An optimized recommender system with enhanced hybrid filtering approach

A Thesis submitted in the partial fulfillment of the requirements
for the award of the degree of
MASTER OF TECHNOLOGY
(INFORMATION SYSTEMS)

Submitted By:
SANCHIT JAIN
(Roll No. 2K12/ISY/23)

Under the esteemed guidance of
Mr. RAHUL KATARYA
Assistant Professor

DEPARTMENT OF INFORMATION TECHNOLOGY
DELHI TECHNOLOGICAL UNIVERSITY
BAWANA ROAD, DELHI-110042
SESSION: 2012-2014

CERTIFICATE

This is to certify that work entitled "**An optimized recommender system with enhanced hybrid filtering approach**" submitted by **Sanchit Jain (2k12/ISY/23)**, to Delhi Technological University, Delhi for the award of the degree of Master of Technology is a bonafide record of research work carried out by him under my supervision.

The content of this thesis, in full or parts, have not been submitted to any other institute or university for the award of any degree or diploma.

Mr. Rahul Katarya
Project Guide
Assistant Professor
Department of Information Technology
Delhi Technological University
Shahbad Daultpur, Bawana Road, Delhi-110042

Date:----------

ACKNOWLEDGEMENT

I would like to thank my project guide, **Mr. Rahul Katarya** for his valuable guidance and wisdom in coming up with this project. I humbly extend my words of gratitude to **Dr. O. P**. **Verma**, Head of Department, and other faculty members of IT department for providing their valuable help and time whenever it was required. I thank all my friends at DTU who were constantly supporting me throughout the execution of this thesis.

Special thanks to the Almighty Lord for giving me life and the strength to persevere through this work. Last but not least, I thank my family for believing in me and urging me. May you all be blessed.

Sanchit Jain
Roll No. 2k12/ISY/23
M.Tech (Information Systems)
E-mail: **omsanchitjain@gmail.com**
Department of Information Technology
Delhi Technological University

ABSTRACT

The tremendous development of computer and internet technology has really changed the life of human in most of the aspects. Internet has change the way in which information used to be processed earlier, it has not only provided as excellent way for information flow but provided new platform for social networking, business, entertainment etc. On internet now companies can market as well as sell their products that have revolutionized the way business has done earlier. It saves a lot human effort and processing time, but the problems is how to search the relevant data from the large bulk of data which is present on the internet. The solution is the recommender system, they are the software solution used by most of the web sites, they search the all the data and apply some forms of pre and post filtering, and then provide more relevant results to the user by eliminating the irreverent data. But after the drastic popularity of smart phone and as the number of smart phone users raised up to 1.6 billion, now there is a need to develop better recommender system using more enhanced filtering approaches which are based on attributes like personal factors, current location of the user and mobile usage pattern of the user which are called psychographic factors.

Table of Contents

List of Figures

CHAPTER 1

Introduction

1.Introduction

Long distance communication is a very important process as far as human development is concerned, in order to understand this we can start from the landmark in the human development where human first used pigeons for carrying latters. Then comes the postal system where a complex working mechanism is defined and exercised like country code, states code, city code, block code, street code, housecode, etc. Radio was the first electronic invention which was used as a simplex communication to inform the large population part. Even till date radio is one of the most important and efficient mean of communication.

1.1 Evolution of Communication

After that in early 19th century telephone comes into market which was the first time electrical signals are used to carry the signal of the human voice from one place to another, all these things has revolutionized the way human communication was done before. Mobile was another development after telephones as they remove the requirement to be present at any particular point. It has changed the way business has done before.

Figure 1.1 Data communication[53]

After the drastic development of electronic devices which finally leads to the development of computer has changed the way in which the general day to day work used to be done earlier. In 80's it become a common practice to feed the data into thecomputer instead of keeping the hard copy which first occupy a lot of space and

second waste human effort and money. And slowly the tradition changed and nowadays approximately every organization irrespective of the field that is either business or a government firm process all their data in computers. There are a number of advantages, like first it is easy to process the data into computer, second it is easy to maintain the data in the computer and security of data in a computer is also easy [39].

1.2 Computer Network

The miracle happens when all the computers in the world were getting connected though internet technology. Thereafter data started flowing free on the internet, with this information revolution, world wide web provided a new platform for not only information exchange but also business. Internet technology has really changed the life of humans in most of the aspects. The internet has changed the way in which information used to be processed earlier, it has not only provided as an excellent way for information flow, but provided a new platform for social networking, business, entertainment etc. [38]. On the internet now companies can market as well as selltheir products, which has revolutionized the way business has done earlier.Finally internet comes on the market in late 90es which has provided a platform to share knowledge. After than that, everything changed. It's not only a revolution in the field of technology or human communication, but it has also changed the business used to be done.

Figure 1.2 Computer network[53]

1.3 E-Business

As now days computers have a remarkable partnership in approximately every field in which human is developing for e.g., whether it is education and science, business, travel industry etc. computers have really reduced human effort and made the work easy because of which not only produced cost reduces but also time spends is also reduced. Although it is not only the computers which have revolutionized the current era, but the ability of technology to connect different computers together thought a communication link. The internet is a network of the entire computer in the world that is connected through a communication link.

So it's not only the computer, but the information revolution that has changed the entire scenario, now information is easy to access but a click away from your computer. Now we can access any required information through internet, because of which a lot of human effort is saved which was earlier waited to make the information accessible.

As a result of this information revolution, the internet has provided a new platform where we're is not just for information sharing, but also acts a new business market called e-business. In this new place now there is no need of a physical shop or a place from where a businessman can trade, here a business man advertises its product online and if any customer show interest to buy something the payment can also be made online, and then the product is delivered to the customer's address [35].

It saves a lot human effort and processing time, but the problems is how to search the relevant data from the large bulk of data which is present on the internet.

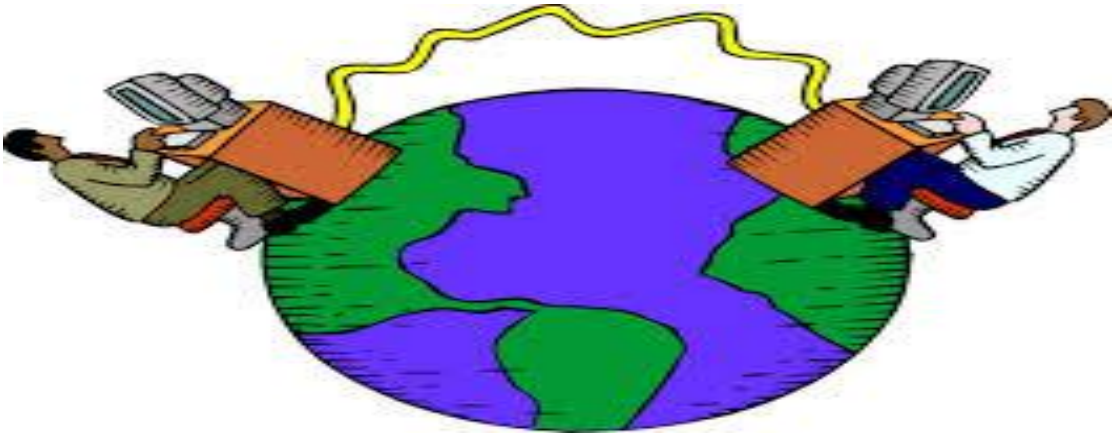

Figure 1.3 E-business relationship[53]

That is where the recommender system comes into the picture, they are the software solution used by most of the web sites, they search the all the data and apply some forms of pre and post filtering, and then provide more relevant results to the user by eliminating the irreverent data. They apply a number of data mining and clustering technique on the web data which is present on the web. And then narrow down the list of recommended item[36].

1.4 Web Mining

When we apply different data mining techniques to find out any information or pattern on the web data it is called web mining. Here we try to analyze patterns and collate in the data collection of the web site which may be even dynamic in nature by using traditional data mining techniques and attributes such as clustering and classification, association, and examination of sequential patterns.Web Mining tools analyze web logs for useful customer-related information that can help personalize web sites according to user behavior. Web mining tools are also used to search the web for key words, phrases, or other.

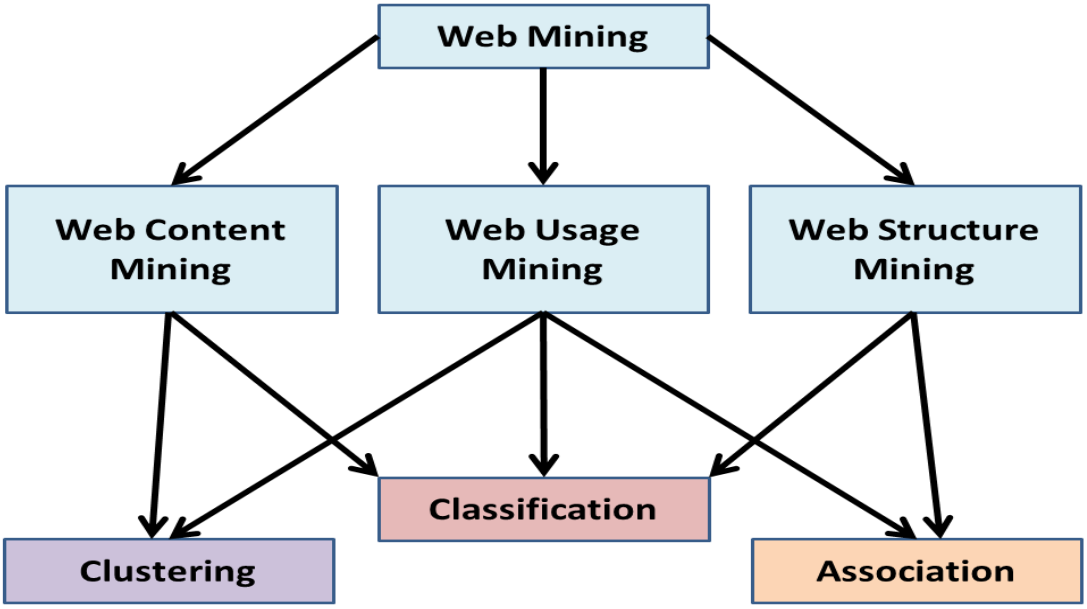

Figure 1.4 Web mining classification[54]

.Open source software for web mining includes Rapid Miner, which provides modules for text clustering, text categorization, information extraction, named entity recognition, and sentiment analysis. Rapid Miner is used for example in applications like automated news filtering for personalized news surveys[37]. It is also used in automated content-based document and e-mail routing, sentiment analysis from web blogs and product reviews in internet discussion groups. Information extraction from web pages also utilizes RapidMiner to create mash-ups which combine information from various web services and web pages, to perform web log mining and web usage mining.

1.4.1 Web Usages Mining

Web usage mining is the process of extracting useful information from server logs that is user history. Web usage mining is the process of finding out what users are looking for on internet. Some user might be looking at only textual data; whereas some other might be interested in multimedia data. This type of mining allows us to gather the collection of all web access information of the web.

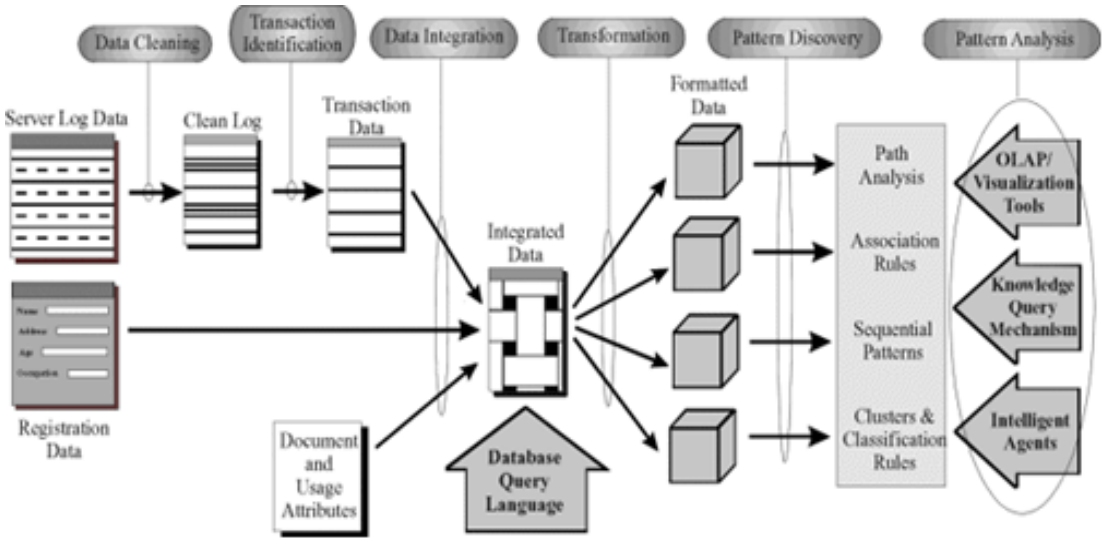

Figure 1.5Web mining process[53]

Web usage mining allows companies to produce productive information which is very useful for company growth. Usage mining is valuable to e-business whose business is based solely on the traffic provide through search engine[33]. Web usage mining helps to gather the important information from customer visiting the site. This enables an in-depth log to complete analysis of a company's productivity flow. E-business depends on this information to direct the company to the most effective web server for promotion of their product or service.

Web usage mining enables web based business to provide the best access routes to service or other advertisements. Analysis of this usage data will provide the companies with the information needed to provide an effective presence to their customers [34].

Usage mining has valuable uses to the marketing of business and a direct impact on the success of their promotional strategies and internet traffic.

1.4.2 Web Structure Mining

Web structure mining is a tool used to identify the relationship between web pages linked by information or direct link connection. This connection allows a search engine to pull data relating to a search query directly to the linking web page from the web site the content rests upon. This completion takes place through use of spiders

scanning the web site, retrieving the home page, then linking the information through referral links to bring forth the specific page containing the derived information[30].

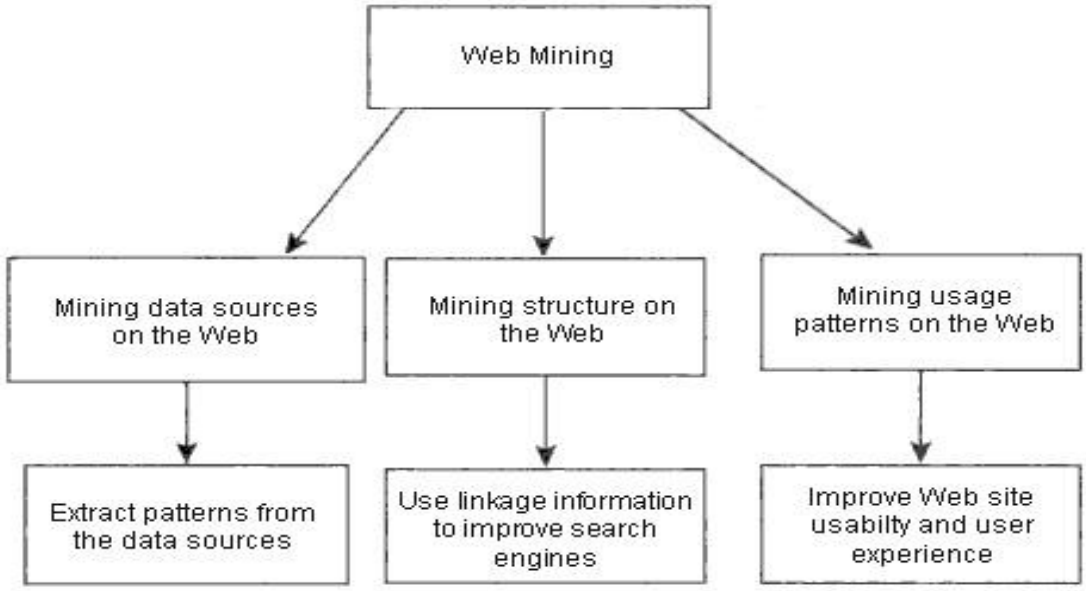

Figure1.6 Data mining on web[53]

Structure web mining try to minimize to main problem of World Wide Web due of its vast amount of information.

