DESIGN AND DEVELOPMENT OF SOCIAL COMMUNITY FRAMEWORK USING DATA MINING

Submitted in partial fulfilment of the requirements for the degree of

Doctor of Philosophy

In

Department of Computer Science & Engineering

By

Sachin Papneja
2K11/Ph.D./CO/11

Under the Guidance of

Prof. Kapil Sharma
Professor & Head of Department
Department of Computer Science & Engineering
Delhi Technological University, Delhi - 110042, India

Dr. Nitesh Khilwani(co-supervisor)
Technical Architect
Round Glass, Noida

Delhi Technological University,
Shahbad Daulatpur, Main Bawana Road,
Delhi - 110042, India.
2020

Contents

CERTIFICATE

DELHI TECHNOLOGICAL UNIVERSITY (Govt. of National Capital Territory of Delhi)
BAWANAROAD, DELHI-110042

Date:____________________

This is to certify that the work embodied in the thesis titled “Design and Development of Social Community Framework Using Data Mining” has been completed by Sachin Papneja under my supervision and guidance towards fulfillment of the requirements for the degree of Doctor of Philosophy of Delhi Technological University, Delhi. This work is based on original research and has not been submitted in full or in part for any other diploma or degree of any university.

Prof. Kapil Sharma	**Dr. Nitesh Khilwani (co-supervisor)**
Professor & Head of Department,	Technical Architect
Department of Information Technology,	Round Glass, Noida
Delhi Technological University,	
Delhi-110042.	

DECLARATION

I Sachin Papneja, Ph.D. student (RollNo.2K11/Ph.D./CO/11), hereby declare that the thesis entitled **"Design and Development of Social Community Framework Using Data Mining"** which is submitting for the award of the degree of Doctor of Philosophy in Computer Science & Engineering, is a record of Bona fide research work carried out by me in the Department of Computer Science & Engineering, Delhi Technological University. I further declare that his work is based on original research and has not submitted to any university or institution for any degree or diploma.

Date:

Place: New Delhi

SACHIN PAPNEJA

2K11/Ph.D./CO/11

Department of Computer Science & Engineering,

Delhi Technological University,

Delhi-110042.

ACKNOWLEDGMENT

First and foremost, I would like to thanks Prof. Kapil Sharma, who has always extended unconditional support and guidance in carrying out my research work. He has always been very co-operative and understanding, and I am thankful for his constant support and encouragement throughout this work.

I would now wish to extend my deep sense of gratitude to Dr. Nitesh Khilwani, who has been instrumental in shaping my research and who has guided me from time to time by giving his valuable inputs and suggestions which helped me in continuing with my work smoothly. Whenever I was in stress or found myself demotivated, he always gave courage and motivated me and brought the best out of me. Besides being my mentor, he has been a pillar of strength, and I am deeply indebted to him for guiding me with his extensive knowledge and insightful discussions.

I would also like to thank all my teachers and friends for their direct and indirect help and support in this work.

Lastly, I would like to thank my parents, wife and daughter, who has been very loving and supportive throughout the research period.

Sachin Papneja

ABSTRACT

With the dawn of the living web 5.0 and Ontological semantic networks, open source social interaction Platform popularity and dependence has gained attention of researchers for both industry and academician. This new area of research focuses on social behavior of netizens. A digital avatar along with net information about user's choices like membership of sports group, financial group, political or entertainment society, video gaming society etc. is available processing and recommending different products, services and friends. The main challenge associated with a Recommendation System is to recommend useful information to the user at right time. Friend Recommendation, which is the one of the indispensable feature of Social media, has taken it to new height. Facebook, Twitter, LinkedIn, MySpace have captivated millions of users now a days. But the antecedent research work on Friend Recommendation cynosure on user current relation in Social Networking. Facebook, one of the most prominent social networking platforms provides the personalized friend recommendation based on FOAF (Friend of a Friend) ontology. MySpace is based on PYMK (People You May Know) friend recommendation. Basic perception behind it is that probability of a person knowing a friend of friend is more than unknown person. This work proffers a unique approach of friend recommendation based on the user's interest and based on user current location. The proposed system recommends friends based on user interest. Further, user interest keeps on changing. To overcome this challenge, recommendation System is proposed using Ontology and Spreading Activation. we developed a recommendation system which provide content recommendation to user based on user interests which gets

changes over the period of time and system learns this using the spreading Activation algorithm. User interest is being captured using the Spreading Activation. Spreading Activation has been used to overcome variation in user interest. Our experimental results have shown the benefits of considering Spreading activation and ontology in friend recommendation in as social networking.

LIST OF FIGURES

LIST OF TABLES

NOMENCLATURE

OC	online community
VC	virtual community
ICT	information and communication technology
FB	facebook
IF	information filtering
IR	information recovery
CF	collaborative filtering
kNN	k –nearest neighbor
SVD	singular value decomposition
MF	matrix factorization
CBF	content-based filtering
AHP	analytic hierarchy process
FOAF	friend of a friend
LDA	latent dirichlet allocation
VOIP	voice over IP
CSSN	computer-supported social networks
RDF	resource description framework
OWL	web ontology language
KB	knowledge base
SW	semantic web

OPO	online presence Ontology
PIMS	personal information management system
W3C	world wide web consortium
RDF	resource description format
XML	extensible markup language
RDFS	resource description format schema
SPARQL	SPARQL protocol and RDF query language
DAML	DARPA agent markup language
API	application programming interface
HTTP	hypertext markup language
PIM	personal information management
TIM	task information management
SA	spreading activation
GPS	global positioning system
WOA	whale optimization algorithm
SOM	self-organizing maps
GA	genetic algorithm
PSO	particle swarm optimization
ABC	artificial bee colony
PCA	principal component analysis
MAE	mean absolute error

API	application programming interface
HTTP	hypertext markup language
PIM	personal information management
TIM	task information management
SA	spreading activation
GPS	global positioning system
WOA	whale optimization algorithm
SOM	self-organizing maps
GA	genetic algorithm
PSO	particle swarm optimization
ABC	artificial bee colony
PCA	principal component analysis
MAE	mean absolute error

Chapter One: INTRODUCTION

1.1 Background

"Man is a social creature". Socrates had said about the person who is independent of his fellow beings and is unable to live in community or society is either Beast or God. As per dictionary meaning "a group of people living in the same place" means it is based on some shared location. That's what community used to be, historically. They have something in common in such as religion, belief, faith, interest. A people group is said to exist when cooperation between people has the motivation behind gathering singular needs and acquiring bunch objective, a constrained geological region is another aspect, the highlights of social communication, framework for the satisfaction of social and physical needs, and restricted land zone are essential to the meaning of community. T.S. Eliot wrote "What lives have you if you have not life together? There is no life that is not in community". Offline community is defined as (1) A gathering of individuals (2) who share social communication (3) and some basic ties among themselves and different individuals from the gathering (4) who share a territory for probably a portion of the time .But due to industrialization, people are migrating from one place to another place so there location or neighbourhood is no longer a key identity for them. As a result, people are migrating from offline community to online community. Online people group is characterized as online network ought to be comprehended here as comprising of (1) People who connect socially as they endeavour to fulfill their own needs or perform extraordinary jobs, for example, leading or moderating. (2)A common reason, for example, passion, need, information exchange or administration that gives motivation to the community. (3) Policies such as customs, conventions, rule and regulations that guide individual’s cooperation. (4) Computer systems to help and resolve social association and encourage a feeling of harmony. "Online Communities (OCs) are part of the web where individuals can discover and afterward electronically "talk" to others with comparative interests. Online group perform an enormous job in numerous parts of a part's life – from framing and keeping up companionship and exotic relationship, to research, shaping conclusions, buying, and devouring items and services. Specialists are occupied with finding an appropriate plan of action for e-commerce business. Online Community can upgrade trust among individuals along these lines decreasing danger and empowering them for more noteworthy e-commerce business. Online Communities are in this manner perfect apparatuses for web based business, e-advertising, information building and e-learning exercises. In case of online community, communication can be synchronous or asynchronous where as in case of Offline community, it’s always synchronous. Online networks have existed on the Internet for right around a fourth of a century. One reason why individuals are moving from offline to online network is that if there should be an occurrence of online networks, individuals have more options when compared with offline which is limited because of geological limitations. The Well (http://www.well.com), began in 1985, and Usenet newsgroups, began 1979, are broadly viewed as the initial virtual networks on the Internet.

