Index Terms: Big Data Analytics, Dynamic users, Collaborative filtering, Recommendation system

I. INTRODUCTION

Big Data is defined as- "The quantities, characters, or symbols on which operations are performed by a computer, which may be stored and transmitted in the form of electrical signals and recorded on magnetic, optical, or mechanical recording media. ".Big Data is an emerging technology which acts as a "fuel" for the future, an important engine driving social and economic development. It is a term that describes large volumes of data.This data can structure as well as unstructured. Due to explosive rise in data produced, big data is used to handle overwhelming data in business on a day to day basis. Big Data not only manages data but performs analysis on the data sets to get better insight to develop better decisions as well as strategic business moves. The uniqueness of Big Data lies in its structure and size. For Decades, Companies had been seeking for techniques that extract information to improve their business capabilities [1]. The New York Stock Exchange produces in an average of one terabyte of data per day. Single Jet engine can produce 10 plus terabytes of data in 30 minutes of a flight span. With thousands of flights daily, generation of data elevates up to even Petabytes. Statistics reveal that 500 plus terabytes of new data gets ingested into the databases of social media site Facebook, every day. This data generated consists of photo and video uploads, message exchanges etc. With Big Data finding information, whose extraction seems impossible by traditional means, has become fluid. Big Data comprise of 3V namely, high velocity, high volume, high variety data. Velocity is defined as the speed with which data spreads and changes among various data sets. Variety describes the heterogeneity present in diverse data sets. Volume describes the amount of data present to process or manage. Big data has proved to be a promising technology among various sectors of industries. Big data provides huge contribution under sectors for example public sector , Healthcare, Learning and Insurance Services, Industrialized and Natural Resources, Transportation Services, Banking Sectors and Fraud Detection. Recommendation systems are simple algorithms which aim to provide most relevant , accurate, preferential items to the user by filtering useful information from a huge pool of information base. Recommendation engines discovers data patterns in the data set by learning user's preferences and generates the outcomes that tallies to their needs ,interests and prefernces. As the days are passing by, there is huge amount of increase in amount of information in the www and at the same time there is a tremendous increase in number of users accessing these information .So, it becomes even more significant for the companies to search, map and generate them with relevant chunks of information according to their choices , needs, tastes and preferences. For this, recommendation systems are needed .These play a huge role in today's world. These are most importantly used for user satisfaction/contentment. Nowadays, companies are building smart, intelligent recommendation systems that can predict the users' preferences by studying the users' past behavior. Hence they provide users with recommendations and their tastes of interests as per "Movies of Interest" , "Suggested Videos", "People whom you may know on Facebook" , "Relevant Job Offers" , "Product recommendations on online sites like Amazon while buying a product" etc and is illustrated in Fig.1.
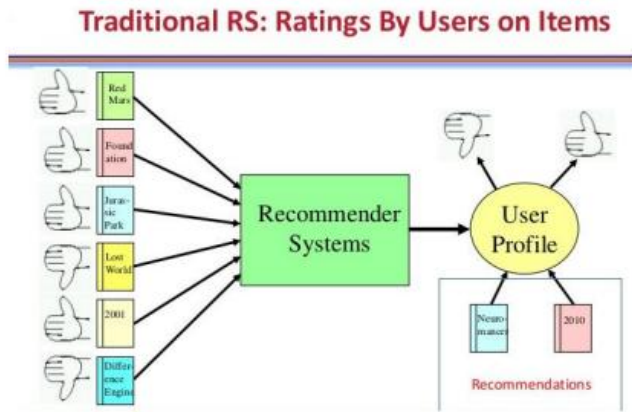
Figure 1: Traditional Recommendation system

There are many approaches to build recommendation systems .The most important approach is the collaborative filtering which is a widely used approach. This approach is based on gathering and evaluating a huge pool of information in terms of user's preferences, needs, activities and also predicting user's future needs based on their similarity analysis among other users. Collaborative Filtering method has two different approaches namely, memory based Collaborative Filtering approach and user based Collaborative filtering approach. The below mentioned scenario tries to explain a scenario in which collaborative filtering can be implemented. Let us suppose a person wants to fear a song can ask for recommendations from his/her friends. The recommendations of friends with similar interests can be used to give recommendation about other songs that the person might like to listen or songs that could be similar to his/her taste. This information is used in the decision on which movie to see. This approach doesn't depend on machine analyzable content [2]. The examples include k-nearest neighbor and Pearson Correlation approaches.