(i) Irrelevant search result

(ii) Inability to index the vast amount if the information provided on the web.

Structure data mining provides use for a business to link the information of its own web site to enable navigation and cluster information into site maps. This allow its user the ability to access the desired information through association and content mining. This mining can be quite useful in determining the connection between two or more web sites [24].

1.4.3 Web Content Mining

Web content mining is the scanning and mining of text, pictures and graphs of a web page to determine the relevance of the content to the search query[48]. With the massive amount of information that is available on the World Wide Web, content mining provides the results lists to search engines In order of highest relevance to the keywords in the query.

Text mining is direct towards specific information provided by the customer search information provided by the customer search information in search engine. This allows for the scanning of the entire web to retrieve the cluster content triggering the scanning of specific web pages within clusters. The results are pages related to the search engine through the highest level of relevance to the lowest.

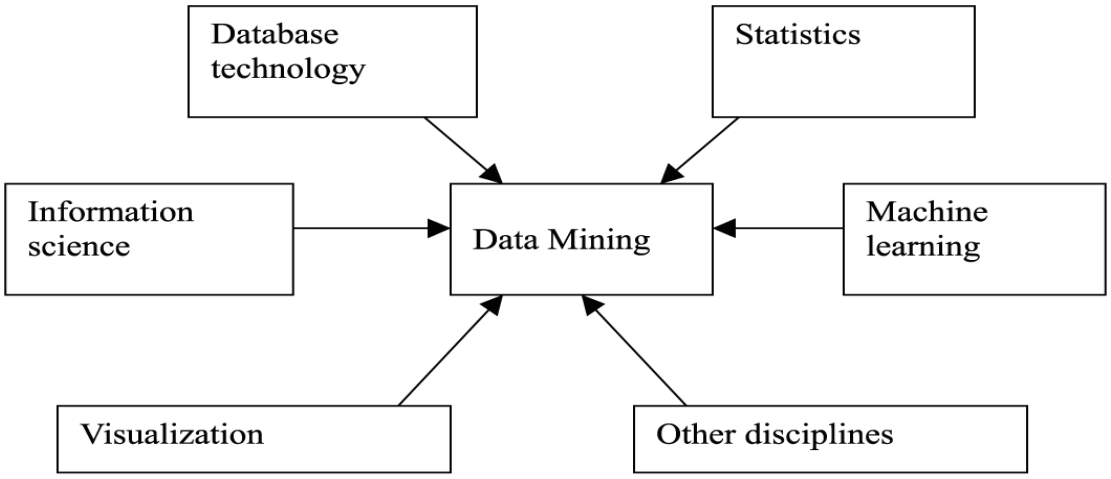

Figure 1.7 Data sources[53]

The main use of this mining is to gather, categorize, organize and provide the best possible information available on the WWW to the user requesting the information. Business uses of content mining allow for the information provided on their site to be structured in a relevance order site map used as a marketing tool that provides additional traffic to the web pages.

1.5 Recommender System

As the internet has provided a new market for business, there is a competition among the trading web sites to see their products, but it is not an easy task. There are also some short coming of internet, like as it is a huge amount of data on the internet[50] , it becomes difficult to search the relevant data or information. The software solution which was introduced in the market was recommender system. They are the software's which provide various types of filtering to refine the search domain of the user in order to provide him the most relevant result.

Recommender systems or recommendation systems (sometimes replacing "system" with a synonym such as platform or engine) are a subclass of the information filtering

system that seek to predict the 'rating' or 'preference' that a user would give to an item [27].

Recommender systems have become extremely common in recent years, and are applied in a variety of applications. The most popular ones are probably movies, music, news, books, research articles, search queries, social tags, and products in general. However, there are also recommender systems for experts, jokes, restaurants, financial services, life insurance, persons (online dating), and twitter followers.

Recommender systems typically produce a list of recommendations in one of two ways - through collaborative or content-based filtering. Collaborative filtering approaches build a model from a user's past behavior (items previously purchased or selected and/or numerical ratings given to those items) as well as similar decisions made by other users; then use that model to predict items (or ratings for items) that the user may have an interest in[29]. Content-based filtering approaches utilize a series of discrete characteristics of an item in order to recommend additional items with similar properties. These approaches are often combined (see Hybrid Recommender Systems).

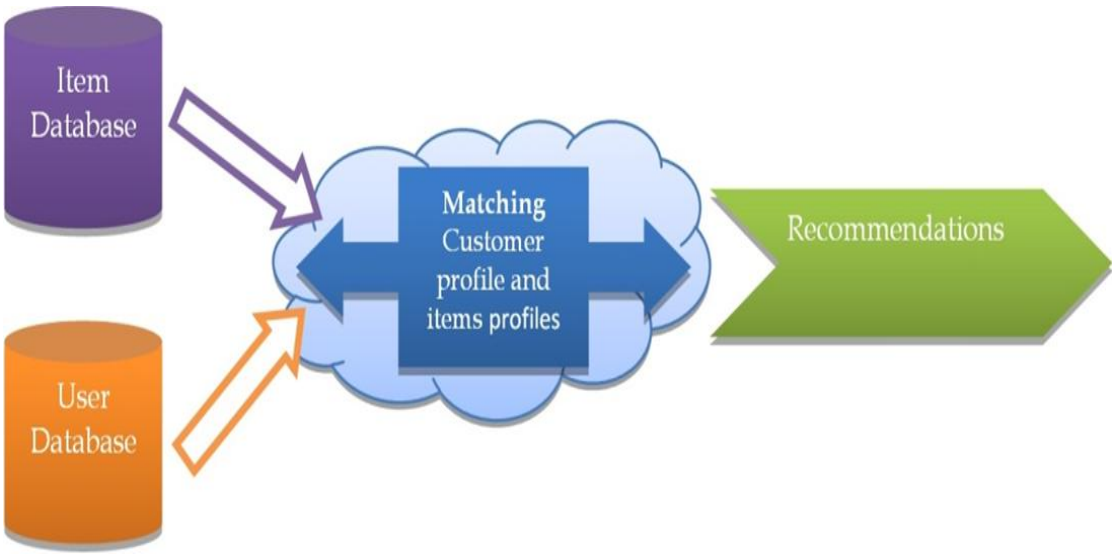

Figure1.8 Recommendation process[53]

1.6 Collaborative Filtering

Collaborative filtering is the process of filtering for information or patterns using techniques involving collaboration among multiple agents, viewpoints, data sources, etc. Applications of collaborative filtering typically involve very large data sets. Collaborative filtering methods have been applied to many different kinds of data including: sensing and monitoring data, such as in mineral exploration, environmental sensing over large areas or multiple sensors; financial data, such as financial service institutions that integrate many financial sources; or in electronic commerce and web applications where the focus is on user data, etc[26].

collaborative filtering is a method of making automatic predictions (filtering) about the interests of a user by collecting preferences or taste information from many users (collaborating). The underlying assumption of the collaborative filtering approach is that if a person *A* has the same opinion as a person *B* on an issue, A is more likely to have B's opinion on a different issue *x* than to have the opinion on x of a person chosen randomly. For example, a collaborative filtering recommendation system for television tastes could make predictions about which television show a user should like given a partial list of that user's tastes (likes or dislikes).

The motivation for collaborative filtering comes from the idea that people often get the best recommendations from someone with similar tastes to themselves.

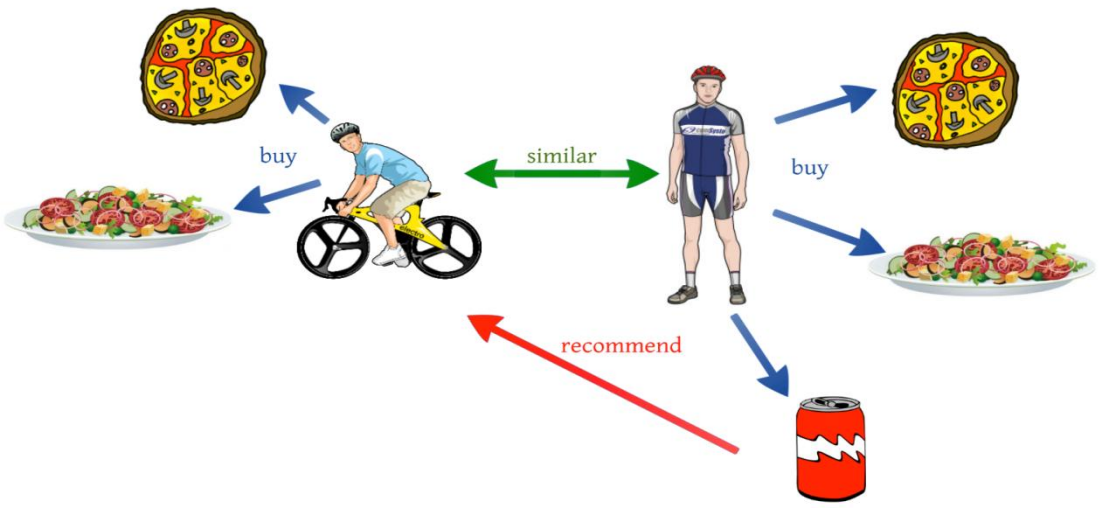

Figure 1.9 Collaborative filtering[55]

Collaborative filtering explores techniques for matching people with similar interests and making recommendations on this basis. A key problem of collaborative filtering is how to combine and weight the preferences of user's neighbors. Sometimes, users can immediately rate the recommended items. As a result, the system gains an increasingly accurate representation of user preferences over time. Collaborative filtering systems are able to generate more personalized recommendations by analyzing information from the past activity of a specific user, or the history of other users deemed to be of similar taste to a given user. These resources are used as user profiling and help the site recommend content on a user-by-user basis. The more a given user makes use of the system, the better the recommendations become, as the system gains data to improve its model of that user.

1.7 Content Based Filtering

Content-based filtering methods are based on a description of the item and a profile of the user's preference. In a content-based recommender system, keywords are used to describe the items; beside, a user profile is built to indicate the type of item this user likes. In other words, these algorithms try to recommend items that are similar to those that a user liked in the past (or is examining in the present)[52]. In particular, various candidate items are compared with items previously rated by the user and the best-matching items are recommended. This approach has its roots in information retrieval and information filtering research.

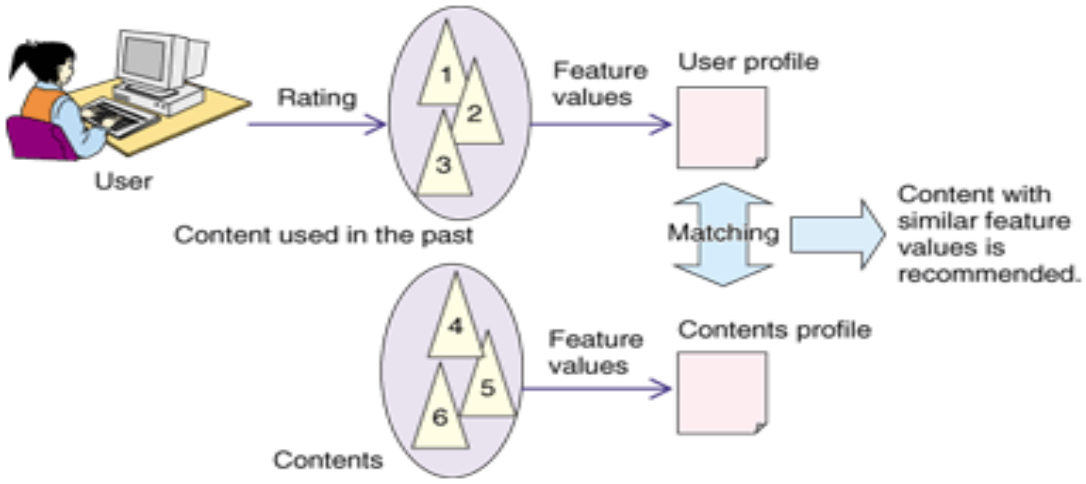

Figure 1.10 Content based filtering[56]

To create a user profile, the system mostly focuses on two types of information:

1. A model of the user's preference.

2. A history of the user's interaction with the recommender system.

Basically, these methods use an item profile (i.e., a set of discrete attributes and features) characterizing the item within the system. The system creates a content-based profile of users based on a weighted vector of item features. The weights denote the importance of each feature to the user and can be computed from individually rated content vectors using a variety of techniques[46]. Simple approaches use the average values of the rated item vector while other sophisticated methods use machine learning techniques such as Bayesian Classifiers, cluster analysis, decision trees, and artificial neural networks in order to estimate the probability that the user is going to like the item[22].

A key issue with content-based filtering is whether the system is able to learn user preferences from a user's actions regarding one content source and use them across other content types. When the system is limited to recommending content of the same type as the user is already using, the value from the recommendation system is significantly less than when other content types from other services can be recommended. For example, recommending news articles based on browsing of news is useful, but it's much more useful when music, videos, products, discussions etc. from different services can be recommended based on news browsing.

Advantages and Disadvantages of CF and CBF

	Collaborative	Content based
Knowledge engineering effort	minimal	considerable
User rating effort	higher	lower
Sparse data problem		
cold start user	✗	✗
cold start item	✗	✓
Concept drift	✗	✓
Serendipity	✓	✗
Contextualization	✗	✓
Recommendation quality	higher	lower

Figure 1.11 Comparison between CF and CBF[53]

1.8 Hybrid Filtering

Content-based filtering research has demonstrated that a hybrid approach, combining collaborative filtering and content-based filtering could be more effective in some cases[45]. Hybrid approaches can be implemented in several ways: by making content-based and collaborative-based predictions separately and then combining them; by adding content-based capabilities to a collaborative-based approach (and vice versa); or by unifying the approaches into one model. Several studies empirically compare the performance of the hybrid with the pure collaborative and content-based methods and demonstrate that the hybrid methods can provide more accurate recommendations than pure approaches. These methods can also be used to overcome some of the common problems in recommender systems such as cold start and the sparsity problem[43].

Netflix is a good example of hybrid systems. They make recommendations by comparing the watching and searching habits of similar users (i.e. Collaborative filtering) as well as by offering movies that share characteristics with films that a user has rated highly (content-based filtering).

A variety of techniques have been proposed as the basis for recommender systems: collaborative, content-based, knowledge-based, and demographic techniques. Each of these techniques has known shortcomings, such as the well-known cold-start problem for collaborative and content-based systems (what to do with new users with few ratings) and the knowledge engineering bottleneck in knowledge-based approaches. A hybrid recommender system is one that combines multiple techniques together to achieve some synergy between them[44].

Collaborative: The system generates recommendations using only information about rating profiles for different users. Collaborative systems locate peer users with a rating history similar to the current user and generate recommendations using this neighborhood.

Content-based: The system generates recommendations from two sources: the features associated with the products and the ratings that a user has given them. Content-based recommenders treat recommendation as a user-specific classification problem and learn a classifier for the user's likes and dislikes based on product features.