1.2 Motivation

PCs, cell phones and other data and communication innovations have become a significant piece of the regular daily existence in well-to-do social orders, yet huge socio segment differences stay in their utilization. Youthful grown-ups specifically keep on being significantly more dynamic user of communication technologies than the old ones. In the mid-2000s, the rising new communication innovations incited a range of openings with respect to the potential results of the new technological improvements. Mark Prensky, for example, laid a case over a division between the people born into an "advanced era", and those brought into the world before. As indicated by this view, the youthful are on an advance level not the same as the old one as far as their technological capacities, relational abilities and the manner in which they are mingled. This new digital era has led to paradigm shift from offline communities to online communities and social networking. Some of famous social networking sites are Facebook, Myspace, LinkedIn whereas popular social communities are Yelp, YouTube, and Digg etc. Difference between social networking and social communities is that unlike social networks, Communities are held together by normal intrigue. It might be a leisure activity, something the community individuals are energetic around, a shared objective, a typical undertaking, or just the inclination for a comparable way of life, topographical area, or profession. People in the social networking platform either know each other in advance through some pre-established relationship e.g. family, school friends, college friends, colleagues or have a link with other person through some intermediate friends link, for e.g. Friend of a Friend in case of Facebook. In the field of psychology, a wide collection of communities have been characterized based on different criteria. "“No man is an island, entire of itself; every man is a piece of the continent, a part of the main”. Our reality is community. The entirety of our connections with one other, paying little mind to the medium we use to impart, are about community. Community must be viewed "whole of itself". T.S. Eliot expressed “What lives have you if you have not life together? There is no life that is not in community”. Communities of training are gatherings of individuals who share an energy for something that they realize how to do and who associate routinely to figure out how to improve.

The main Issue with the existing social networking platform is that either they are based on Friend of a friend based recommendation or based on the user location. The main motivation for this research proposal is that generally end user interest are keep changing especially in younger generation and there doesn’t exit any platform which cater to this user need .By discovering user changing interest , proposed platform will provide required friend recommendation to end user.

1.3 Research Objectives

The proposed research work goal is as follows:

- To design Social Community Framework using Semantic based User Profile and dynamic User contexts.
- To design Friend recommendation System based on proposed Social Community Framework.
- To design Content Recommendation Framework to Social Community users based on his/her current preferences.

Chapter Two: LITERATURE REVIEW

In this chapter literature review related to research area is presented. The purpose of literature review is to have the better understanding of the field of research and give information to help in the research. Also, the idea behind is to give the simple overview about the terms, concept, issues and techniques studied during the literature survey.

2.1 USER PROFILE ACQUISITION

User profiles assume a significant part in suggestion forms as the information provided by them speak to the user's requirement. Many personalization frameworks required to frame an end-user profile or a user inclination model so as to distinguish the necessities of each consumer (Jannach, Zanker, Felfernig, & Friedrich, 2010). The underlying advance to provide customized proposals is to discover customer interests or inclinations so as to create a required profile. One of the way to gather User's inclinations from their prior interaction with the platform being referred to (Rendle, Freudenthaler, Gantner, & Thieme, 2009). These user connections comprise of either explicit or implicit data about the end-user's inclinations or enthusiasm for items. The profile permits user to be displayed, which can be depicted as the way toward building individual inclinations (Bhowmick, Sarkar, & Basu, 2010). At the end of the day, the user model is commonly spoken to a profile that catches individual inclinations of the end-users as far as the user's information regarding the item or subject where they are intrigued (Liang, Xu, Li, & Nayak, 2010) (Middleton, Shadbolt, & De Rour, Jan. 2004). Profiles speak about passion or inclinations of both a single user as well as a gathering of users: a single user profile gives just individual user's inclinations and data, though a gathering profile portrays regular interests or objectives of a gathering of users (Liang, Xu, Li, & Nayak, 2010). Profiling is the technique of collecting data regarding the points or area in which a user is intrigued (Gauch, Speretta, Chandramouli, & Micarelli, 2007).Precision or adequacy of profiling influences the presentation of recommender frameworks. The vital part of profiles is their capacity to speak to user's present interests. According to (Gauch, Speretta, Chandramouli, & Micarelli, 2007), the user profiling process comprises of three primary stages:

• **The 1st stage includes data gathering. This procedure is used to accumulate basic data about the end-user**

The initial step for making a suggestion framework is by collecting data of the user. For creating a profile or account of the user, the framework necessitates important data regarding the user's inclinations or choices. Various kinds of user data references and systems are present in the world that can continue utilized to find user's inclinations or interests. Essentially, frameworks further accumulate user choices or inclinations of consumer criticism. Criticism can signify express or certain, being clarified further in subsection 2.1.1.

• 2nd stage remains user profile development and depiction

A fundamental of the customized recommender frameworks is the means by which to construct customer profile, which includes the data needs and inclination of customer and has extraordinary effect on the presentation of suggestions. One significant thought while developing a profile is that more authentic the client profile is, the more powerful the suggestions will be. Client profiles are built utilizing various procedures dependent on the client profile portrayals. Client profiling is either information based or conduct based (Middleton, Shadbolt, & De Rour, Jan. 2004). Knowledge-based methodologies emphasize complete domain information regarding products and implicit information regarding the customer. The methodologies are rule-based in proposing things which precisely meet guidelines to users, for instance utilizing choice standards to order user's very own advantages or inclinations dependent on their segment attributes (Amini, Ibrahim, & Othman, 2011). Widely used recommendation frameworks practice action dependent on methodology to develop profile for user. Behavior methodologies utilize user's action as a model and find the helpful examples of that action by utilizing AI (Middleton, Shadbolt, & De Rour, Jan. 2004).User action can be characterized to by means of a few sorts of examples, for example, frequency patterns, sequential patterns, neural network models, and graph models (Liang, Xu, Li, & Nayak, 2010). Similarly manner profile can be created concerning either certain information of user (For instance, series of catchphrases, web utilization information, and content and basic data about visited site pages) or direct user information (e.g., polls or meetings with the user). In light of distinctive data acquired concerning user, profiles of user can be spoken to assets of weighted catchphrases, subjects, ideas, appraisals. Every catchphrase speaks to a subject important to the user. These catchphrases can be separated from the substance of items or given legitimately by the user. The level of user enthusiasm for the keywords can be weighted utilizing the tf-idf technique in the vector-space approach (Gauch, Speretta, Chandramouli, & Micarelli, 2007) (Lops, Gemmis, & Semeraro, 2011). Different methodologies taken to speak to user profiles incorporate the history-based model (which consider user buying details and valuation), the vector space model (which uses highlight vectors to show things), weighted n-grams (which shows as a word graph, with having weight on the nodes and the links), or semantic systems (where the profiles might be represented by a weighted semantic system wherein every node of semantic graph denotes to a concept). Further choices are weighted associative networks (where each node of the associated network represent individual user profile), rating matrices about the user-item (where each rating represent or show the interest of the user about that particular item) and demographic features (which build end user information through user attributes) (Montaner, Lpez, & Rosa, 2003). In rating-based user recommendation system (Liang, Xu, Li, & Nayak, 2010) (Montaner, Lpez, & Rosa, 2003), user assessment for items are mostly employ to express their item inclination. But the performance of the traditional collaborative recommender systems remains cool commencement difficulty due until availability of user ratings at initial level. To solve the cold-start problem, Zhang and Koren (Zhang & Koren, Efficient bayesian hierarchical user modeling for recommendation system, 2007)suggested to use Bayesian hierarchical linear models based analytics in order to boost content based user profiles. The main benefit is that even though there is no rating data in the system, still the user preferences can be achieved. In addition, user profiles can be shown to by means of a concept based, which is like the catchphrase based technique, aside from the information is introduced as vectors of weighted features. At the beginning, to show the content of web pages, the concept hierarchy was used. All the more as of late, few research have used concept hierarchy to show user profiles that echo the content of a given user's interest in the hierarchical structure (Nanas, Uren, & Roeck, Building and applying a concept hierarchy representation of a user profile, 2003) (Singh, Shepherd, Duffy, & Watters, 2006) (Yu, Liu, & Zhao, 2012). (Singh, Shepherd, Duffy, & Watters, 2006)Proposed a news filtering recommendation system which uses end user interest captured based on explicit feedback to model user's interest. Broadly, the reference taxonomy or vocabulary is being used to build the concept-hierarchical profile. Many research studies attribute ontology as a concept-hierarchy whereas the relationship between the concept is 'is-a' relationship. The hierarchical link between different concept nodes has either a super-concept or a sub-concept. User profile is constructed based on a set of concepts and relationships between them. A collection of concepts related with an end user is knows as user profile, whereas a collection of concepts related with an item is called as an item profile (Jannach, Zanker, Felfernig, & Friedrich, 2010). The users' inclination or interests can be seen as concept vectors in this approach. The utilization of concept hierarchy to show users' interests is clarified further by (Kim & Chan, 2008). (Nanas, Uren, & Roeck, Building and applying a concept hierarchy representation of a user profile, 2003)Came up with technique for building a concept hierarchy under documents' topic to show the users' topic interests. In light of the connection between concept or ideas in a various levelled structure, numerous examinations likewise looked into the new ways to deal with build end user profile, for example, utilizing the hierarchical connections between topics in scientific categorization to show user's ordered subject interests or inclinations (Liang, Xu, Li, & Nayak, 2010;