II. RELATED STUDY

A. Big Data Environment

Several papers were studied to understand the concept of Big Data and its application in real world. The paper "Big Data for Social Transportation" written by Xinhu Zheng, Wei Chen, Pu Wang, Dayong Shen, Songhang Chen, Xiao Wang, Qingpeng Zhang proposes a social transportation system which is data driven. Survey is performed upon various data sources, analytical approaches and application system for social transportation [3]. Another paper titled "Big Data Security and Privacy" discusses various challenges faced regarding data confidentiality, privacy and trustworthiness while working with big data. The paper "Challenges and Solutions of Information Security Issues in the Age of Big Data" by Yang Mengke ,ZhorXiaoguang , Zeng Jianqiu, XU Jianjan published in IEEE publication tells us that the wellfare of Big Data brings satisfaction to people's daily life as well as brings multiple challenges in information security. This paper also focuses on issues of information security and gives the necessary solutions for this. However, people can't enjoy the benefits of big data if the security is not strong/guaranteed [4]. The paper "Small Data Big Impact" was published by AislingKelliher, Virginial Tech in IEEE publication describes that sometimes having the correct amount of small data at the right time in the world of big data can pack a powerful punch. It has also presented its multistage design process and lessons from its ongoing project. An efficient framework for exchanging matching records using private record linkage protocol is discussed in the paper "A Hybrid Private Record Linkage Scheme: Separating Differentially Private Synopsis from matching record". Eli Collins' "Big Data in the Public Cloud" talks about the development of big data in public cloud environment [5]. Companies like Google,Yahoo,Microsoft were the first ones who introduces big data to the IT industry. "Kluster: An Efficient Scalable Procedure for Approximating the number of cluster in unsupervised learning" proposes a kluster procedure that applies statistical cluster methods on subsets of data drawn randomly to obtain optimal number of clusters. The above mentioned paper research is based upon the application of big data under the domain of Health. The paper titled "Chronic Diseases and Health Monitoring Big Data: A survey" performs a survey on chronic diseases and health monitoring big data technologies [6]. HuanKe, Peng Li,SongGho, Ivan Stojmenovic's "Aggregation on the Fly: Reducing Traffic for Big Data in the cloud" published in IEEE have described how to reduce the large amount of network traffic by designing their own architecture and using framework that is MapReduce. The results obtained from the experiments conducted proves that the given proposal is efficient in reducing the network traffic overall ArdiImawan, JoonhoKwon;'s "Time Visualization System for Road Traffic Big Data" published in IEEE is trying to describes how a timeline visualization system is implemented using road traffic related big data. This in turn provides ample opportunities in the field of road traffic [7].

B. Comparison on Recommendation techniques

"Social Media Recommender Systems : Review and Open Research Issues" provides an extensive review of Recommender System on Social Media based research articles published between the years 2011 and 2015 (inclusive). "Multirelational Social Recommendation via Multigraph Ranking" provides solution for the sparsity problem that is

faced while performing Collaborative Filtering rating by providing different user relationship by implementing. Hadi Habibzadeh, Andrew Baggio Dandry, Tolga Soyata,Burak Kantarci, Hussein T.Mouftah's "Software Sensing in Smart Cities :Handling 3Vs using Recommendation Systems, Machine Intelligence and Data Analytics" describes that in today's world smart transportation systems are very where the intelligent roads warn the drivers about the heavy traffic conditions. It is also provides way for intelligent ways of parking [8]. Recommendation systems, statistical reputation systems as well as context analysis are implemented here.. Xiaoyan Zhu, RipeiHaotian Chi, XiaojiangDu's "Fine Route Personalized and Time-Aware Route Recommendation Based on Check Ins" was published in IEEE talks about proper planning of routes which will enable people to enjoy travelling rapidly and save time and energy costs. Using data from Web 2.0 technologies, location based networks can be used for user's preference of routes and generate time information for recommending routes. "Web Services Recommendation via Exploiting location and QoS Information" proposes a personalized Recommender System for web services to provide users with optimal QoS (Quality of Service) Web Service by developing Collaborative Filtering approach based REcommendation System that employs context aware data such as user location and perform clustering user-wise as well as service-wise. The paper titled "System to Recommend the best place to live based on wellness state of user employing heart rate variability" evaluates various locations wellness rate in regard to the user by measuring user's Heart Rate Variability(HRV) to recommend user best suited house or apartment location where he/she could settle [9]. Another paper under the domain of recommender system proposing an application in the field of HealthCare was written by Paolo Pilloni, Lucas Piras, Salvatore Carta, Gianni Fenu, FabrizioMulas titled "Recommender System: Lets Coaches identify and help athletes who begin losing motivation" presents a web based application that proposes a monitoring approach to detect whether an athlete is losing motivation towards training and provide an explanation for it [10]. "Healthy Routes in the Smart City (A context- Aware Mobile Recommender) is another paper that proposes context aware recommender system under the field of Health Care that is used to recommend that healthy daily exercise routes for people in accordance to their health conditions by determining polluted routes and also determining measure that can be taken in order to reduce pollution in the concerned area by informing the concerned authorities [11]. "User Recommender System Based on Knowledge, Availability and reputation from interaction in Forums" proposes model for a recommender system that provides

the target user (asking solution for a problem P) with a list of candidate users who can provide solution [12]. The results obtained by the recommender System in the above mentioned paper are based upon reputation, knowledge and availability of user. The paper titled "Scientific Article Recommendation: Exploiting Common Author Relations and Historical Preferences" written by Feng Xia, Haifeng Lu, Ivan Lee, Longbing Cao proposes a recommender system that resolves problems such as providing user with articles written by an author whose article, user has already read as well as the model determines whether the target user wants author based recommendation been shown to him/her or not.