Demographic: A demographic recommender provides recommendations based on a demographic profile of the user. Recommended products can be produced for different demographic niches, by combining the ratings of users in those niches.

Knowledge-based: A knowledge-based recommender suggests products based on inferences about a user's needs and preferences. This knowledge will sometimes contain explicit functional knowledge about how certain product features meet user needs.

The term hybrid recommender system is used here to describe any recommender system that combines multiple recommendation techniques together to produce its output. There is no reason why several different techniques of the same type could not be hybridized, for example, two different content-based recommenders could work together, and a number of projects have investigated this type of hybrid: News Dude, which uses both naive Bayes and KNN classifiers in its news recommendations, is just one example[29].

CHAPTER 2

Related Work And Research Motivation

2.1 Important literatures

In paper [1] author explained that how recommender system gained importance after some initial paper on collaborative filtering in 1990's. The basic motivation behind the paper was the deficiency of good comprehensive literature reviews and classification of the research field. Here author mainly concentrated on last 10 years' journals and then classified them according to their data mining techniques, their application field, the journal in which they appeared and the year of publication. They mainly tried to divide them in the categories like music, movies, images, documents, books, shopping, TV program and other.

Here the author also tried to identify the main data mining technique link analysis, k-nearest neighbor, clustering, association rule, decision tree, neural network, regression, and other heuristic method [2]. Here the author tried to find the trend of the research in the fields of recommender system and tried to help and simulates the research in the correct direction. In general, recommender systems directly help users to find content, products, or services(such as books, digital products, movies, music, TV programs, and web sites) by aggregating and analyzing suggestions from other users, which mean reviews from various authorities, and users.[3]

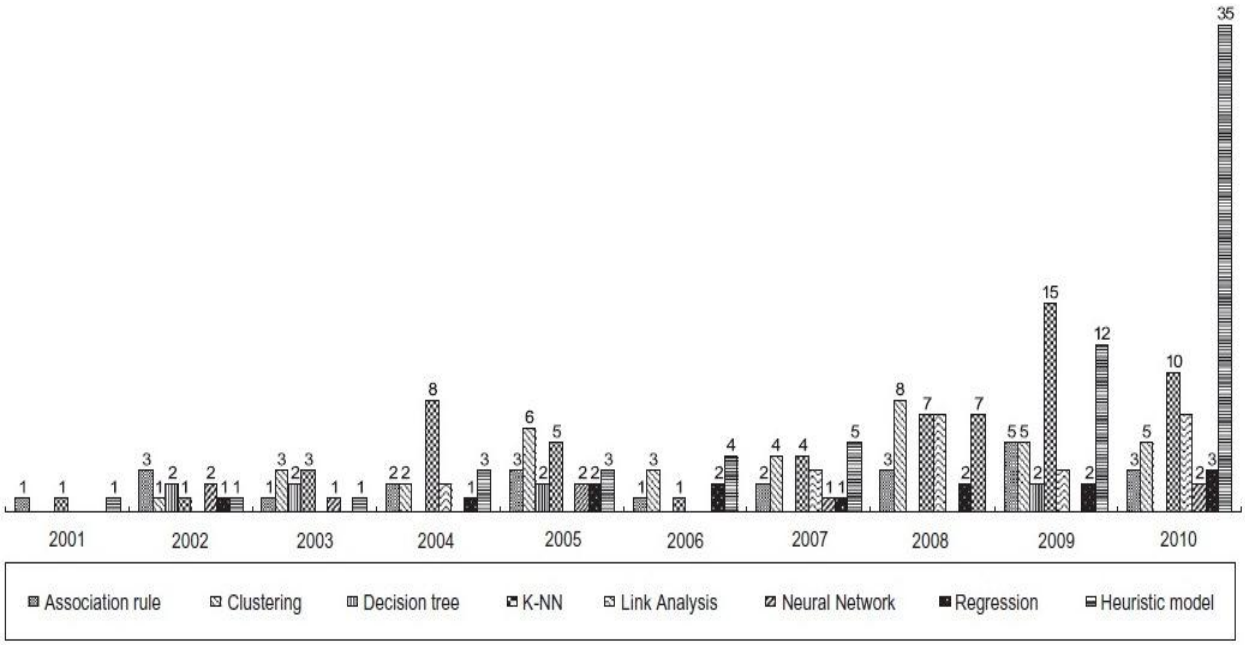

Figure. 2.1. Distribution of research papers by publication year and application fields.[1]

Here we tried to develop a analytic technology where the system tries to predict the probability the user will purchase certain items after a number of recommendations. Normally recommendation system is based on collaborative and content based filtering, basically collaborative filtering try to identify the patter from user's previous purchase, but there are two important issues first scarcity problem and the scalability problem.[4]

Basically content based filtering trialsto identify various products, services and documents rated by individual and the trees to infer that its profile can be used how to select the next recommendation.[5] But the problem with content based filtering id's that sometimes because of an overemphasis on the detection of similarities between the items, it predicts only those item items that are very similar to each other.[6]

2.2 Recent Changes in Research Methodology

But in the recent time, researchers tried to over the problems associated with collaborative and content based filtering and tried to implement them in real world situations. All this will help to provide personalized information to the user by analyzing the behavior or usage pattern. In their research author first explained the methodology they have used in the study, then they explained the criteria based on which they classified the papers followed by the conclusion which comes and their limitation. The authors mainly used keywords like recommender system, recommendation system, personalization system, collaborative filtering and content filtering. The authors did not consider either master or doctoral dissertations.

One author explained that recommender system is nothing but the software solution which helps a user to purchase a product in a better way [7]. Even today after a good amount of research it is not easy to classify papers in this area. Data mining techniques can be understood as a process of extracting meaningful information from the data which can be used in a constructive manner [8]. Here the author basically tried to classify data mining techniques into eight categories [42].

First is association rule mining with the help of which we try to understand normal behavior of the user towards the type of products. a set of transactions in which each transaction contains a set of items, an association rule applies the form X)Y, where X

and Y are two sets of items[9]. The cluster tries to identify a finite set of categories or clusters to describe data. Among some of the clustering methods some of the most popular are K-means and self-organizing map. Basically, K-means take k input parameters and then divide the data into k number of clusters [8], while on the other hand self-organizing map uses the methods of unsupervised learning based on artificial neurons clustering technique. One very popular technique is decision tree, A decision tree is a tree in which each internal (non-leaf) node represents a test on an attribute, each branch represents an outcome of the test, and each terminal (leaf) node represents a class prediction [10].

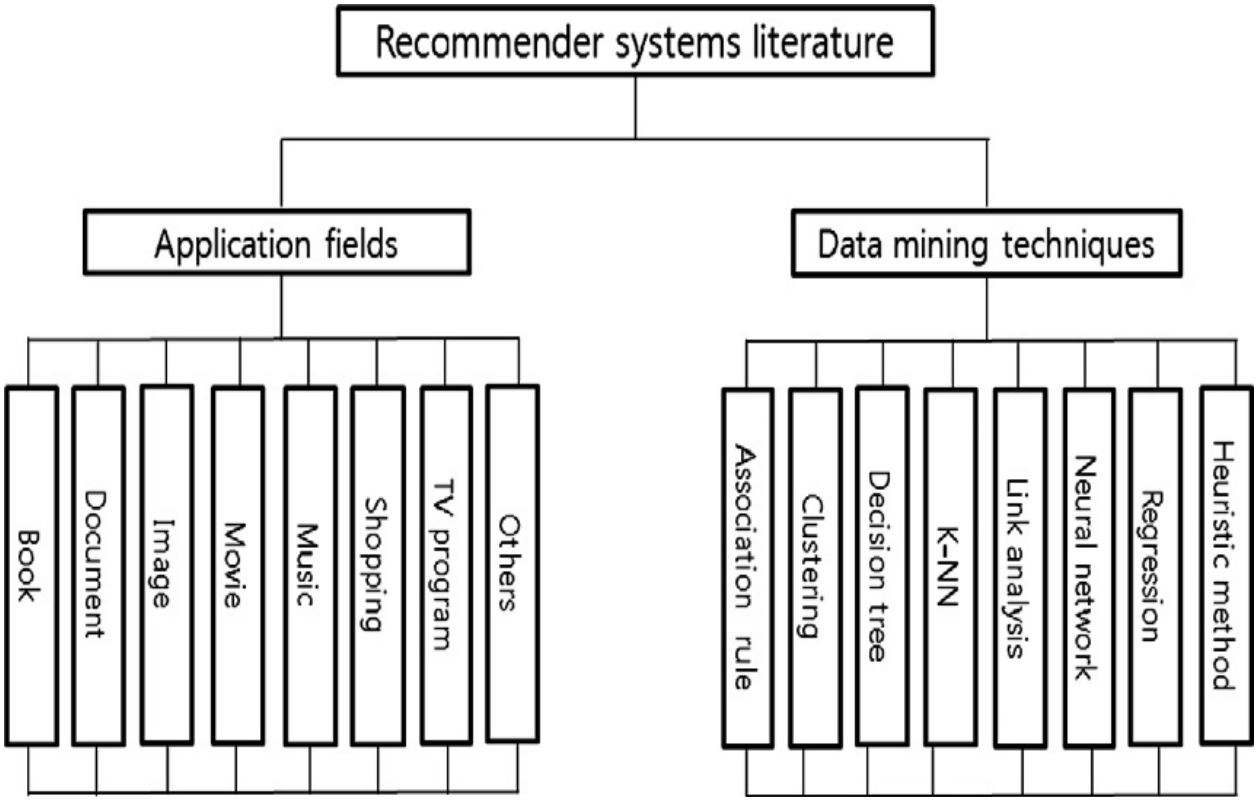

Figure2.2 Classification framework [1]

A K-nearest neighbor is also a traditional collaborative filtering based recommendation system, which normally works in three phases. In the very first step recommender system try to construct a user profile by either understanding the rating given by the user, to either a product or service or by understanding the purchase pattern of the user. After that recommender system tries to make k clusters with a specific dissimilarity pattern, and then try to classify the other users in these clusters based on their similarities with the cluster centroid. Then in the final step

recommender system try to predict the future behavior of the user in any particular cluster based on the behavior of the other user in the group[11].

2.3 Use of Neural Networks

In the approach of using a neural network, a parallel distributed information system is used which is able to learn and self-organize its self. It can further be understood a large number of entities which are interconnected to perform any computational task[12]. Mostly neural networks are used in varieties of applications such as classification; non-linear regression and prediction. They are a very pliable model that can be applied in verities of practical application [13].

The next approach is link analysis where we try to discover the relation between the domains of large database. One approach of doing this is social network analysis, which is a sociological approach where we try to find out the fundamental social structure by finding the social relationships and intersections among the social actors. It has been understood after a significant amount of research that link analysis can be used in predicting and improving the accuracy of the web search. Most of the link analysis algorithms use HITS algorithms and page ranks and they uses web page as a node in the web graph [15]. In another approach where regression is used to find the relationship between a dependent variable and one or many independent variables, mostly this technique is used in other practical applications like testing, systematic hypothesis about relationships between variables, prediction and sometime curve fitting[15].

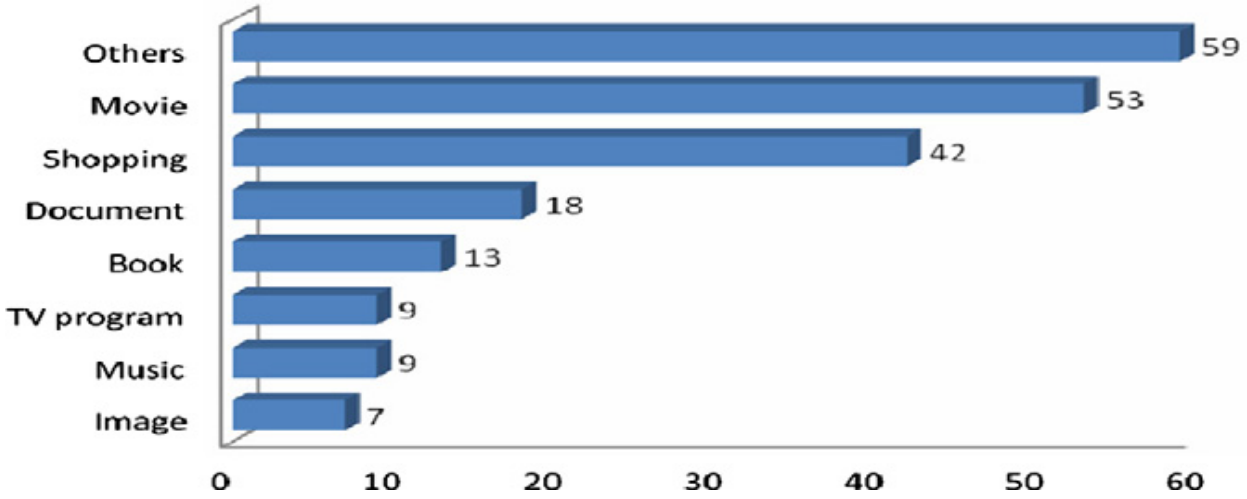

Figure. 2.3. Distribution of research papers in application fields [1]

2.4 Areas of Application

On last some recommender system uses other heuristic methods that have been developed by mixing any two previous methods or adding some new logic to some previously existing method like ontology. It is understood that there was a steady growth in the field of recommender system from the year 2000 to 2004, then in 2005-06 it is realized that recommender system can be applied to various types of field starting from a movie and shopping recommendation then the growth of recommender system increased rapidly after 2007[51].