Weng, Xu, Li, & Nayak, 2008; Ziegler, Lausen, & Schmidt-Thieme, Taxonomy-driven computation of product recommendations, 2004; Zhang, Ahmed, Josifovski, & Smola, 2014). In Section 2.3, will discuss about the recommender system based on Taxonomy-based user profiles. These days, the improvement of Web 2.0 and Semantic Web innovation give plentiful literary substantial data, interactive media content data, and system data including labels, audit, remarks, posts, pictures, tweets, recordings, sound clasps, and long range interpersonal communication which give an important asset while building user profiles (Tang, Yao, Zhang, & Zhang, 2010; Liang, Xu, Li, & Nayak, 2010). (Li & Chang, 2005) Introduced strategy of combining information to build user profile. To showcase user's real-time preference for Web personalized services, (Yu, Liu, & Zhao, 2012) build user profile dependent on concept and relation. In today's era of Web 2.0, textual content information has taken in form of tag or folksonomy, and has become a significant research area. Core point of research toward tag learning is to focus on the semantics of words to enhance the collaborative filtering recommendation outcome, along with to overcome the issues such as the cold-start problem (Lops P. , Gemmis, Semeraro, & Musto, 2009). The primary thought of utilizing labels or folksonomy is to profile users' topic interests or inclination when the amount of available user ratings information availability is too low (Djuana, Xu, Li, & Josang, 2011; Liang, Xu, Li, & Nayak, 2010). Due to new development in web domain, developers or researchers are capable to accomplish extra data to build and represent user profiles.

- **The 3rd stage comprises of using user profile information to provide personalized services**

Once user profile happen to be developed, same has been used to offer customized assistance in different domain areas, for example, as customized recommender methods, customized explorations, questions, and trust aware based recommender systems. There are three primary strategies used in recommender framework for user profile exploitation: content-based, collaborative, and hybrid methods, which will all be subsequently discussed in Section 2.2. Furthermore, collection of labels, catchphrases, or ideas are utilized to give customized searches in this phase.

2.1.1 User Information Collection

Primary stage of user profiling remains towards gathering data regarding user. Having the option towards recognizing user's requirements, recommendation method has to understand something regarding user needs. Subsequently, gathering knowledge regarding user requirements remains one crucial activity concerning creating customized recommendations. User inclination continue experienced from users' associations among items. These inter communication comprise of straight forward and derived information, generally designate as explicit feedback and implicit feedback (Jawaheer, Weller, & Kostkova, 2014). Subsequent subsection continues explain about explicit and implicit feedback in more and more detail.

2.1.2 Explicit user information collection

Mainly explicit knowledge about used is based on what information user had given at time of providing the information. This data is procured straightforwardly by means of registration forms or surveys, or by questioning end users to rank products, by following users' questions (Jawaheer, Weller, & Kostkova, 2014). Mostly websites gather user inclination through giving customized assistance towards end users and afterward straightforwardly requesting that they give individual data to make a profile. This explicit user information mirror the actual user desires. For instance, eBay request users to give their impression and to rank for the services and commodity on offer. The organization at that point uses this data to improve the customized suggestions that it provides for customers. This information incorporates demographic information (e.g., sexual orientation, background related to educational, date of birth, area, and work), information regarding choices further more inclination (e.g., topics of interest, tastes, preferred products and brand preferences), and opinion-based information (e.g., reviews, comments, and feedback) (Gauch, Speretta, Chandramouli, & Micarelli, 2007; Liang, Xu, Li, & Nayak, 2010; Montaner, Lpez, & Rosa, 2003). In recommender systems, explicit information mainly extensively used to profile users' inclination (Jawaheer, Weller, & Kostkova, 2014; Bell & Koren, Improved neighborhood-based collaborative filtering, 2007; Deshpande & Karypis, 2004). A few sites work by utilizing this type of information, for example, Netflix, which uses film rating to create well known film recommendations for users (Pil´aszy & Tikk, 2009). Albeit explicit evaluation is successful and easy to gather, and has less commotion, there are few disadvantages. To begin with, users need to put time and exertion in communicating their inclinations or interests through activities (Liang, Xu, Li, & Nayak, 2010). In different terms, specific feedback puts extra weight toward every user. On the off chance that users don't give individual data, user profiles can't be constructed. Another concern regarding the security concerns; it may be possible that users may not be ready for ex. changing private data or towards allowing precise knowledge to system. These issues influence accuracy concerning recommender systems further more present it hard towards profile users correctly. In this way, reassuring users to give adequate express data is a difficult assignment.

2.1.3 Implicit user information collection

Explicit information is not everlasting accessible furthermore does not generally have ample knowledge concerning a fair user profile towards its construction. Due to issue, implicit user knowledge usually gathered and this collected information is based on user actions. There are many ways to collect implicit information, some of them are through web utilization logs, click streams, perusing narratives, buy records, and content or basic data from visited website pages (Daoud, Tamine, Boughanem, & Chebaro, 2007). Browsing histories are a typical wellspring for getting implicit data (Gauch, Speretta, Chandramouli, & Micarelli, 2007). (Yu, Liu, & Zhao, 2012)Computed user’s real-time preference by scanning user’s browsing contents of each web page. The principle advantage of getting implicit data is that it doesn't require any additional exertion with respect to users during the way toward building profiles (Jannach, Zanker, Felfernig, & Friedrich, 2010). It likewise permits simple and nonstop access to information; it is consequently refreshed when users associate with the system being referred to. Be that as it may, it is hard to change over user conduct into user inclinations, as the exactness relies upon whether the user conduct is deciphered accurately. For instance, user may purchase things, for example, music for another person. One downside concerning this issue of data collection is that it requires advancement of great applications or modules, that web designers need for introduction. Understood user implicit knowledge that's rich wellspring information which permits customized suggestions to be made. (Weng, Xu, Li, & Nayak, 2008; Ziegler, Lausen, & Schmidt-Thieme, Taxonomy-driven computation of product recommendations, 2004)Suggested the utilization of item taxonomy in depicting user's interesting topics for making the recommendations. (Kim & Chan, 2008)Proposed that watchwords or topics ought to be caught from user perusing chronicles so as to evaluate their enthusiasm.