III. RESULTS AND OUTPUTS

Use The preferences of user tends to keep on changing with time in regards to certain ar- eas of interest, namely genres of movies, songs and reading. This volatile nature of preferences need to be handled while developing a recommendation system in order to gain more efficiency. UPD that is User preference dynamic is one such algorithm that intents to find the interest of user for an item in a quantitative value. Thus by applying this algorithm users top k areas of interest in certain field can be computed. Thereby instead of calculating recommendation under the area that user is not even interested in, focus of find recommends can be diverted towards these top k interest only thereby giving this security that whatever recommends that are giving to be generated.

RES tries to implement the phenomenon of resonance to find similar user. RES also uses jaccard factor to find the similarity between users. As seen from paper regarding RES it has been found that the efficiency of RES as compared with other similarity measures such as Cosine Similarity and PCC Similarity is higher and it also achieves success over many of the drawbacks of the latter similarity measures. In this project firstly the filtered out preferences for user are calculated with the help of UPD algorithm Now based upon these filtered out preferences similar users are extracted. After finding top k similar users , their watch list is searched through and all the items/movies that they have shown interest in and are yet not seen by the target users are filtered out as the recommendations.

The results obtained are the obtained from the movie lens data set given below:

53

Figure 2: Movielens Dataset

The prediction analysis used here have been generated using the formulae:

Precision=r/i

where

r=relevant instances and

i=retrieved instance

Recall=(r)/(Total amount of relevant instances)

F measure=1/(((1/Recall) +(1/Precision))/2)

Hence,the accuracy obtained is **84.65%**

Screenshots of the outputs generated are:

Figure 3: The first pop screen asking for user id

The output of the results is generated as follows:

Figure 4: Generating recommendations to the target users

CONCLUSION

After performing review on recommendation systems it is realized that present methods are just not very efficient in providing a detailed recommendation result. In dynamic users, there exist multiple details into which they are categorized as. But, users may not be sometimes satisfied with the recommendations as the interest may not stick to one required detail. So, to work upon this it is to improvise the attributes distinguishing the dynamic user interests and find a solution which reduces the problems faced by many techniques such as the Cold Start problem, Sparsity problem etc. We would be exploring in this context to determine whether such changes could be implemented to improvise the recommendation experiences for a user. We expect better results as it tends to capture both explicit and implicit features both.

REFERENCES

1. Rajatish Mukherjee, ParthaSarathi Dutta, GerdurJonsdottir&Sandip Sen "Movies2go-An Online Voting Based Movie Recommendation System".

2. Christoph Beckmann, Tom Gross "Towards a Group Recommender Process Model for ad-hoc Groups and On-Demand Recommendations".

3. Yan Tang, Mingzheng Li, Wangsong Wang, Pengcheng Xuan, Kun Geng"Quality-aware movie recommendation system on big-data".

4. Goral K. Godhani, MaulikDhamecha"Simulation of Genre Based movie recommendation system using

54

hadoop map reduce technique" 2007 IEEE, ICECDS-2017.

5. Lakshmi TharunPonnam, Sreenivasa Deepak Punyasamudram, Siva NagarajuNallagulla, Srikanth Yellamati "Movie Recommendation System Using Item Based Collaborative Filtering Techniques" 2016 IEEE.

6. Daniel Pomerantz, Gregory Dudek "Context Dependent Movie RecommendationsUsing a Hierarchical Bayesian Model" SpringerVerlag Berlin Heidelberg 2009.

7. ShinhyunAhn, Chung-Kon Shi "Exploring Movie Recommendation System UsingCultural Metadata" Springer-Verlag Berlin Heidelberg 2009.

8. Sonali R. Gandhi, Prof. JaydeepGheewala "A Survey on Recommendation System withCollaborative Filtering using Big Data" 2017 IEEE , ICIMIA 2017.

9. Anan Liu, Yongdong Zhang, Jintao Li "Personalized Movie Recommendation", 2009 ACM.

10. Panagiotis Symeonidis, Alexandros Nanopoulos, YannisManolopoulos "MoviExplain: A Recommender System with Explanations", 2009 ACM.

11. Suhang Wang, Jiliang Tang, Yilin Wang, Huan Liu "Exploring Hierarchical Structures in Recommendation System" IEEE Transactions On Data and Knowledge Engineering, Vol 30, June 2018.

12. YehudaKoren, Robert Bell, Chris Volinsky, "Matrix Factorization Techniques for Recommender Systems",2009 IEEE.

1. AUTHORS PROFILE

R. Lavanya completed her Master of Technology in Pondicherry University and currently pursuing Ph.D. Her research area is Data science and recommender systems. She has published more than 15 papers in reputed journals. She is a member of ISTE, IET etc. She is currently working as Assistant professor at SRM Institute of Science & Technology, Chennai.

Rithika Lahari completed her Graduation in Bachelor of Technology (Computer Science) from SRMIST, KTR with first class with distinction. She's currently working under the field of Data Science and currently working as Trainee decision scientist in Mu Sigma.

Palak Gupta completed has Graduation in Bachelor of Technology (Computer Science) from SRMIST, KTR. She's currently working under the field of Data Science and has job as Software developer engineer at Zoomrx.