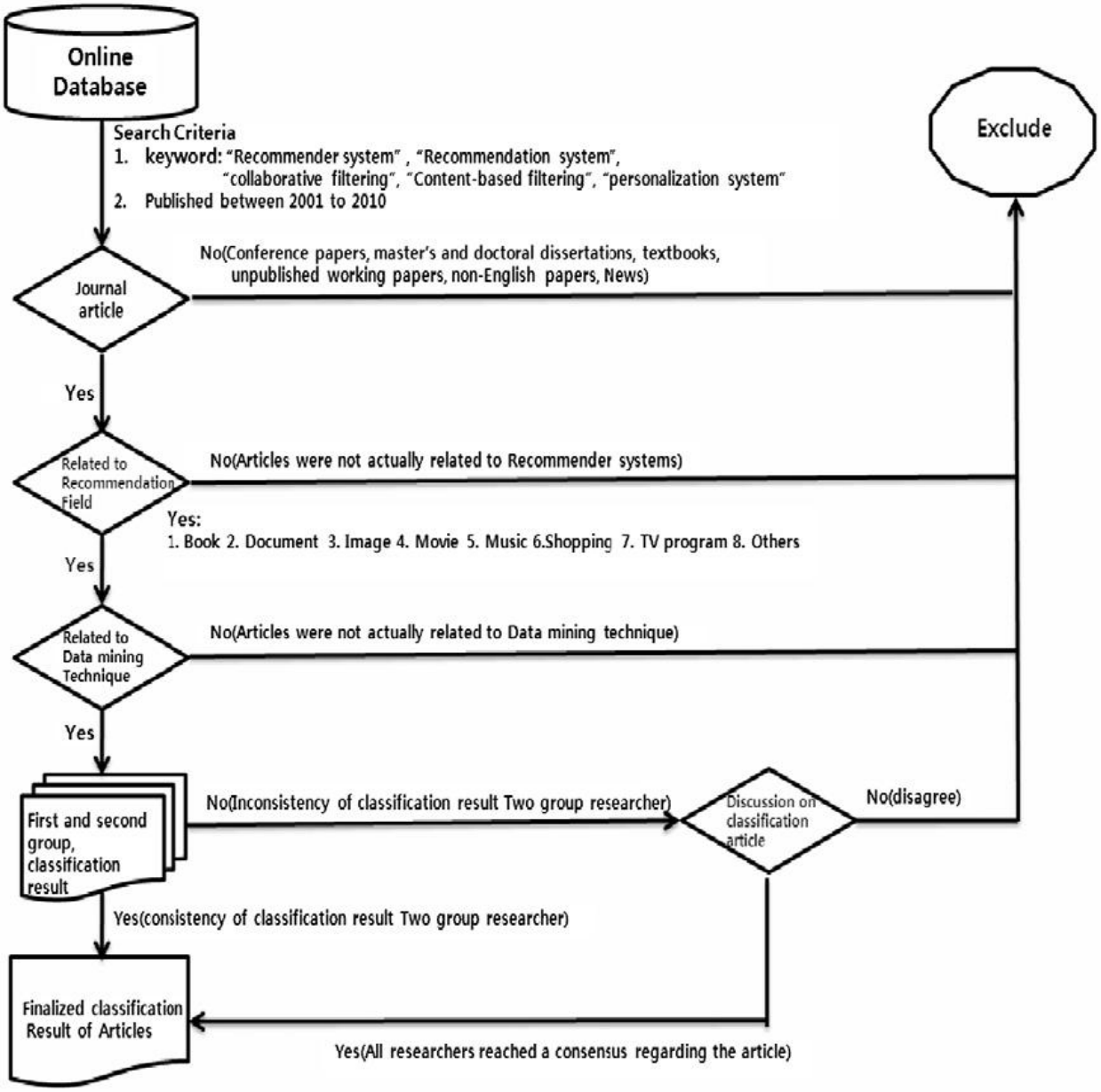

Figure. 2.4. Selection criteria and evaluation framework[1]

It is found under the literature survey that out of total recommendation systems approximately 25% of the research papers are movie recommender systems, remaining 20% are shopping recommender system, as a large majority of papers and practical system are related to movie and shopping recommendation so a very few papers are found in other areas like images, movies and TV programs, as the movie lens provide free api of their data, it is also a reason why most of the research paper deals with movie recommendation[41].

2.5 Most Popular Algorithms

By a logical and proper review of the research work done in the previous 10 years it can be understood that what are the major techniques that are used for designing sophisticated recommender systems like heuristic models, clustering algorithms, association rules, decision trees and neural networks.

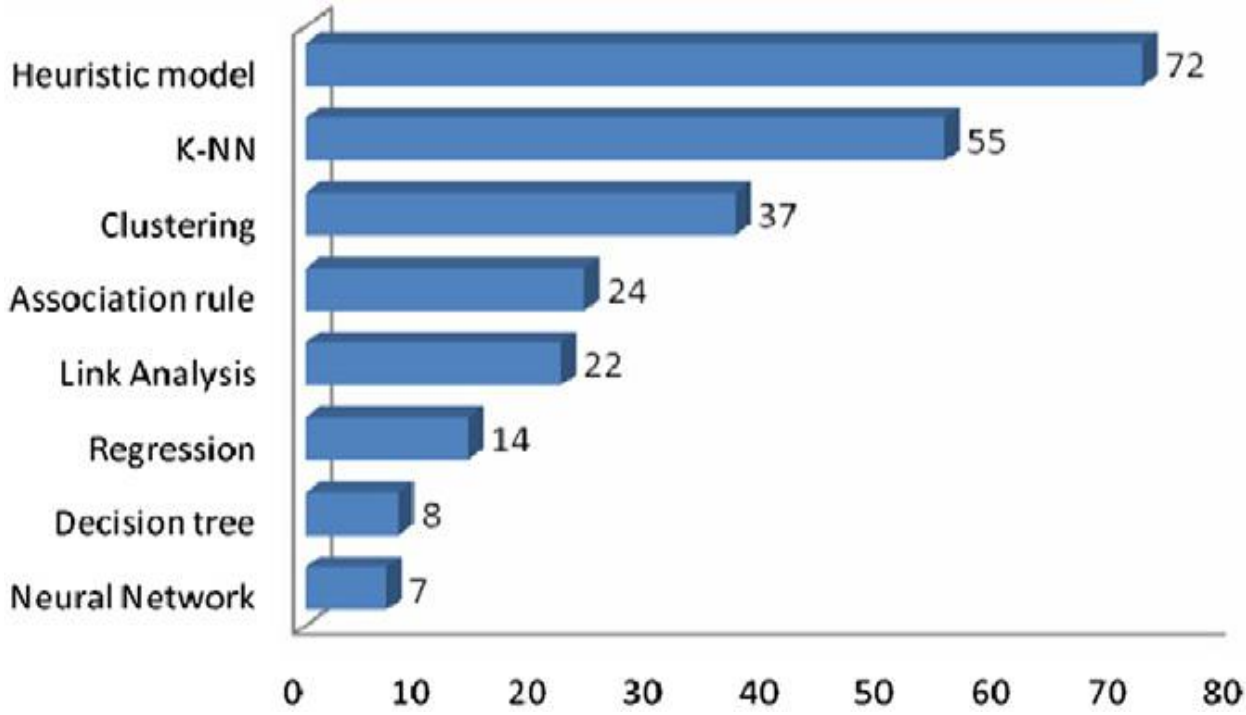

Figure.2.5 Distribution of research papers by application fields[1]

After all this analysis we can understood that recommender system have attracted the attention of academics and practitioners. One very common thing which can be understood is in the future the importance of the recommender system will further increase. By the literature survey we can also understand that as most of the papers are related to movie recommendation that is a need to motivate research in the field of

image, music and tv recommendation. Most of the research papers were based on clustering, linkage analysis, association rule mining etc.

Now a days social networks are getting more and more importance day by day, but a very less amount of work is done in this area till this time, so it required to have more practice and theoretical work in a recommender system relating to social networks.

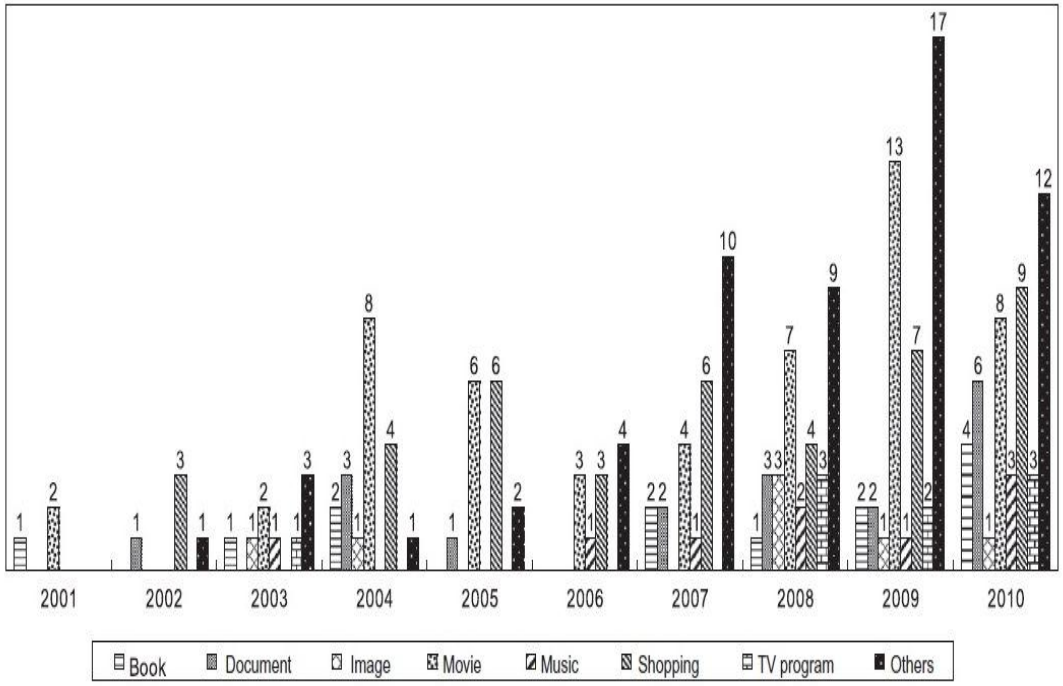

Figure. 2.6. Distribution of research papers by publication year and data mining technique[1]

Apart from these important issues we can understand some other important issues from this literature survey.

2.6 Changes InThe Trend of Application With Time

In starting mostly recommender system were using demographic, content based and collaborative filtering. Now a days these systems are trying to incorporate social information so that these systems can be made more effective, by taking social information we mainly try to incorporate personal information, local information and other implicit information. Now we will try to understand recommender system in deep with collaborative filtering method and algorithms. By starting from evolution of recommender technology and coming to the current level of technology in a systematic manner we can understand the current area of importance in the field of

recommender system and can also identifies the future area of importance, on which further work can be done to make the system more effective.

The basic logic of recommender system is to collect information about the user from different resources like preference of the user for a set of items like movie, TV programs, songs, travelling places, etc. other source of information may be like net browsing behavior, application downloaded on the phone etc., and some time explicit asking of information and ratings. Gps and some other latest technology are also sometimes used now a day to provide better recommendations [40]. The recommender system tries to balance different criteria's like novelty, disparity, accuracy and stability. The basic concept of content based filtering is same a like a human made decisions, first we consider past experience and then knowledge which is not confined to individuals, but also knowledge of our friends and known once to make the best decision possible. The logic of filtering used by most of the recommender system in the beginning (demographic, collaborative and content-based) were described in[16].

Then in one of the most important papers on collaborative filtering [17]. Explained the basis and evolution of collaborative filtering.Through the process of evolution of recommender system we slowly understood the importance of hybrid filtering, in one of the very important survey of on the hybrid filtering technique of recommender system [18].Although in that survey they did not give part to social mining which has become very popular in the recent year with the advancement in the social networking.

In starting collaborative filtering neighborhood based method was very popular in starting [19], here the author provided a guideline for neighbor based method for collaborative filtering. In paper [20] author explained about numerous problems of collaborative filtering like scarcity of collaborative based filtering methods, cold startproblem, etc. there are already number of papers in the field of collaborative filtering. If we start understanding the evolution of the recommender system in starting most the papers were based on traditional web based application, followed by the second phase where the concentration was more on the social web, and finally the third phase the internet technology. If we try to understand every logic in one by one

fashion like recommendation system foundation, k- nearest neighbor algorithm, cold start problem, similarity metrics and the evaluation of the recommender system.

Now we will try to understand some the issues in deep and explore different application of recommendation system in different areas, first we will try to classify the approach related to social information like social tagging, taxonomies, content based filtering of social data, followers, followed, trust, reputation, credibility etc[21].Now we will also discuss some of the latest advancement of the recommender system technology like bio-inspired approaches and using location of the user to provide better recommendation. We will also try to understand some other important issues like health parameters, security, privacy, teleoperations, Telepresence, p2p information and raw data. The normally recommender system makes use of the following sources of information for better prediction. Content, demographic, social, collaborative, geographic, knowledge-based, sensor, implicit and explicit data acquisition etc. [22].

In a systematic process in this thesis I first selected the most important papers and then first tried to understand the basic methods, algorithms and models which are most popular in the fields of recommender system. Mostly we tried to divide the papers and system based on their filtering approach like collaborative, content based, hybrid, etc., then again systematically I tried to understand what are the measure and quality criteria's based on which we can judge a paper or a system. First, we want to discuss the basic concepts and key issues which are very necessary for understanding the recommendation system in deep like data bases, filtering technology, methods, algorithms and their taxonomies. First, wemust understand the quality and type of data which is available like user registration information, the rating given by the user, features and content of the item that can be ranked, location of the user and social relationship among the user. Then next thing which is important is what is the filtering algorithm should be used for an efficient recommender system that must have most exact and efficient predication about the user behavior[24].

2.7 Basic Approaches of Filtering Using in Most Papers

The basic concept of any recommender system is based on the filtering method it uses that is what is the basic algorithm on which the entire system is dependent on. Commonly collaborative, demographic, content based and hybrid filtering the most common approaches used. Content based filtering mostly uses the previous behavior of the user. For the next future prediction for e.g. if a user purchase a ticket for a cricket match then when the next time a cricket match is held then the recommender system will recommend the match ticket in the user based on its previous purchasing behaviors [25]. On the other hand the content based recommender system tries to find similarity between your previous selection and other objects or services. And then provide new recommendation. Demographic filtering approach is that the individuals with common personal attributes like age, country, and sex while have the common choice or behavior. Collaborative filtering is also another interesting approach where first a user is asked to give ratings to a set of elements such as films, songs, videosetc. After this phase when enough information is stored on the system, then we can make recommendation to the other user if we find that a lot of things are similar between them[49].

One of the most commonly used algorithms in collaborative filtering is the k Nearest Neighbor approach. Where first the k number of neighbors are decided of a user, then an item of finding out for which the main user did not give any rating. Theme based on the rating given by the neighbors of the user the rating values of the item are predicted. Hybrid filtering is mostly the combination of the collaborative filtering with demographic filtering or the combination of collaborative or content based filtering [26]. By this technique we try to exploit the merits of different filtering approaches. One another taxonomy, which is also widely accepted is to divide the recommended methods into a model based and memory based methods.

Memory based methods first concentrate on the matrix of the user rating for the items which are properly maintained in the system. After that memory based methods try to find the similarity based on their ratios between any two users and between any two items. On the other hand, model based methods use the information which is given in the recommender system so that we can create a model that can give recommendations. Some of the most commonly used model is matrix classification,

latent feature, genetic algorithms, neural network, fuzzy system and Bayesian classifier. WE found out that some of the studies have also user dimensionality reduction technique, so that we can reduce problems from large level of scarcity from the recommender system data base. Matrix factorization is used for reduction methods, in case a large database and providing scalable approaches matrix factorization is specially used[27].