Besides, after arrival of Web 2.0, remarkable different kinds of user data can utilized as part of implicit information, for example, labels, remarks, pictures, recordings, posts, and snaps. This information gives rich data regarding connection between users, things, furthermore substance moreover both additionally infer user inclinations. Labelling among a user for instance, catchphrases on labels can be utilized to catch the user's subject advantages.Amazon.com utilizes utilization logs concerning users to prescribe books to their end-users (Jannach, Zanker, Felfernig, & Friedrich, 2010). Information resources talked about above are significant, yet there stay a few impediments identified with information gathering. Security concerns may make a few users retain data or carry on distinctively when signed in to the recommendation system. Be that as it may, the benefit of user profiling lies in the entrance to both implicit and explicit user inclination data. Few highlights concerning user data that can gathered and used to recover things that are important to users. Nevertheless, finding out new users' interest is difficult as not much knowledge is accessible, and even that may be incorrect, as aforementioned. Significant subject of research comparable to recommender system and personalization system.

2.2 RECOMMENDER SYSTEMS

Recommender frameworks have become a notable subject of research, as the amount and scope of data accessible on the Internet is colossal. This abundance of data implies that it very well may be hard for users to decide. This abundance of data implies that it very well may be hard for end-users to decide. Recommendation system are well known applications used for Information filtering (IF) and Information recovery (IR) systems in light of the fact that the proposals that they make help users with data they are looking for, by making customized suggestions comparable to data, items, and services. Recommendation systems attempt to foresee what things would be fascinating to users and to meet their needs. System make suggestions moreover forecasts dependent upon data about user, for example, profiles, inclinations, history. These days, information about the users, product they are using or rating they are giving related information are being used to provide recommendations. This information based way to deal with creating proposals includes thinking about which items or subjects will meet user's prerequisites. According to (Chen, Wu, Xie, & Guo., 2011), recommender systems comprises of three main modules: (1) a user information gathering module (2) reasoning module, which investigate user behaviour; and (3) Core module to provides recommendations. In the following section will present the fundamental algorithm broadly and effectively utilized in recommender systems.

Algorithm used in Recommender system may be separated comprehensively within 3 classifications: collaborative based, content based and hybrid one. These classifications are analyzed beneath.

2.2.1 Collaborative Based Recommendation System

Famous websites such as LinkedIn, Facebook, Twitter, Google, Netflix and Amazon are based on Collaborative based recommendation (CF) (McLaughlin & Herlocker, information retrieval; Xu, Bu, Chen, & Cai, 2012)to recommend occupations, buddies, and organization in which users may be intrigued (Schafer, Frankowski, Herlocker, & Sen, 2007). The fundamental undertakings in CF are appraising forecasts and Top N suggestions (Deshpande & Karypis, 2004; Herlocker, Konstan, Terveen, & Riedl, 2004). To predict user behaviour, CF approaches use historical information about the user, along with information about the current users in the community. Fundamentally, CF methods make suggestions regarding the choices concerning any user-dependent covering likenesses among those inclinations of users in a system. The fundamental concept regarding CF implies that those users who rated similar things are bound to have similar inclination. The procedures use user rating against different users to decide the connection among user and item, and convert the inclination of a user for item to a user-item matrix (Breese, Heckerman, & Kadie, 1998; Herlocker, Konstan, & Riedl, Explaining collaborative filtering recommendations, 2000; Koenigstein, Dror, & Koren, 2011; Ricci, Rokach, & Shapira, 2011; Su & Khoshgoftaar, 2009). The best outstanding algorithm in CF are the neighbourhood method and the latent factor model (Kanagal, et al., 2012; Koren, Bell, & Volinsky, Matrix factorization techniques for recommender systems, 2009).

You would hope to get the best proposal from somebody with comparable taste. The issue, however, is the means by which to discover such an individual. You may need to take part in numerous connections with loads of various individuals, through which you gradually find out about each other's inclinations, before you begin to get proposals you are certain about.

Collaborative based systems investigates strategy for coordinating individuals with identical passion and afterward making suggestions on this premise. Three challenges of this methodology are (1) numerous individuals must take an interest (making it likely that any given individual will discover others with comparable inclinations), (2) there must be a simple path for individuals to showcase their inclinations to the system, and (3) algorithm should have the option to find individuals with comparative interests.

Collaborative based system has made the end user task very straightforward: you express your inclinations by ranking things (like food or films or sports) that the recommendation system presents to you. These ratings at that point fill in as a surmised portrayal of your preference. The system at that point coordinates these rankings against evaluations put together by every single other end user of the system. The outcome is the arrangement of your "closest neighbors"; this formalizes the idea of individuals with comparative taste. At long last, the system prescribes things that your closest neighbors evaluated exceptionally that you have not appraised (and probably are in this way curious about); a key issue is the manner by which to join and weight the inclinations of your neighbors.

You can quickly rate the suggested things on the off chance that they don't intrigue you; in this manner, after some time, the framework gains an undeniably exact portrayal of your inclinations.

GroupLens (Resnick P. , Iacovou, Suchak, Bergstrom, & Riedl, 1994), the Bellcore Video Recommender (Resnick P. , Iacovou, Suchak, Bergstrom, & Riedl, 1994), and Firefly (Shardanand & Maes, 1995) are also based on collaborative filtering. The recommender system changed the way they weighted the ranking of various clients (i.e., figured out who your neighbors were and how close they were) and how they clubbed the evaluations.

CF methods can be arranged as per the algorithm methods into 2 classes: memory-based CF (or neighbourhood-based CF) and model-based CF (Hahsler, 2011).

2.2.1.1 Memory-based CF algorithms

These algorithms utilize the entire collection of items previously rated by a user to make recommendations (Liang, Xu, Li, & Nayak, 2010; Adomavicius & Tuzhilin, 2005). Memory-based collaborative filtering algorithms are commonly referred to as neighbourhood-based algorithms (Recommender systems, 2010). They can be divided further into user-based CF algorithms and item-based CF algorithms (Deshpande & Karypis, 2004; Sarwar, Karypis, Konstan, & Riedl, 2001).User-based CF makes recommendation based on the similarities between an active user and other users, while item-based CF makes recommendations based on the similarities between a target item and other items (Sarwar, Karypis, Konstan, & Riedl, 2001). In memory-based algorithms, a user’s preferences for an item are evaluated based on the ratings data of other users who have similar behavior to the user. These k-Nearest Neighbour (kNN) techniques are widely used in CF based algorithms to identify a group neighborhood of users and items that are similar to a user or an item. The algorithms use the given rating data by similar users for many items to predict missing ratings or create a Top-N recommendation list for the active user. To form a neighborhood for the active user, a similarity measure is required. The top-k neighbor users and items for the active user can be selected by calculating the similarity between the active user and all other users or all other items. The similarity measure can be calculated by various kinds of proximity computing approaches. The most common methods utilized for determining the similarities between users or items are the Pearson correlation and vector cosine similarity measures. There are also several other similarity measures used in the literature, including adjusted cosine similarity, Euclidean distance and the Jaccard coefficient (Chen, Wu, Xie, & Guo., 2011; Herlocker, Konstan, Terveen, & Riedl, 2004; Su & Khoshgoftaar, 2009). The rating data plays an important role in CF techniques to form the neighborhood. Bell and Koren (Bell & Koren, Improved neighborhood-based collaborative filtering, 2007) proposed the neighborhood-based approach to improve the accuracy of kNN approaches without meaningfully affecting running time. However, when the amount of user direct rating data in the system is too small it would result in poor neighbor formation and recommendations. Besides using explicit ratings and implicit ratings, the similarities between users, items, or user-items can be measured based on other features, such as users' topic interests or users' tagging behavior (Liang, Xu, Li, & Nayak, 2010; Weng, Xu, Li, & Nayak, 2008; Zhang, Ahmed, Josifovski, & Smola, 2014; Nadee, Li, & Xu., 2013). (Weng, Xu, Li, & Nayak, 2008)Utilized the taxonomy information of the item incorporated with the existing user's explicit rating in the neighborhood formation, instead of using only the rating data.