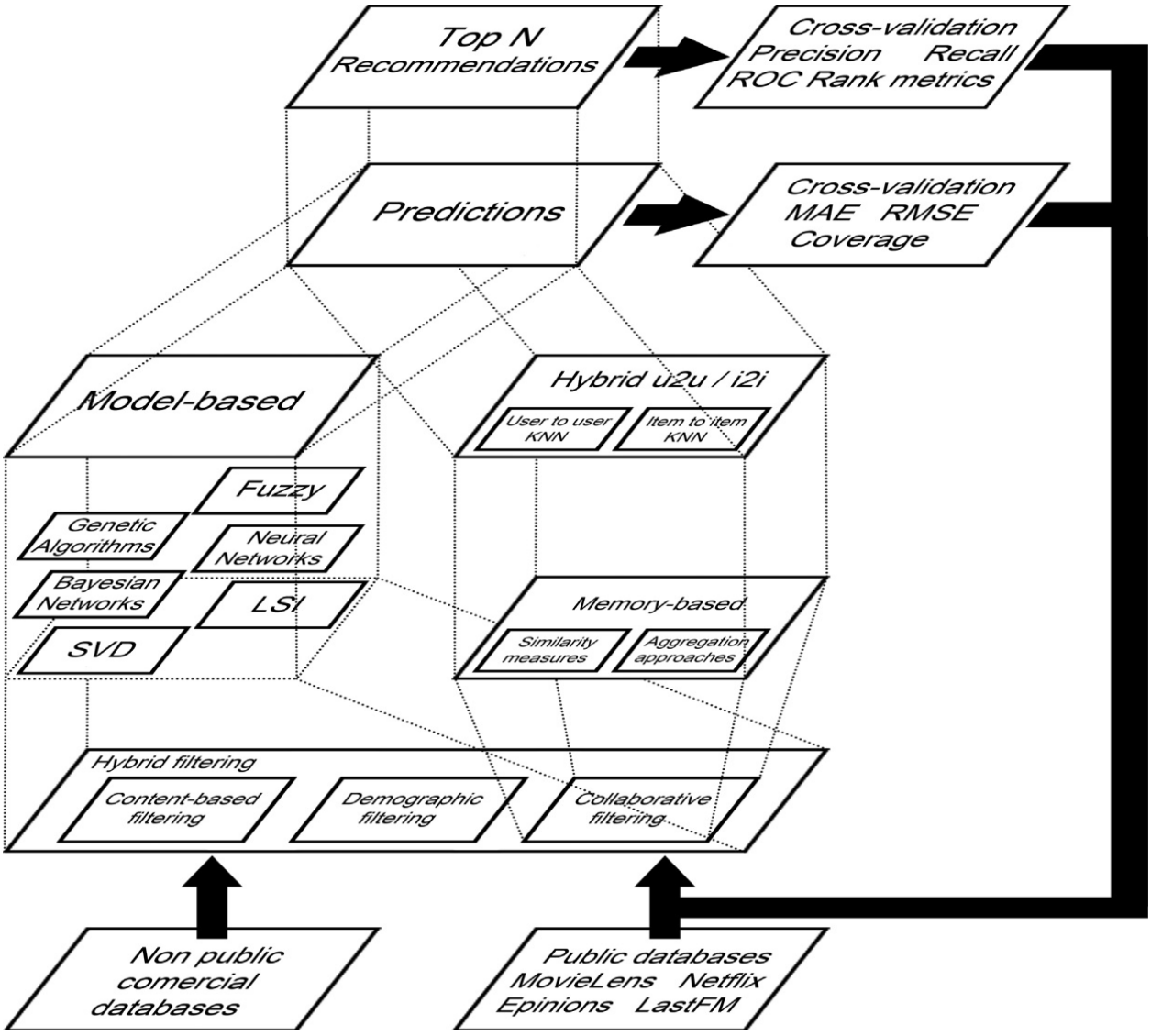

Figure. 2.7. Traditional models of recommendations and their relationships[56]

Model based technique simply try to combine the reduction method singular values decomposition and latentsemantic index.it is understood that singular value decomposition provide good results but they are computationally very expensive, they are mostly used in a place the known preference do not change with time like static

off line setting. Sometimes the recommender system can also use clustering approach to improve the prediction quality and reduce the cold start problem[28]. Some approach may also use clustering for both items and users. Recommender systems which require social information have already used clustering to improve the quality of prediction. Model-based technology like neural network or genetic algorithm makes use of different kind of information. While the typical memory based approach depends on item to item or user to user, etc. although the main aim of both the process is to get the most accurate prediction of the user thought. Normally the methods which are used to measure the accuracy of the prediction through this recommender process is like precision, MAE and recall[29].

2.8 Major Problem Encountered By Recommender System

One of the major problems encountered by most of the recommender system is the cold start problem, means at the starting it is not possible to give a suitable recommendation because we do not have the complete profile of the user that we don't have the knowledge weather the user like certain thing or not. Sometimes, problems also come because of the new item in the list or the new community. The new community problem actually means that when a new recommender system starts working, then it does not have the complete information about either the user or the item, one of the methods maybe be solve the problem is to inspire the user to provide number of ratings of the items it will help us in two ways first it creates a profile of the user that is whether the user like certain thing or not and second we get the routing information of the items [30]. To understand the new item problem in deep we will find when a new item enters in the recommender system they will not have any previous rating, and therefor they will have a very little chance to get recommended for some user. There is only one solution that we must inspire the user to give a rating to the new item so that the chance of their recommendation can also be improved further. Finally, we can understand that the new user problem is the most difficult problem to be taken care of in the recommender system. as in storing user didn't provide any rating, then we did not know much about the user preferences and choice so recommender system cannot provide a good quality personalized recommendation[31].

Some of the authors presented different approaches and techniques to resolve cold start problem. Like long proposed a new king of hybrid approach which sues content based filtering and uses cross level association rules to integrate content information about domain items. Kim in its paper proposed a collaborative tagging so that the filtering of the user preference can be done properly, they also took the advantage of collaborative tagging for cold start problem and data sparseness[32]. Weng on the other hand tried to find the relationship between the additional taxonomies preferences and the user item preferences all this resulting into better recommendation and also solved the cold start problem up to some extent. Leo presented the profile of the recommender system user with the information extracted from some scientific publication.

Martinez represented a hybrid recommender system in which they merged to filtering approaches knowledge based filtering with collaborative filtering. While Chen and he introduced different type of common work which are based on collaborative filtering and related to demographic vector. Saranya proposed a different type of hybrid recommender system that try to use inherent features which can be understood from the item shown in multi-attributed record using a model based on probability. Park also discussed a new approach where they used filterbots and surrogate that give a predicted rating based on item attributes and user. I found out that most common and widely used method to implement collaborative filtering is K Nearest Neighbor algorithm [33]. Commonly this algorithm works in three steps, first it tries to identify the most close k neighboring users based on similarity measure, then it uses an aggregation approach with those ratings of the item which are not rated by the selected user[47].

Finding a logical and efficient measure for assessing the similarity is also an important issue, which can tell similarity between the pair of items or the pair of users. Normally there are two methods used which comparing the rating of two different items given by all the user and second comparing the rating of all the items given by two users.the method which uses the most basic logic for finding the similarity of statistical origin is KNN algorithm.

2.9 Use of Psychographic Factors And Research Motivation

After the systematic literature survey, in which I studies almost all major filtering approaches used, and the basic framework used by most of the recommender system. One major problem which is discussed by most of the papers is the cold start problem. A number of researchers have tried to combine different filtering approaches and sometimes used other methods to reduce the effect of the cold start problem. Another issue is now days most of the mobile phone users a using smart phones and their location now can be used as a feature can be used to predict better recommendation.

But now days most of the mobile phone users are using android platform and the development of the platform is open and anybody can develop their own application and make it public. Most of the application developing companies are now days using the psychographic factors to give better recommendation, but as far as research field is concerned till now no major work is done in the field.

So I tried to use the psychographic factors to give better and more relevant recommendation by using the psychographic attributes of the mobile phone user like number of call done in previous times, number of messages send or received by the user, the total amount of time for which the mobile screen was on and total internet gprs data used the mobile user. Using these attributes we we predict the behavior of the user like weather he or she is more social or introvert based on which we can assume weather should we give more priority to personal features or to the community feature.

CHAPTER 3

Proposed System Architecture

In the system we used the data set of the world's most popular food portal four squares. This is one of the most famous databases of the total food locations around the world. It covers not only the main food locations, but also the small and little food corners. Here we also implemented our logic in mobile app ‘chatora’. The major problem which we have tried to solve here is the cold start problem and secondly to make the system more efficient I have combined psychographic factors, along with other factors. So here when a user installs the mobile app of the recommender system first the application will ask the permission to access the pattern of usage of different things like number of calls, number of messages, total time for which the mobile screen was on and the total internet data used by the system. When the user first login on the system, there will be two navigation paths.If you have already a four square account, then the system will fetch all your details from the four square accounts, if you do not have an account then system will ask your basic information like name, age, sex, location etc.

After taking the data from the user then the system applies all the filters like location, demographic, personal and psychographic.Basically when a user log on using the mobile application first encounter the start screen, then, according to the architecture, there two navigation possibilities, first log in with four square if the user has already had a foursquare application. Otherwise can directly enter the details on the chatora interface. In the first case if the user enters through the foursquare interface, then all the information about the user will be taken directly from the foursquare data base after the permission of the user. But if the userdoes not have a four square application, then the userhas to enter all the details separately through the chatora interface. After the process of taking the information about the user, now system try to fetch the data set from the foursquare data base of the restaurant, then the process of filtering is applied to the data set where we take care of the psychographic factors to give the refined set of recommendation.

In the process of recommendation, we have taken some assumption based on a proper literature survey like if the user has already been to the restaurant, then we will give a bump up in value by 2 units and if the user have already liked the restaurant the it will give a bump up value by 3. Some other consideration is also taken like if the gender of the user is a girl then we will a bump up of 2 units and is male than only 1 unit. Because it is basically learned that girls are more foodie than boys. Some other

assumptions are also taking like an age for a valid prediction must be between 15 and 40, the system will treat the user income above the threshold of 30 lakhs. These assumptions are taken to make the computation less complex and easy to understand.

3.1 Filtering Process

Now will understand the process of filtering in deep and also examine how psychographic factors can improve the system prediction by understanding the mobile phone usage pattern of the user. After the mobile application is stored on the user smart phone the application automatically fetch all the required data from the smart phone, then one by one it uses data gained from the user smart phone to calculate the value of the psychographic factor 'X' here system will use four different factors to calculate the value of the psychographicfactor ('X").

$$(\text{Psychographic Factor})X_{(0<x\leq 0.9)} = \frac{\text{number of calls per day}}{10} + \frac{\text{screen usages per day in minutes}}{180}$$

$$+\frac{\text{mobile internet data in MB}}{512} + \frac{\text{message per day}}{15} \qquad (1)$$

Using these values system tries to understand how social the user is, that is if the user is more social that it is common that he or she will make number of calls, send or receive more messages, will use more internet data.

3.2 Derived Equation

After calculating the psychographic factor, the system finds out two different values first is the collaborative score and second is the demographic score. Demographic score basically uses the personal attributes of the user like income, age and gender and previous behavior of the user like number of past visits in the restaurant and whether the restaurant was previously like by the user or not.

$$\text{Demographic Score} = \frac{\text{income per year in lacks}}{5} + \frac{25}{\text{age}} + \text{Previous chaek in value}$$

$$\text{Previous like value} + \text{Gender value} \qquad (2)$$

Than system takes care of some social attributes like what is the current community rating of the restaurants and how many numbers of previous chicken in the restaurants this factor is called

$$\text{Collaborative Score} = \frac{\text{community rating}}{10} \times \frac{\text{no of check in}}{100} \times 5 \quad (3)$$

We then apply the most important step of this hybrid filtering approach which also consider psychographic factor. Depending upon the score of the psychographic attributes system will give different wattage to different types of factors, that is if the value of the psychographic factor is higher than more importance will be given to the collaborative score as the user is more social so he or she will more depend on the social attributes if the user's psychographic score is less than it means the user is less social than system gives more importance to personal attributes. All the logics are will understood by the system in the terms of equations to calculate the net score of the user.

$$\text{Net Score} = (x * \text{collaborative score}) + ((1-x) * \text{demographic score}) \quad (4)$$

$$\text{Final recommendation score (frs)} = 50 \times \text{scaled adjustment factor} \quad (5)$$

$$\text{Scaled adjustment factor} = \text{net score} \div 10 \quad (6)$$

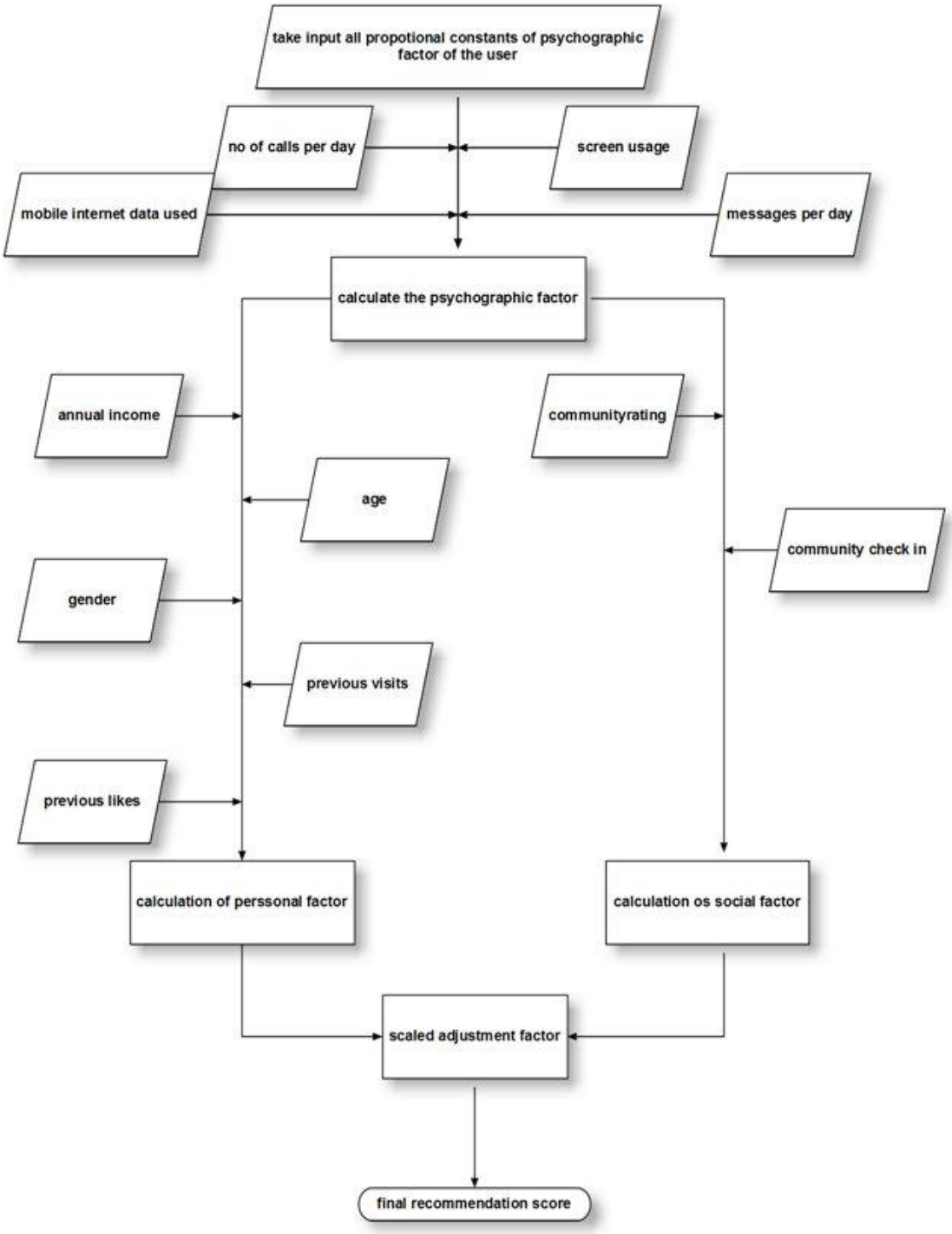

figure 3.1 Chatora(system) architecture

value of x comes more then .9 then consider it as .9.

Gender value = 1. If male

2. If female

If previous check in there then value is 2, otherwise 0.

If previous like is there then value is 3, otherwise 0.

Income per year cannot be more then 30.

Age cannot be less than 15 and greater than 40.