2.2.1.2 Model-based CF algorithms

These algorithms use a collection of ratings to learn the pattern of ratings, and then make intelligent rating predictions based on the learned models. The models are developed using data mining techniques and machine learning algorithms to explain the rating pattern. In contrast to memory-based CF algorithms, model-based approaches are not subject to heuristic prediction rules. There are many model-based CF algorithms, including Bayesian network-based models, clustering models, linear regression models, latent factor models, linear regression, singular value decomposition models (SVD) and matrix factorization (MF) models. A key advantage of the model-based approach is that it improves prediction performance. Recently, the use of matrix factorization and latent factor models has become popular in recommender algorithms, both for implicit and explicit feedback. Typically, matrix factorization classifies both items and users via factor vectors that are gleaned from item rating patterns. Many studies have utilized matrix factorization and neighborhood methods to improve the performance of Collaborative Filtering and alleviate or solve the cold-start problem (Cremonesi, Koren, & Turrin, 2010; Koren, Factorization meets the neighborhood: multifaceted collaborative filtering model, 2008; Tak´acs, Pil´aszy, N´emeth, & Tikk, 2008).Most memory-based CF approaches can be used for both rating predictions and Top-N recommendation tasks, whereas model-based CF approaches focus on rating predictions. Standard CF algorithms, including user-based and item-based examples, are popular benchmark baseline models. These includes latent factor and matrix factorization models, which have emerged as state-of-the-art methodologies in recommender techniques (Recommender systems, 2010). The advantages of collaborative filtering are that it is easy to implement and incorporate with other information sources, but the next section explores its drawbacks.

2.2.1.3 - Challenges and limitations of collaborative filtering

Collaborative filtering approach has few issues, for example, versatility, the dark sheep issue, protection concerning privacy, divergence, sparseness, moreover cold start obstacle (Koenigstein, Dror, & Koren, 2011; Su & Khoshgoftaar, 2009; Adomavicius & Tuzhilin, 2005; Koren, Factor in the neighbors: Scalable and accurate collaborative filtering, 2010). These problems can decrease the recommendation accuracy. In order to have great outcome, it is important that CF algorithms will overcome these challenges. The sparsity issue happens when the quantity of items is incredibly huge (Guo, Resolving data sparsity and cold start in recommender systems, 2012; Guo, Improving the performance of recommender systems by alleviating the data sparsity and cold start problems, 2013; Papagelis, Plexousakis, & Kutsuras, 2005). Users generally give ratings to small number of items as compared to large amount of total products contains inside dataset, causing less quantity concerning ratings per product. Accordingly, it leads to highly sparse user-item matrix; finally lessen the efficiency of the Collaborative Filtering systems at time of making recommendations. Thus, CF system must be able to manage exceptionally meagre information. The cold-start issue happens whenever a new user or item arrive in the system first time causing the system to come to any conclusion between users and items due to lack of information about the user or item (Chen, Wu, Xie, & Guo., 2011; Su & Khoshgoftaar, 2009; Recommender systems, 2010; Guo, Improving the performance of recommender systems by alleviating the data sparsity and cold start problems, 2013). For good and precise recommendation, recommendation system have to increase few insight about both users and items (Papagelis, Plexousakis, & Kutsuras, 2005). This issue can to a great extent deteriorate the accuracy of the conventional recommender systems in giving customized predictions, especially when making a nearest group of active user's neighbourhood towards energetic user which depends upon users' ratings information.

2.2.2 Content-Based Recommendation System

Content-based recommendation system make use of user-profiles information and the item contents primary information and match new item information towards account concerning users. Products that matched as per user's profile are recommended (Pazzani & Billsus, 2007) to the end-user. Content-based recommendation create a model towards basis concerning items details which user earlier liked or given a particular specific rating moreover feedbacks (Nanas, Roeck, & Vavalis, What happened to content-based information filtering?, 2009).User's benefit or inclinations can be depict in terms of item qualities, for example, points, traits, or classifications. Thusly, a user profile may comprise of the user's inclinations, needs, and verifiable or express interests, (for example, sets of things, subjects, ideas, or evaluations) (Lops, Gemmis, & Semeraro, 2011). For instance, the users' preferences can be captured by the action performed by the user earlier on the items that user evaluated, clicked, scan, or bought. Content based filtering prescribes using these profiles to discover different things with comparative substance to the article favoured by users. Inclination between user and item can be find out by determining the content of item alongside utilizing the implicit and explicit ratings. Set of features which describe item content relate to collection of keywords, topics, terms, or concepts. Most content-based methodologies created utilizing information retrieval methodology (Lops, Gemmis, & Semeraro, 2011). These procedures attempt to match query words or other user information with item features (Schein, Popescul, Ungar, & Pennock, 2002). There are two well-known ways to represent item: heuristics-based and model-based technique. In a heuristics-based technique, both user and item profile is shown as a vector of weights for every feature. The most broadly-utilized weighting function is tf-idf which is used to find out the significance of word related with an item (Lops, Gemmis, & Semeraro, 2011; Pazzani & Billsus, 2007). A user's inclination for an item can be determined by making use of cosine similarity. In the model-based technique, user inclination determined by using probability and naive Bayes, language models, machine learning, decision trees, and linear classifiers (Pazzani & Billsus, 2007; Zhang & Koren, Efficient bayesian hierarchical user modeling for recommendation system, 2007). The benefit of content-based based recommender systems is that they don't need users' ratings information. User inclination is represented using the items content rather than item rating. But, there are likewise a few constraints and difficulties to content-based recommender system, as talked about beneath.

A fruitful recommendation system needs to address various inherent difficulties that respectively establish an innovative field. In light concerning changing characteristics of news details a few difficulties possess deeper significance than some others. Accompanying difficulties are totally related however not totally explicit to news region. The majority concerns difficulties as usual difficulties regarding recommender system wherever a portion regarding particular test may not be an issue in different spaces.

Data sparsity: When there are not sufficient rating provided by the users then the matrices used can be very scant. The probability regarding knowledge inadequacy grows if the count of columns or rows is excess than the other. For instance; if the quantity of items is significantly more than the quantity of users then it requires an excessive number of evaluations to fill the matrix. Information sparsity causes a reduction in the accuracy of the system.

Recency: It is one of the most challenging issue in news recommendation area. A large portion of users wants to view latest news rather than old dated ones. Significance of news diminishes over a period of time. Then again, some news articles might be associated with one another such that the user might need to view the old news identified with the one user as of now viewing or might need to keep educated about that topic (Li L. , Wang, Li, Knox, & Padmanabhan, 2011).

Implicit user response: Users response are very essential to make increasingly exact suggestions. In absence of clear-cut feedbacks it may not be conceivable towards judging if the user liked the article she read or not (Fortuna, Fortuna, & Mladenic, 2010). Yet, it isn't viable for the recommendation system to communicate with the user consistently. Ultimately system must have the option to gather user inputs successfully while securing the user privacy.