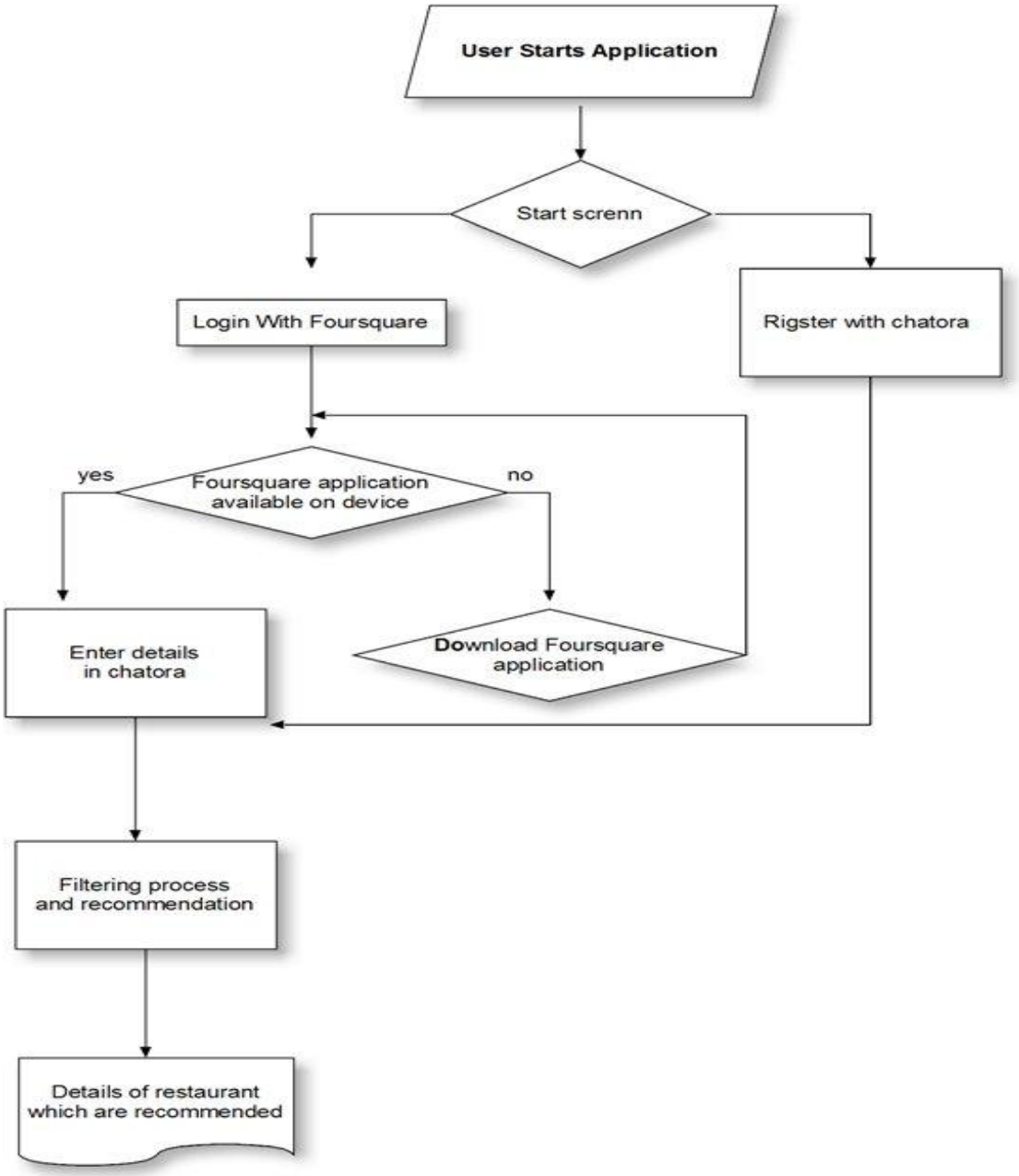

Figure 3.2 Flow control graph of system

CHAPTER 4

RESULTS

As we have used the database of four square web site with the help of application program interface provided by them and also used the psychographic factors along with demographic, location and personal attributes of the user. By using the psychographic factor we got success in recommending a better and more confined recommendation to the user, as by using psychographic attributes we can predict whethera user is more social or introvert based on which we can give different priority to different attributes which we are using in our filtering process. The most popular and orthodox problem which we tried to solve is the cold start problem by which most of the recommender systems suffer by gaining the data of users previous behaviors from his or her smart phone. As whenever either a new item or a new user is entered into a database, recommendation system could have not recommended them because they do not have data either about the item or about the user.

Because of psychographic factor consideration we can understand that the recommendation given by our recommender system 'chatora' will give better and more appropriate recommendation to the user, as we have shown that how close the attributes value of the restaurants which are in the top among the recommended list.Here we have used the most basic method to find the closeness of the items in a cluster, first by using the community rating and number of check in of the top five recommendations given by four square first with the system second without the system. In order to compare the performance the newly designed system we here first consider the top five recommendation without considering the new algorithm then we consider two attributes community rating and previous check-in the restaurant and then computes the mean of the top five recommendation, then in order to compute the variance we first computed the standard deviation and then computed the variance, the smaller variance value shows the cluster of the recommender value is very close to each other that is cluster is very compact, but if the variance value is large then it means that the values are not close to each other and cluster will not predict the good quality recommendation.

Here we have used the most popular method which is used in mathematics to find the closeness of different items in a cluster of recommended to the user is standard deviation, we have used an attributes community rating and number of checks in of the top five recommendations given by four square first without using psychographic factors and then with psychographic factors. In order to compare the performance of our newly designed system 'chatora' first we calculated the arithmetic mean of the attributes value of the restaurants which are in the top recommended list. Then in order to calculate the standard deviation using root mean square method, first we calculate the square of the difference of every individual value with the arithmetic mean, then again we took the average of the squared errors and then we took the square root of the final value, finally resulting into the standard deviation. The smaller standard deviation value shows the cluster of the recommender value is very close to

each other that is cluster is very confine, which that our hybrid filtering scheme is very impressive, but if the variance value is large then it means that the values are not close to each other and cluster will not depict the good quality recommendation.

Arithmetic mean (AM)

The *arithmetic mean* (or simply "mean") of a sample space $x_1, x_2, \ldots, x_n$ is the sum of the total sampled values divided by the total number of items in the sample:

$$\bar{x} = \frac{x_1 + x_2 + \cdots + x_n}{n}$$

For example, the arithmetic mean of five values: 4, 36, 45, 50, 75 is

$$\frac{4 + 36 + 45 + 50 + 75}{5} = \frac{210}{5} = 42.$$

After calculating the arithmetic mean of the given attribute we calculate the standard deviation of all the values with the mean value.

For a finite set of numbers, the standard deviation is found by taking the square root of the average of the squared differences of the values from their average value. For example, consider a **p**opulation consisting of the following eight values:

$$2,\ 4,\ 4,\ 4,\ 5,\ 5,\ 7,\ 9.$$

These eight data points have the mean (average) of 5:

$$\frac{2 + 4 + 4 + 4 + 5 + 5 + 7 + 9}{8} = 5.$$

First, calculate the difference of each data point from the mean, and square the result of each:

$$\begin{aligned} (2-5)^2 &= (-3)^2 = 9 & \qquad (5-5)^2 &= 0^2 = 0 \\ (4-5)^2 &= (-1)^2 = 1 & \qquad (5-5)^2 &= 0^2 = 0 \\ (4-5)^2 &= (-1)^2 = 1 & \qquad (7-5)^2 &= 2^2 = 4 \\ (4-5)^2 &= (-1)^2 = 1 & \qquad (9-5)^2 &= 4^2 = 16. \end{aligned}$$

Next, calculate the mean of these values, and take the square root:

$$\sqrt{\frac{9+1+1+1+0+0+4+16}{8}} = 2.$$

So this standard deviation metrics help us to find that how close the values of the recommended restaurants are with each other.

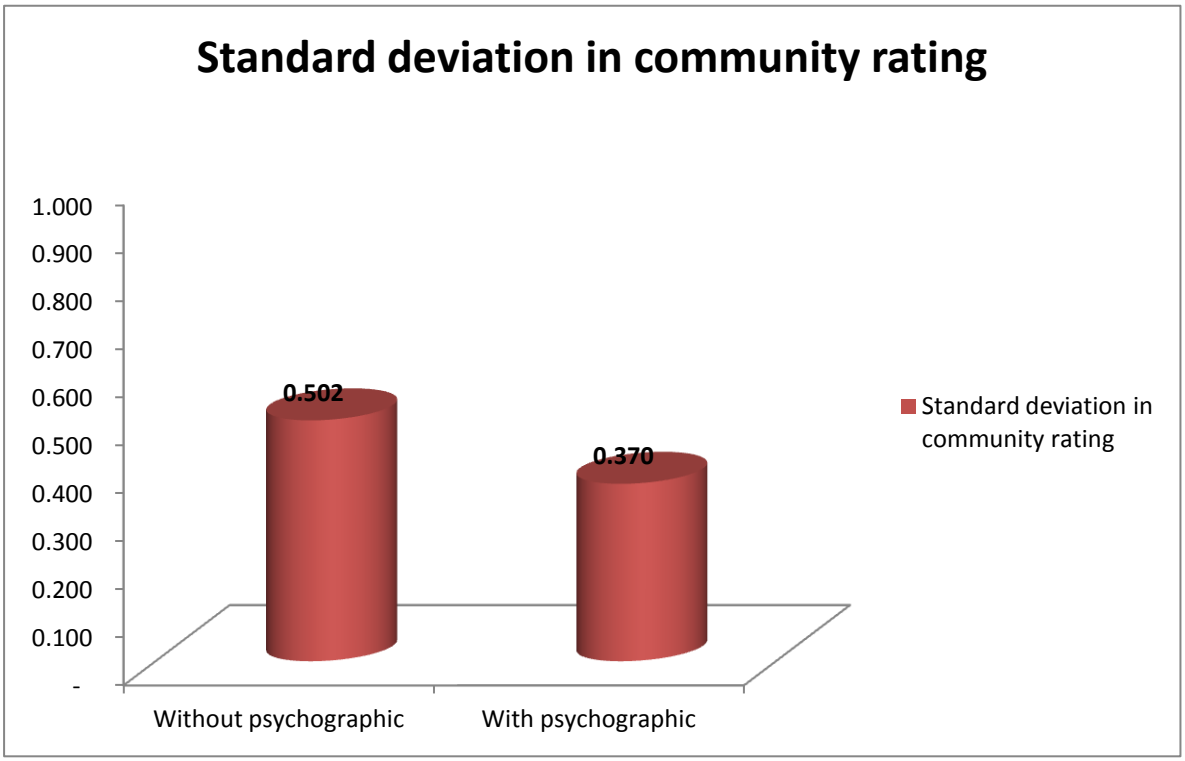

Figure 4.1 Comparisons of performance with and without psychographic in community rating

Because of psychographic factor consideration we can understand that the recommendation given by this recommender system will give better and more appropriate results to the user here we have user the variancemetrics to show that how close the recommended list of restaurants to each other in properties with the help of which we have done the filtering.

	variance	Standard deviation	Root mean square value
With psychographic factor	0.1367	.3697	8.0885

Without psychographic factor	0.252	.5019	8.1133

Table 4.1 Comparing community rating values with different mathematical function of with and without psychographic factors

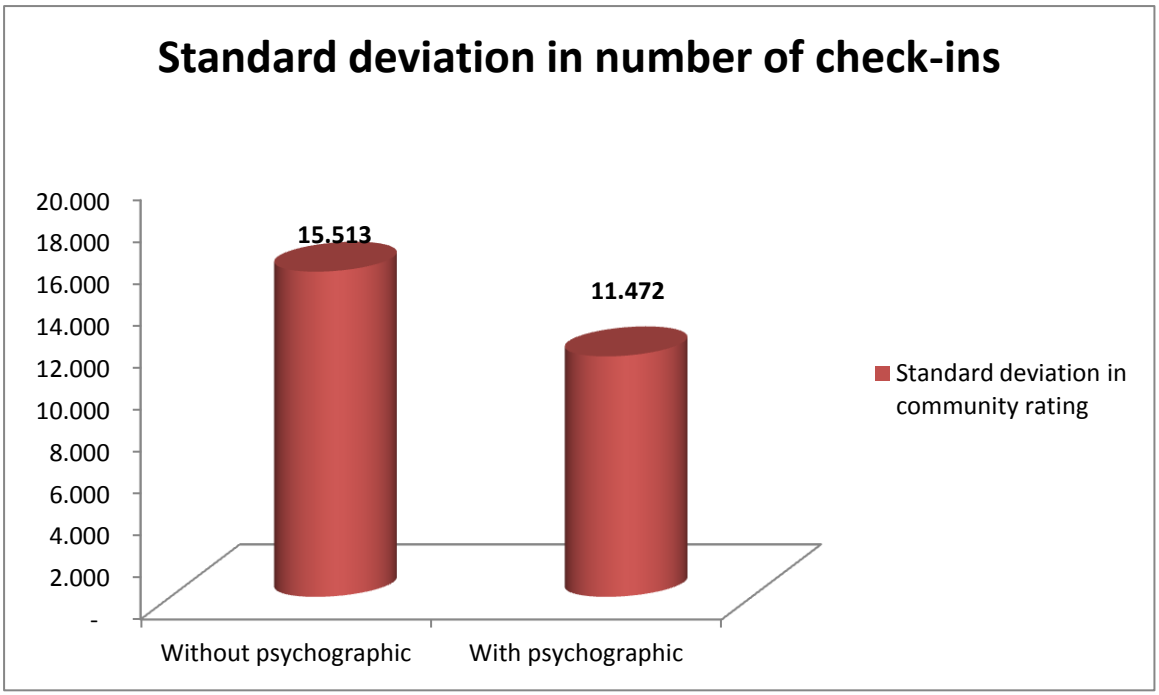

Figure4.2 Comparisons of performance with and without psychographic in check-in

Here we have used the most basic method to find the closeness of the items in a cluster, first by using the community rating and number of check in of the top five recommendationsgiven by four square first with the system second without the system.

	variance	Standard deviation	Root mean square value
With psychographic factor	131.6	11.4717	109.602
Without psychographic factor	240.64	15.5125	112.4748

Table 4.2 Comparing checkin values with different mathematical functions of with and without psychographic factors

Root mean square value is a important function used in statistics in which first we square all the individual values and then we compute their mean and final the square root of the resultant values is taken. RMS can depict very significantly that how close the values are with each other in a cluster.