Changing interests of users: User continuous changing interest is also a challenge in foreseeing the future interests of user for better recommendations (Liu, Dolan, & Pedersen, 2010). For certain areas like film or book suggestions, the fluctuation in user passion happens all the more gradually. Be that as it may, for the news space it is extremely difficult to anticipate the changes. Additionally a few people may peruse the news not on the grounds that he/she is interested in the topic but may be because its significant.

User modeling/profiling (Knowledge of user preferences): A significant factor of recommender systems. It is required to build user profile to build user interest specific recommendation.

Serendipity (over-specialization, portfolio) problem: Another issue arises when the system suggest comparative or similar things with the as of now suggested ones. In case of the news recommender system, a recent news composed diversely in various news sources might suggested by recommender system regarding various articles. Clearly the users would not be glad to get the equivalent or comparative suggestions. The system ought to consistently have the option to find new things to prescribe by maintaining a strategic distance from similar things.

	Recency	**Virtual Feedback**	**Dynamic User Interest**	**Sparsity of Data**
(Resnick P. , Iacovou, Suchak, & Bergstrom, 1994)	O	N/A	N/A	N/A
(Li L. , Wang, Li, Knox, & Padmanabha, 2011)	O	O	N/A	N/A
(Fortuna, Fortuna, & Mladenic, 2010)	O	O	N/A	N/A
(Lin, Xie, Li, Huang, & Li, 2012)	N/A	N/A	N/A	O

Table 2.1 works on news recommendation with the different challenges

(Li L. , Wang, Li, Knox, & Padmanabhan, 2011)Suggested versatile 2 phase customized news recommendation approach with a two-stage representation, which thinks about the elite qualities (e.g., content to be shown, way of access, prevalence and recency) of news items when giving suggestions. Additionally, a principled structure for choosing news depends on the underlying interest of user is presented, with a decent harmony between the uniqueness and distinctiveness of the recommended outcome. The viability and proficiency of methodology is proved the broad experiments performed on the collection of new articles collected from different new sites.

(Khandelwal, Shanbhag, Shriyan, Thorve, & Borse, 2018)Presented the event-based methodology for News Recommendation based on both users inclination. News stories are prescribed based on ML techniques like grouping of comparable articles, anticipating their classification, comparability and watchword extraction. There are different criteria's that system takes as input, like time spent by user on reading the article, article liked or not-liked by user towards determining interest concerning user. Proposed system recommends new articles based on different events and show them in chronological array & attaches new up and coming related articles, in this way giving an excursion of that event.

(Yeung & Yang, 2010)Introduced one news recommender system which further proactively pushes in the nick of time customized news stories to users dependent on user's relevant data just as news content. User's information requirement are assessed depending on Bayesian network technique. An "Analytic Hierarchy Process (AHP) Model", which underpins both Content-based as well as Collaborative filtering, is developed to rate the significance of news articles.

(Sadasivam & Praveen, 2014)Suggested a news recommendation system based on four personalization traits: profile of a user, interest of a group, along with location and feelings. The percentage of these traits isn't the equivalent for all users. Therefore genetic algorithm is utilized to recognize the percentage of the characteristics and customize it for a specific user.

2.2.2.1 Challenges and limitation of content-based filtering

In this particular techniques, if number of features are limited then it leads to a limited content analysis which results in wrong recommendations (Pazzani & Billsus, 2007).Accuracy of content based recommendation system depends on the availability of sufficient information to analyze as they don't require user rating data to learn. For instance, there isn't adequate catchphrases to show the end-user interests in books. In addition, the over-specialization issue happens when a user is just prescribed things that are like things that were evaluated or purchased previously. For instance, assume that at one stage in a user's life she is keen on the regular day to day existence of individuals surviving in age of medieval period. She purchases book towards that topic moreover afterward doesn't utilize Amazon.com again for a year. In the wake of returning to the system she sees an enormous number of books on medieval life, however little else in the suggested items area. Not exclusively is this individual being ineffectively comprehended by the system, yet her enthusiasm for this point passed quite a while in the past. To be sure, it was fulfilled by that one book, and she needn't bother with any more. The over-specialization issue has happened and no valuable suggestions are being made to the end-user. (Abbassi, Yahia, Lakshmanan, Vassilvitskii, & Yu, 2009) Investigated the over-specialization issue based on item regions. This is on the grounds that content-based recommendation have no intrinsic strategy for producing or discovering things not quite the same as things users have seen previously; they can just suggest things that score exceptionally based on user's profile (Montaner, Lpez, & Rosa, 2003; Su & Khoshgoftaar, 2009; Adomavicius & Tuzhilin, 2005; Kim Y. S., 2013). In another instance, an end-users whose profile contains no involvement in Thai cuisine could never get a suggestion for a Thai eatery, regardless of whether it was their preferred sort of cuisine. One approach to take care of the over-specialization issue is to make various and fortunate things show up on the recommendation lists (Lops, Gemmis, & Semeraro, 2011). A fortunate proposal encourages the user to find unforeseen yet intriguing things with a high level of curiosity that the user probably won't have gone searching for autonomously (Hurley & Zhang, 2011; Vargas, 2014). (Ziegler, McNee, Konstan, & Lausen, 2005)Suggested a taxonomy-based approach to recommend decent variety of items based on the variety of taxonomy classification. Another problem same as collaborative filtering methods is the novice user or cold-start problem but it is less afflicted than the collaborative filtering techniques (Weng, Xu, Li, & Nayak, 2008). In order for recommendations to be made both user and the item information is required (Lops, Gemmis, & Semeraro, 2011; Adomavicius & Tuzhilin, 2005). With regards to content based recommendations, additional assessment acts demanded concerning those methods that utilized upon excerpt article information moreover finally recommend articles those meet a user's choices. Another significant inquiry concerns how the known qualities of items can be harnessed so as to make valuable recommendations.

2.2.3 Demographic Based Recommendation System

Demographic based Recommender Systems employ end-user traits, categories as demographic information, so as to create their suggestions, occasionally with the support of pre-created demographic groups. Krulwich (Krulwich, 1997)and Pazzani (Pazzani M. J., 1999) have introduced systems which depend on demographic information. This class of system recommend things dependent on the segment profile of end-user. It tends to be utilized to recognize the sense of end-user that are associated to some specific group. In this way, to plan these systems, we need a few data about end-users to classify them into clusters. At that point, if a few end user in a specific cluster like or request a thing, it is conceivable that different end-user of this cluster will in general do likewise. It ought to be noticed that in spite of the fact that it may be smarter to utilize progressively organized data about end users, there is a trade-off between the algorithm complexity and the nature of demographic filtering. Pazzani (Pazzani M. J., 1999) performed a research dependent on demographic filtering on information about eating outlets and he asserted that by and large, 57.5% of the best three suggested eating outlets were enjoyed by end-users.

2.2.4 Knowledge Based Recommendation System

This category of recommender system is utilized in explicit areas where the buying history of the end user is limited. In this type of systems, the logic take into account about the information about the things, for example, highlights, inclinations or interest asked from the user directly, and suggestion rules, before giving proposals. The efficiency of the model is decided on how helpful the prescribed thing is to the end user. Take, for instance, a situation wherein you are building a recommender system that suggests automotive, for example, newly launched car, where a large portion of the end users are novice. For this scenario, the system takes into consideration about the attributes of the product, and end user profiles are created by getting extra data from the end user, and afterward suggestions are made. The main drawback encountered by the Knowledge based recommendation system is the development of the information base, which generally is a challenging work that expect reasonable domain knowledge, and ability in information depiction.

2.2.5 Hybrid Recommendation System

A hybrid recommendation system comprises of distinct recommendation techniques, for example, collaborative based, content-based, information based, and demographic based (Jannach, Zanker, Felfernig, & Friedrich, 2010; Recommender systems, 2010; Burke, Hybrid recommender systems: Survey and experiments, 2002; Pathak, Matharia, & Murthy, 2013). The fundamental objective of this methodology is to improve accuracy in terms of suggestions and to conquer a portion of the problems impacting recommender systems, for example, the cold-start and sparsity problems. Numerous hybrid methods are amalgamation of classical collaborative filtering and content-based filtering (Burke, Hybrid web recommender systems, 2007) (Weng, Xu, Li, & Nayak, 2008; Bakardjieva, 2003). Instance, consolidating collaborative filtering into a content-based method to overpower the cold-start problem (Deshpande & Karypis, 2004; Ricci, Rokach, & Shapira, 2011; Adomavicius & Tuzhilin, 2005). (Barrag´ans-Mart´ınez, et al., 2010)Recommended new hybrid technique by use of content-based and item based collaborative filtering approaches, along with use of “singular value decomposition” so as to recommend TV programs. In other hybrid systems to overcome the sparsity and cold-start problems, mixture of latent factor models and item taxonomy information have been combined to facilitate the development of more personalized recommendations (Zhang, Ahmed, Josifovski, & Smola, 2014; Kanagal, et al., 2012; Ahmed, et al., 2013).Still, the cold-start and sparsity problems still remain an issue as far as recommender systems are concerned. A few strategies suggested concerning managing specific points, including size modification regarding user-item model and by the use of associative retrieval techniques (Chen, Wu, Xie, & Guo., 2011).Content-supported CF approaches can likewise be utilized to increase extra additional information about items, with the end goal of figuring important similarities between them (Gantner, Drumond, Freudenthaler, Rendle, & Schmidt-Thieme, 2010; Melville, Mooney, & Nagarajan, Content-boosted collaborative filtering for improved recommendations, 2002). In any case, there are a couple of shortcomings of hybrid system. In the first place, some of the time there is lack of contextual information to builds system well enough to make recommendations. Also, the versatility issue despite everything exists, due to increase in the quantity of end-users and articles develops quickly, the system need additional time for calculation. Along these lines, the accuracy of hybrid system does not depend only combining the different techniques together. Rather, great outcome and a high quality of suggestions depend on the utilization of suitable contextual information in the recommendation process, for example, both the content and the user behavior.

2.3 TECHNIQUES OF FRIEND RECOMMENDATION SYSTEM

There are various different types of technique to recommend a friend on social media that are as follows:

2.3.1 Potential Friend Recommendation in Online Social Network:

In this system, initially, end user interest is analyzed depending on the domain the user is engaged. This layer is based on the particular domain. In the second layer, users' relationship is calculated. In any particular area, additional domain information can be considered. In the 3rd and final layer, which is, friend recommendation layer, effective recommending policy designed is required based on the below perspective:

Personalized and adaptive associations merging guideline, based upon user's inevitable feedback. Sometimes both this content corporation, including setting associated among a weighted ordinary purpose, require to be viewed, and that particular importance can be improved if a user joins a recommended colleague with the more critical content connection or more powerful connection relationship.

Neighborhood determining policy, to decide wherewith several connected likely companions to continue prescribed instead what some minimal recommending determination. Position concerning most next friend required to be refreshed through randomly determining some portion concerning k nearest neighbours to suggest.

Recommending modernizing policy should be updated regularly as users' concerns may vary as conditions works or the situation might improve.

2.3.2 Shortest Path Based Potential Common Friend Recommendation in Social Networks:

In this method, a universal system signifies viewed as a graph primarily. Wherever every user is a node, and association between every 2 users is considered in terms of links.

In the first round, top-k quickest route are determined performing utilizing various sources Floyd-Warshall algorithm. The pruning approach implemented for optimizing outcome. Formerly possibly familiar friends are achieved towards enlarging the longest common subsequence algorithm to exclude repeated items.

2.3.3 A Social Trust Based Friend Recommender for Online Communities:

The social trust-based structure reflects confidence loyalty highlights hold towards each unique individual, since uncovered by their characters into that online social society. These calculations are totally subject to Social Trust (STrust) and Social Graph (FOAF). Social Trust is made by couple kinds of trust: popularity trust (PopTrust) and engagement trust (EngTrust). This popularity trust of any user is calculated as total trust other members of the community have towards that particular user or member. Conversely, the engagement trust is calculated as the trust of the member towards other individuals. Trust connections are unequal (i.e., A believing B doesn't really mean B trusting in and the other way around). Trust can be dynamic (for example preferring a post) or static (e.g., looking over a post).

2.3.4 Friend Recommendation based on the Similarity of Micro-blog User Model:

In this method, to begin with, each and every assets (end-users account, association, and substance) are gathered. Subsequently, for every asset, related action has to be made as follows:

For end-users profile fundamental data, for example, end-user's basic information such as feminine is to be gathered. For the collected content, after pre-preparing for example joining, stop-word expulsion and highlight determination, content grouping technique is to be utilized to anticipate the subject end-user is keen on. For collaboration asset interface quality is to be resolved. Connection quality among end-users remain determined concerning comments relationship and forward relationship.

One of the difficult assignment with latest SNS Platform is the way by which to recommend appropriate friend to an end-user. A large chunk of them depend on upon adequately existing end-user associations to pick candidate. For example, Facebook rely upon a social association analysis among the people who at now share comparable companions and recommends proportioned end-users as likely companions. Seriously, this technique may not be the most appropriate one which is dependent on friend discoveries. This technique experience the downside of passion mismatch and it is futile to extend the clan of the individuals, since somebody who has numerous basic companions with you likely definitely known to you. As indicated by these investigations the guidelines to gather individuals include:

1) Life style Habits

2) Attitudes towards life

3) Tastes of tongue

4) Moral standards towards other

5) Level of economics and growth; and

6) Already known People.

Obviously, rule #3 and rule #6 are the standard components considered by existing suggestion frameworks.

(Chu, Wu, Wang, Chen, & Chen, 2013)Presented a way to deal with structure and actualize a friend suggestion framework on informal communities by utilizing Voronoi graphs and one's inclinations towards interpersonal organizations. Initial phase is to collect the settle duration near specific area situated by GPS along with user interest information from SNS platforms. Once the required information is collected, the next stage is to figure out the information with our proposed approach, weighted Voronoi graph and feature closeness. Subsequent stage is to build the affinity chart and matrix. As per the available matrix, the one of a distinctive item is utilized to suggest particular friend from the individual's companions.

(Naruchitparames, Gunes, & Louis, 2011)Proposed a friend proposal framework dependent on and metaheuristic technique and graph topology. There are different properties like area, date of birth, cast, language, choices, training that extricated using profile of end-user. There are 2 screening procedure utilizing FOF topology along with Pareto optimal genetic procedure. It will remove unnecessary individuals with the help of complicated graph theory before applying genetic algorithm. The characteristics which follows fellowship criteria are separated towards end-user profile. A social diagram are made inside such that nodes represents end-users. At that point channel dependent on FOF is utilized to diminish number of potential companions. Subsequently those companions are browsed the diagram that have more outlinks and wellness esteem is found for every one of the companions. The arranging in diving request of wellness esteem is finished. Top ten outcomes are given which will be appeared as prescribed friends.

(Yu, Pan, Tang, Li, & Han, 2011)Proposed a friend proposal dependent on physical setting. The physical setting depends on gatherings and experiences here. The technique utilizes the observation that end-users who meet in gathering can be suggested as companions. It will help the meeting participants to all the more likely lead their timetable and amplify their interpersonal organization. It builds up a companion proposal framework which utilizes nearness and homophily. Vicinity characterizes physical setting dependent on gatherings and experiences. The correspondence among various end-users caught by using one of application Find and Connect. That further utilizes area furthermore experiences information, concerning towards gathering essential administrations so as to catch the end-user connections. The loads are doled out for each property utilizing nearness and homophily. At that point the pertinence vector is evaluated for every end-user and furthermore suggestion score is being figured for every end-user. At that point top N end-users with the most noteworthy score will be suggested. Bit of leeway means that here proposal instrument dependent upon physical setting is efficient, useful than FOF approach.