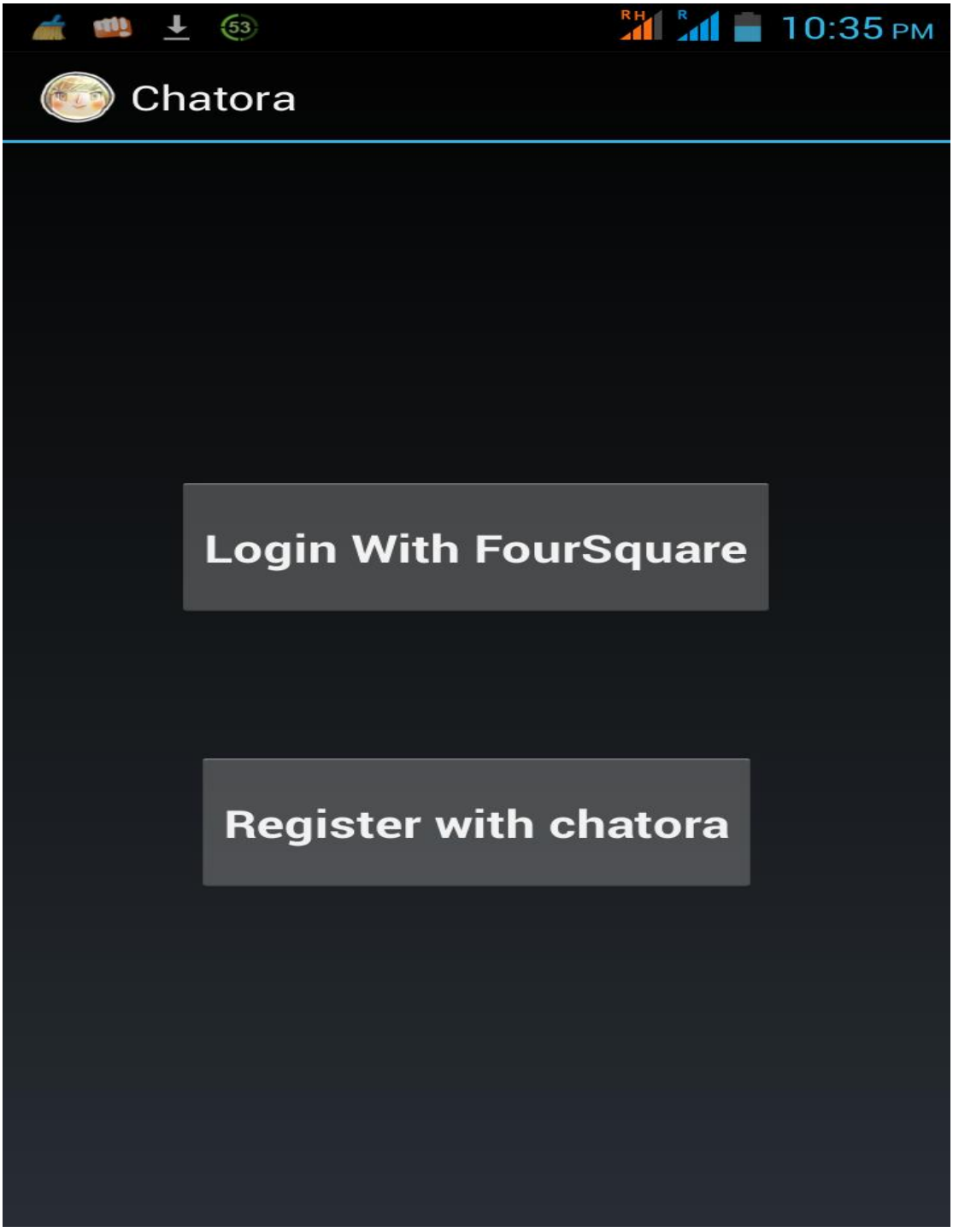

Figure 4.3 Main screen of application

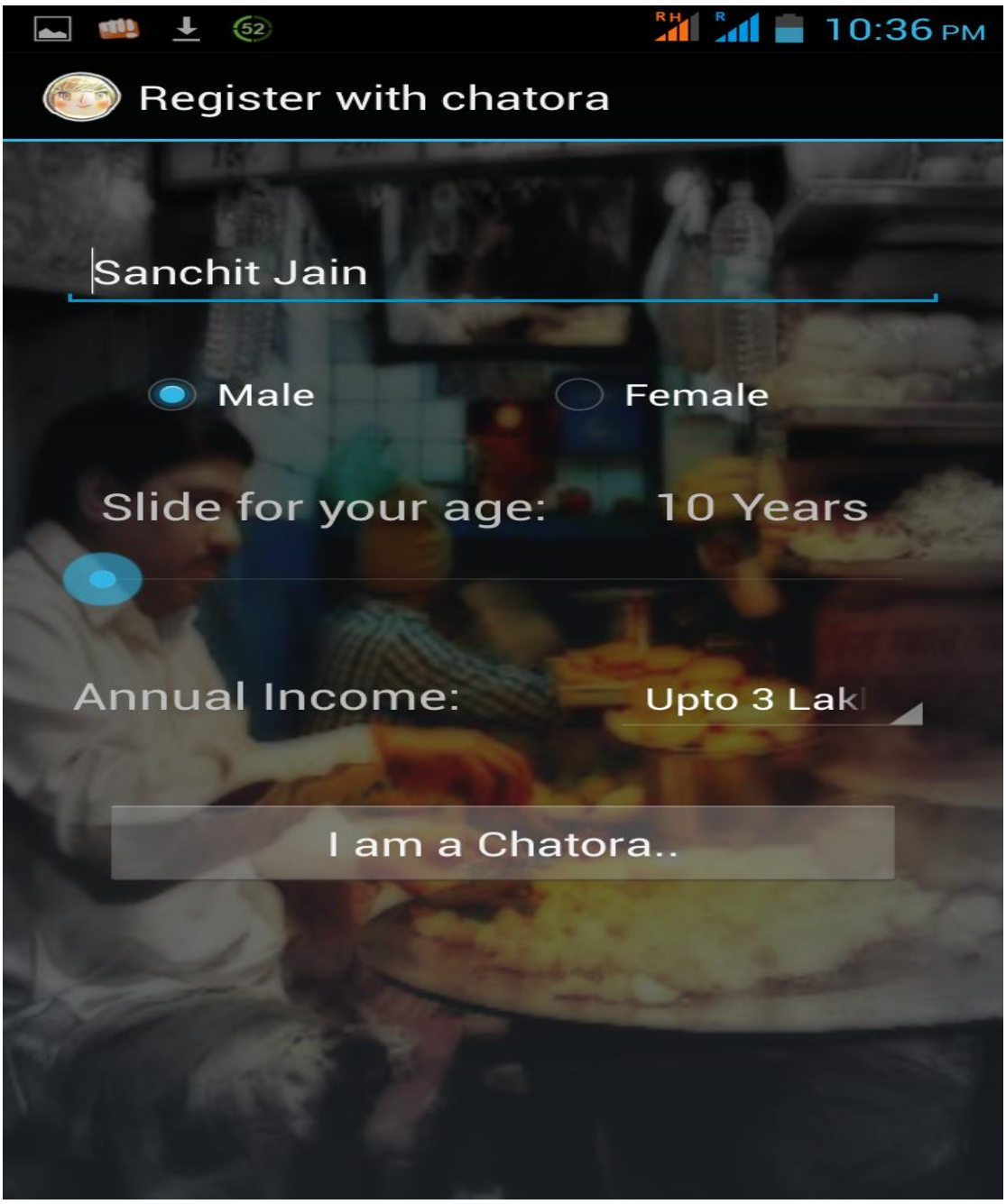

Figure 4.4 Chatora registration screen

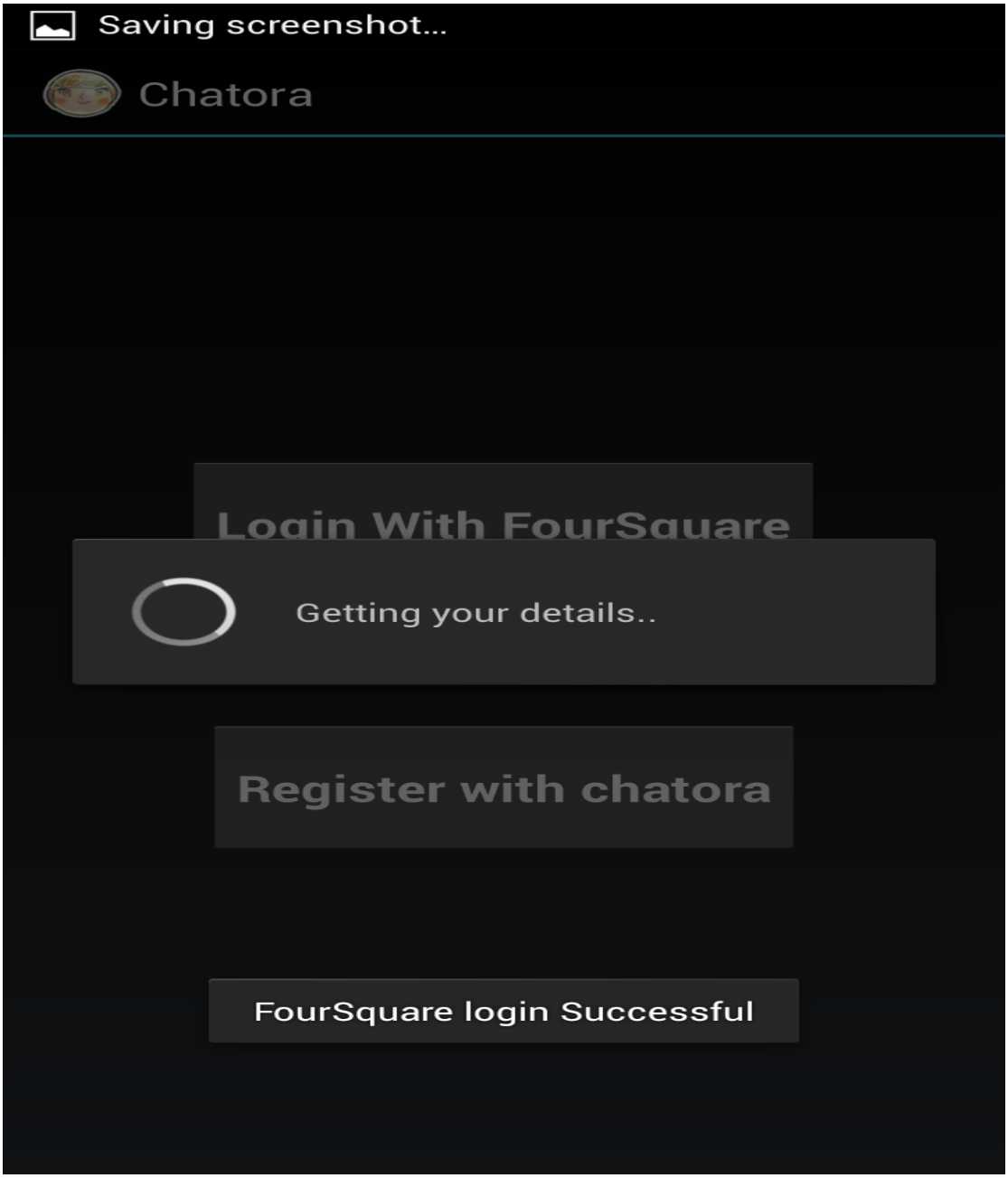

Figure 4.5 System fatching details from foursquare

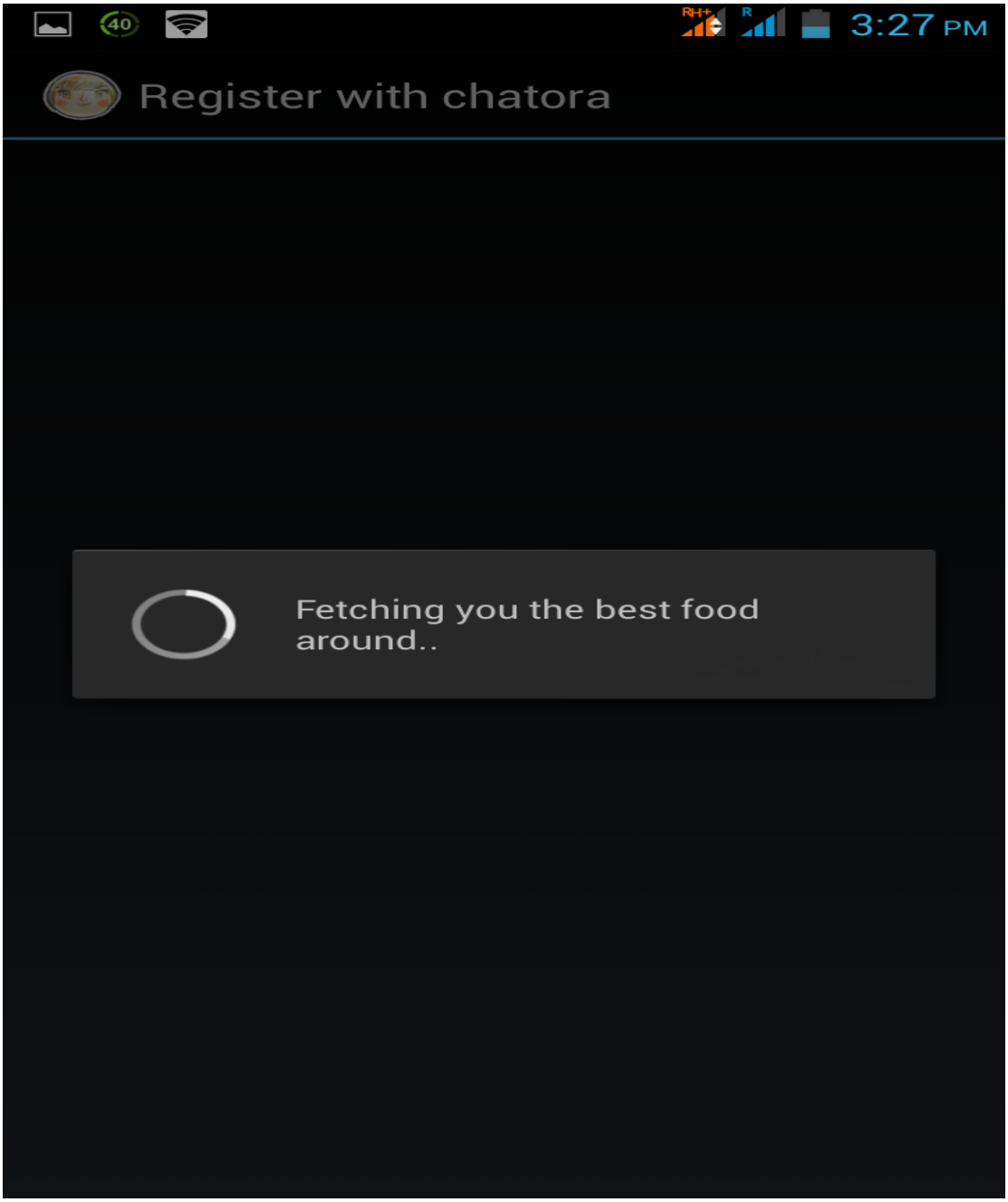

Figure 4.6 Appling the filtering process

Figure.4.7 final suggestion of restaurants

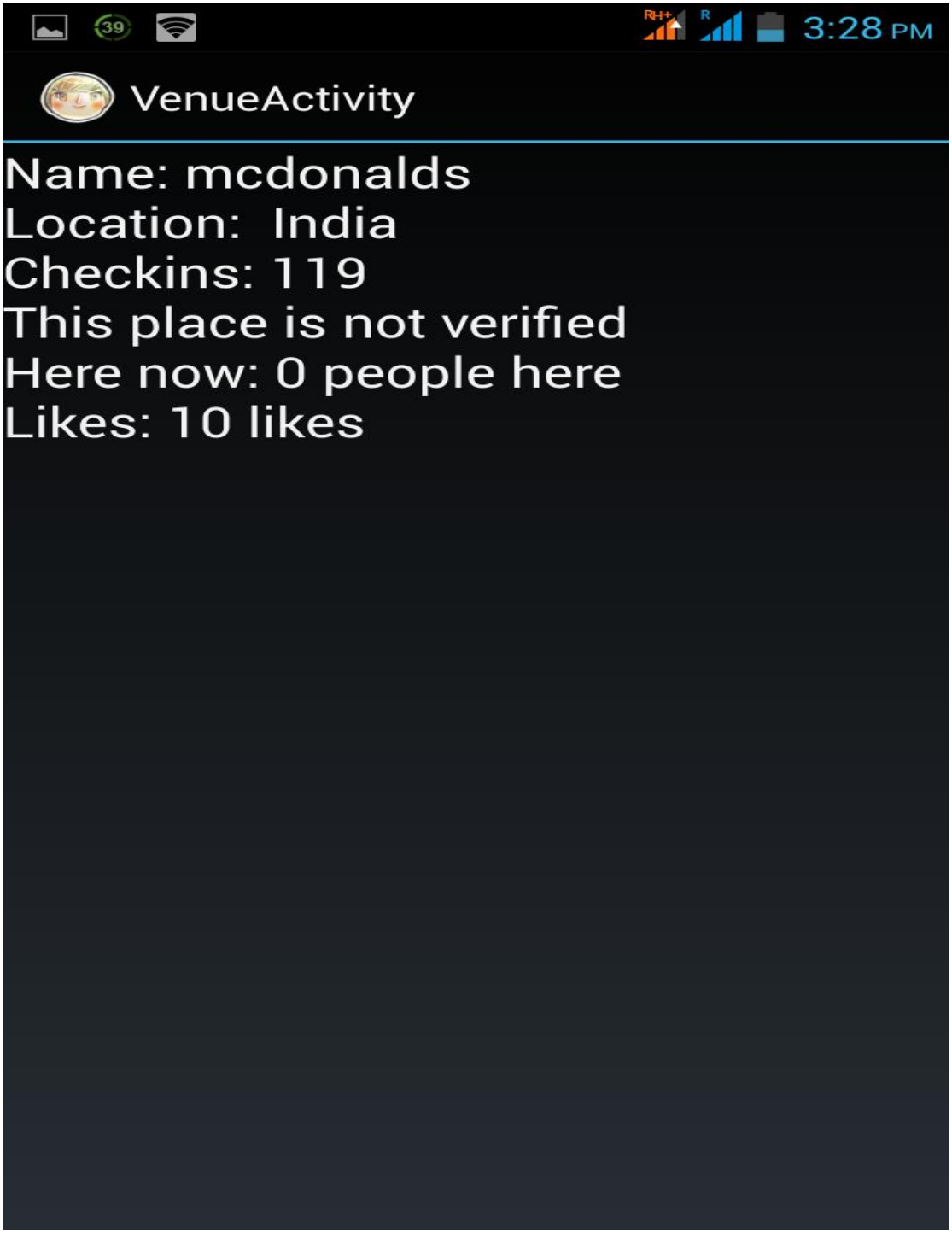

Figure 4.8 details of the selected restaurant

CHAPTER 5

CONCLUSION AND FUTURE WORK

Conclusion and Future Work

Now, after describing and defining the mechanism of the system we can understand that the better recommendation can be provided to the user if we also take care the psychographic factors. Recommender system is very important nowadays as because of the vast amount of data present on the net, it becomes quite difficult for the user to get a confine and exact data which are required. As a recommender system reduces this problem up to a large extent by recommending a better set of data, which is more confined and relevant.

But there are many issues that remain in the process even if we uses current recommender systems, one of the major issues is a cold start problem that is whenever a new item or user enter in the system it becomes very difficult for the recommender system to recommend more relevant and logical data. As most of the mobile phone users are using smart phones now a days. Most of the recommender system uses a feature only like the location of the users and other demographic factors like age, emotions and previous usage behavior. But this hybrid recommender system chatora uses the psychographic factors like which can be taken from the mobile phone of the user like number of calls a user make or attend daily, total time for which the mobile screen was active, total internet data which the user has used and number of messages send or received . All thesefactors we use in our system to predict the behavior of the user in future. So if all these factorsgive high rating, then it means the user is more social and more emphasize should be given to the social factors like community rating given to the restaurant and number of checking of the customers.

If after analyzing psychographic factors is it understood that the user is not social then more emphasize should be given to its personal attributes like income, age, previous visit and gender. So here we reduced the cold start problem up to the extent that is, even if in starting if do not have data about a item or a user still system should be able to recommend appropriate, but for this we must know which factors must be taken into consideration to make an efficient recommendation. Psychographic factor which is actually the behavior pattern of the user help us find the correct the factor on which they emphasize must be given in order to give a better recommendation from the first time.no recommendation system was using factor before.

Form the upcoming researcher in the field we expect from the scientific community that further explore this psychographic factor and try to fetch more information from the behavior of the user so instead of using only four factors we can take a number of other factors in order to identify the psychographic factor of the user in a better way. Some other attributes could also be taken like amount of time a user spend of Facebook or what's app, total number of application installed on the smart phone, monthly bill of the user and many others. These results can further be refined by applying some optimization logic.

REFERENCES

[1].DeukHee Park, HyeaKyeong Kim, Il Young Choi, Jae Kyeong Kim. A literature review and classification of recommender systems research". Expert Systems with Applications 39 (2012) 10059–10072

[2]. Frias-Martinez, E., Chen, S. Y., & Liu, X. (2009). Evaluation of a personalized digitallibrary based on cognitive styles: Adaptivity vs. adaptability. International Journal of Information Management, 29, 48–56.

[3]. Frias-Martinez, E., Magoulas, G., Chen, S. Y., &Macredie, R. (2006). Automated usermodeling for personalized digital libraries. International Journal of InformationManagement, 26, 234–248.

[4]. Claypool, M., Gokhale, A., Miranda, T., Murnikov, P., Netes, D., &Sartin, M. (1999).Combining content-based and collaborative filters in an online newspaper,Proceedings of the ACM SIGIR'99 Workshop on Recommender Systems.

[5]. Basu, C., Hirsh, H., & Cohen, W. (1998). Recommendation as classification :Usingsocial and content-based information in recommendation, In Proceedings of the15th National Conference on, Artificial Intelligence, 714–720.

[6]. Blanco-Fernandez, Y., Pazos-arias, J. J., Gil-Solla, A., Ramos-Cabrer, M., & Lopez-Nores, M. (2008). Providing entertainment by content-based filtering andsemantic reasoning in intelligent recommender systems.IEEE Transactions onConsumer Electronics, 54, 727–735.

[7]. Schafer, J. B., Joseph, A., & Riedl, J. (2001). E-commerce recommendationapplications. Data Mining and Knowledge Discovery, 5, 115–153.

[8]. Berry, M. J. A., & Linoff, J. S. (2004). Data Mining Techniques for Marketing. Sales andCustomer Relationship Management (2nd ed.,). Wiley.

[9]. Cho, Y. H., Kim, J. K., & Kim, S. H. (2002). A personalized recommender system basedon web usage mining and decision tree induction.Expert Systems withApplications, 23, 329–342.

[10]. Cho, Y. B., Cho, Y. H., & Kim, S. H. (2005). Mining changes in customer buyingbehavior for collaborative recommendations. Expert Systems with Applications,28, 359–369.

[11]. Kim, H. K., Kim, J. K., &Ryu, Y. U. (2009). Personalized recommendation over acustomer network for ubiquitous shopping. IEEE Transactions on ServicesComputing, 2, 140–151.

[12]. Ibnkahla, M. (2000). Applications of neural networks to digital communications-asurvey. Expert Systems with Applications, 80, 1185–1215.

[13].Anders, U., &Korn, O. (1999). Model Selection in Neural Networks. Neural Networks,12, 309–323.

[14].Cai, D., He, X., Wen, J.R., & Ma, W.Y. (2004). Block-level link analysis, Proceedings ofthe 27th annual international ACM SIGIR conference on Research and developmentin, information retrieval, 440–447.

[15]. Malhotra, N. K. (2007). Marketing research: An applied orientation (5th ed.). PearsonEducation Inc.

[16].M. Pazzani, A framework for collaborative, content-based, and demographic filtering, Artificial Intelligence Review-Special Issue on Data Mining on the Internet 13 (5-6) (1999) 393–408

[17]. J.L. Herlocker, J.A. Konstan, J.T. Riedl, L.G. Terveen, Evaluating collaborative filtering recommender systems, ACM Transactions on Information Systems 22 (1) (2004) 5–53.

[18]. R. Burke, Hybrid recommender systems: survey and experiments, User Modeling and User-Adapted Interaction 12 (4) (2002) 331–370.

[19]. Yoshii, K., Goto, M., Komatani, K., Ogata, T., &Okuno, H. G. (2008). An efficienthybrid music recommender system using an incrementally trainableprobabilistic generative model.IEEE Transactions on Audio, Speech, andLanguage Processing, 16, 435–447.

[20].Symeonidis, P., Nanopoulos, A., &Manolopoulos, Y. (2010). A unified framework forproviding recommendations in social tagging systems based on ternarysemantic analysis.IEEE Transactions on Knowledge and Data Engineering, 22,179–192.

[21]. Ochi, P., Rao, S., Takayama, L., &Nass, C. (2010). Predictors of user perceptions ofweb recommender systems: How the basis for generating experience andsearch product recommendations affect user response. International Journal ofHuman – Computer Studies, 68, 472–482.

[22]. McGinty, L., & Smyth, B. (2006). Adaptive selection: An analysis of critiquing andpreference-based feedback in conversational recommender systems.International Journal of Electronic Commerce, 11, 35–57.

[23].Nanopoulos, A., Rafailidis, D., Symeonidis, P., &Manolopoulos, Y. (2010). MusicBox:personalized music recommendation based on cubic Analysis of social tags. IEEETransactions on Audio, Speech, and Language Processing, 18, 407–412.

[24]. A. Papadimitriou, P. Symeonidid, Y. Manolopoulus, A generalized taxonomyof explanations styles for traditional and social recommender systems, DataMinning Knowledge Discovery 24 (3) (2012) 555–583.

[25].T.H. Roh, K.J. Oh, I. Han, The collaborative filtering recommendation based on SOM cluster-indexing CBR, Expert Systems with Applications 25 (2003) 413–423.

[26].M.Pazzani, A framework for collaborative, content-based, and demographicfiltering, Artificial Intelligence Review-Special Issue on Data Mining on theInternet 13 (5-6) (1999) 393–408.

[27]. S.B. Cho, J.H. Hong, M.H. Park, Location-based recommendation system usingBayesian user's preference model in mobile devices, Lecture Notes inComputer Science 4611 (2007) 1130–1139.

[28]. J. Wang, A.P. Vries, M.J. Reinders, Unified relevance models for ratingprediction in collaborative filtering, ACM Transactions in InformationSystems 26 (3) (2008) 1–42.

[29]. Z. Yu, X. Zhou, Y. Hao, J. Gu, TV program recommendation for multipleviewers based on user profile merging, User Modeling and User-AdaptedInteraction 16 (1) (2006) 63–82.

[30].C. Porcel, J.M. Moreno, E. Herrera-Viedma, A multi-disciplinarrecommendersystem to advice research resources in university digital libraries, ExpertSystems with Applications 36 (10) (2009) 12520–12528.

[31]. G. Pitsilis, X. Zhang, W. Wang, Clustering recommenders in collaborativefiltering using explicit trust information, Advances in Information andCommunication Technology 358 (2011) 82–97.

[32]. M.C. Pham, Y. Cao, R. Klamma, M. Jarke, A clustering approach forcollaborative filtering recommendation using social network analysis,Journal of Universal Computer Science 17 (4) (2011) 583–604.

[33].L. Martínez, R.M. Rodríguez and M. Espinilla. REJA: AGeoreferenced Hybrid Recommender System ForRestaurants.InWebIntelligence and Intelligent Agent Technologies, 2009.

[34].Fan Yang and Zhi-Mei Wang. A Mobile Location-based InformationRecommendation System Based on GPS and WEB2.0 Services. InWSEAS Transactions on Computers, 2009, Pages 725-734.

[35]. Fan Yang and Zhi-Mei Wang. A Mobile Location-based InformationRecommendation System Based on GPS and WEB2.0 Services. InWSEAS Transactions on Computers, 2009, Pages 725-734.

[36].Raymond J. Mooney and Loriene Roy. Content-Based Book Recommending Using Learning for Text Categorization.InProceedings of the SIGIR-99 Workshop on Recommender Systems:Algorithms and Evaluation, Berkeley, CA, August 1999.

[37].Wan-Shiou Yang, Hung-Chi Cheng and Jia-Ben Dia. A locationawarerecommender system for mobile shopping environments.InExpert Systems with Applications 34, 2008.

[38]. W S Yang, H C Cheng, J B Dia. A location-aware recommender system for mobile shopping environments [J]. Expert Systems with Applications, 2008, 34(1): 437-445.

[39]. M Okabe, A Yoshioka, K Kobayashi. Organizational knowledge transfer of intelligence skill using ontologies and a rule-based system [J]. IEICE Transactions on Information and Systems, 2010, 93 (4): 763-773.

[40]. M J Barranco, L Martinez. A method for weighting multi-value features in content-based filtering [C]// Proceedings of the 23rd International Conference on Industrial Engineering and Other Applications of Applied Intelligent Systems, Cordoba, SPAIN, 2010: 411-418.

[41]. T H. Dao, S R Jeong, H Ahn. A novel recommendation model of location-based advertising: Context-aware bollaborative filtering using GA approach [J]. Expert Systems with Applications, 2012, 39: 3731-3739.

[42]. A. T. Fiore, L. S. Taylor, X. Zhong, G. A. Mendelsohn, and C. Cheshire,"Who's right and who writes: People, profiles, contacts, and replies inonline dating.," in HlCSS, pp. 1-10, IEEE Computer Society, 20 I O.

[43]. A. Zenebe and A. F. Norcio, "Representation, similarity measures andaggregation methods using fuzzy sets for content-based recommendersystems.," Fuzzy Sets and Systems, vol. 160, no. I, pp. 76- 94, 2009.

[44]. L. Li and T. Li, "Meet: a generalized framework for reciprocalrecommendersystems.," in CIKM (X. wen Chen, G. Lebanon, H. Wang,and M. J. Zaki, eds.), pp. 35-44, ACM, 2012.

[45]. R. Hu and P. Pu, “Acceptance issues of personality-based recommender systems.” Proc. of the third ACM conf. on Recommender systems, pp.221–224, 2009.

[46]. Choi, S. H., Kang, S. M., &Jeon, Y. J. (2006). Personalized recommendation systembased on product specification values.Expert Systems with Applications, 31,607–616.

[47]. Han, P., Xie, B., Yang, F., &Shen, R. (2004). A scalable P2P recommender systembased on distributed collaborative filtering.Expert Systems with Applications, 27,203–210.

[48]. Ant Ozok, A., Quyin Fan &Norcio, Anthony F. (2004). Design guidelines for effectiverecommender system interfaces based on a usability criteria conceptual model:results from a college student population. Behaviour& Information Technology.

[49]. Cheung, K. W., Tsui, K. C., & Liu, J. (2004). Extended latent class models forcollaborative recommendation. IEEE Transactions on Systems, Man andCybernetics, Part A: Systems and Humans, 34, 143–148.

[50]. Robillard, M. P., &Dagenais, B. (2009). Recommending change clusters to supportsoftware investigation: an empirical study. Journal of software maintenance, 22,143–164.

[51]. K. Mcnally, M.P. O'mahony, M. Coyle, P. Briggs, B. Smyth, A case study ofcollaboration and reputation in social web search, ACM Transactions onIntelligent Systems and Technology 3 (1) (2011). Article 4.

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

[53]http://images.google.co.in/

[54] http://en.wikipedia.org/wiki/Web_mining

[55] http://en.wikipedia.org/wiki/Collaborative_filtering

[56]J. Bobadilla, F. Ortega, A. Hernando, A. Gutierrez,Recommender systems survey, Knowledge-Based Systems, 46 2013 (109-132)