(Kwon & Kim, 2010)Whereas proposed companion suggestion which is utilized in setting mindful applications. These setting mindful frameworks give the end-user versatile proposals from accessible enormous data. The proposal technique utilizing setting. A difficult research issue in social processing is the suggestion technique utilizing setting. The creator proposes a companion suggestion technique utilizing pair concerning physical furthermore social setting. Principal thought concerning recommended technique are comprised regarding 3 phases; inside initial step that further figures companionship values dependent upon comparable conduct utilizing physical setting. For processing score, conventional ordinary data recovery technique, BM25 weighting plan is utilized. Besides, a social setting is utilized in which the strategy processes companionship score with companion connection in the fellowship chart. Finally, the entirety of the determined kinship scores are joined and afterward suggest companions from positioning of the scoring esteems.

(Bian & Holtzman, 2011)Proposed MatchMaker, a communitarian separating companion suggestion framework dependent on character coordinating. The objective of MatchMaker is to use the social data and common comprehension among individuals in existing informal community associations, and produce companion suggestions dependent on rich logical information from individuals' physical world cooperation’s. Go between permits end-users' system to coordinate them with comparative TV characters, and uses connections in the TV programs as equal examination framework to recommend to the end-users companions that have been casted a ballot to suit their character the best. The framework's positioning diagram permits dynamic enhancement for the character coordinating accord and progressively different fanning of end-users' informal community associations.

(Wang, Liao, Cao, Qi, & Wang, 2015)Presented Friendbook, one novel semantic-based mate approach framework concerning free societies, which further appoints partners through customers addicted over their behaviours concerning living somewhat than social charts. By exploiting sensor-rich cell phones, Friendbook finds ways of life of end-users from end-user driven sensor knowledge, gauges likeness of various approaches towards life among end-users, and designates partners toward consumers if their styles regarding living possess tremendous similitude. Roused by content mining, tried to model end-user's every-day life as life reports, of which his/her habits regarding growth imply disengaged through appropriating significant Latent Dirichlet Allocation(LDA) calculation. We moreover recommend some similitude metric to assess comparability regarding ways concerning life amongst end-users, furthermore determine end-users' impact essentially considerably since ways concerning life including each associate correlating picture. Subsequent holding some solicitation, friendbook restores each summary concerning people including most utmost eminent recommendation numbers over this examination consumer.

(Zhang, Liu, Ding, & Huang, 2015)Proposed another companion suggestion with end-user's data of all out characteristics (FRUITA) and coordinating informal organization end-users' traits with the law of complete likelihood. FRUITA can be handily stretched out to suit new arrangement of end-user qualities in various interpersonal organizations. FRUITA calculation assessment is contrasted and other cutting edge FoF calculations, including Common-Neighbors calculation, Jaccard calculation and Adamic/Adar calculation utilizing genuine information.

(Silva, Tsang, Cavalcanti, & Tsang, 2010)Proposed a friend proposal framework for informal organization dependent on the topology of the system diagrams. The topology of system that interfaces an end-user to his companions is inspected and a nearby informal organization called Oro-Aro is utilized in the investigations. Proposed calculation break down the sub-diagram created by an end-user and all the others associated individuals independently by three level of division. Be that as it may, just end-users isolated by two level of partition are possibility to be proposed as a companion.

Title	Friend Recommendation Technique	Basis of Similarity found	Remarks	Merits	Demerits
Friendbook: A Semantic-based Friend Recommendation System for Social Networks [1]	Probabilistic	Lifestyles and activities	Proposed model which recommends friends based on lifestyle and has fixed threshold for friend matching graph	It extracts lifestyle from user centric data collected from sensors on smartphone	It uses fixed threshold factor in friend matching graph
Friend Recommendation through personality matching and collaborative filtering [2]	Collaborative	Rating given by friends	Application limited to TV shows.	It uses collaborative filtering for friend recommendation system based on personality matching.	It uses personal profiles from social networks for retrieving information of TV characters.
Friend Recommendation using physical and social context [4]	Context	Attributes from user profile	Used in context aware application. Method to extract context based information was not proposed.	It finds friends to satisfy user's present context.	Physical and social context is not clearly defined and how the information is extracted. Also this approach does not consider social interactions of the user
Trust Enhanced friend recommendation [7]	Collaborativ	Attributes from user profile	Solves the sparsity problem	Weights are calculated by real value which improves performance. Alsoit deals with sparsity issue in collaborative based friend recommendation	Limited to user's profile information.
Friend recommendation for Location based mobile social network [5]	Context and Content	Location similarity	Do not capture user's interest similarity	It uses the concept of real time location and dwell time	It failed to track activity of user in a location
Friend recommendation using proximity and homophily [9]	Context	Proximity and homophily	Provides the reason why a user is recommended as friend	This recommendation mechanism based on physical context is better than FOF approach	It supports only indoor activity.

Table 2.2 Comparison of Different Friend Recommendation Approaches

Chapter Three: Social Community

This chapter gives the overview of Community, difference between Offline Community as well as Online Community. It also present an overview of attributes along with the principles of Online Community.

3.1 Introduction

In today’s era, the Internet has become a necessary means of communication in everyone’s life. Human being communicate through video call, VOIP Call, WebEx meeting , doing online shopping , ordering food online and much more. In Year 2019, as per estimate, total number of people using Internet was around 4.13 billion. In schools, students have started learning computer and using internet in the primary classes itself. The young generation have taken internet as granted and there is no difference in internet and television as now Smart TV’s are already in use (Turow & Kavanaugh, 2003). With the introduction of smartphones, usage of Internet has increased dramatically. People used to surf internet, check mails, online banking, ticket booking etc. Internet is expanding all through the world with the development of 3G/4G cellular technologies and Wi-Fi availability (Geer, 2000, p. 11). All over the world, China has maximum number of Internet users followed by India and USA. Even though some of the European countries which didn’t have the legacy communication infrastructure but even then they were able to successfully catch-up the cellular technologies (Markoff, 2002).

(Leckenby & Li)Defines that the Internet gives the way to "communicate" between producer and consumer. The communication includes connecting Human with other Machine, Machine with other Machine and Human with other Human. Number of people on web continues flaring up and end-users present on World Wide Web continues increasing quickly and transmit all over the world. The utilization of web is no longer constrained to those PC enthusiastic who do it for the sake of entertainment or interest. Web is now a fundamental and important part of everyone’s life whether at home or in office. The amount of time and rate of usage of using internet in a day by individuals is getting increasing. The current end-users keep on utilizing it to an ever increasing extent.

(Digital Market overview India Understanding the scale of change of online audiences and digital media in India)Explained that web is an entry point to world information along with a huge stage for national media, encyclopedia, knowledge, information. Researcher further revealed regarding user’s usage of online network in our country. In country like India more than 66% of individuals use it a few times each week or more. This utilization of web changes purchaser habits for all time. The business condition is likewise changing quickly as the need to impart and keep up a relationship with people/groups geologically at various areas is expanding. People using internet for work purpose accounts for 30% of total online population on day by day basis where electronic mail among mostly utilized apparatus or tool or method on web internet. Organizations consistently need to reach out where larger chunk of population is present and pass on the correct, authenticated, unaltered and true information to user using it at due duration in specific correct manner. Earlier e-commerce companies have created their own websites or webpages where end user get to know about the information about the companies and the items they used to sell however that type of connections are unidirectional not bidirectional. The web page is refrained either by single individual or a company with the end goal for essentially forwarding data to end-users. Hence, to achieve two way correspondence Social Networking platform

3.2 What is Community

The word community has a very long historical background. One researcher found 94 unique definitions, and that was over 50 year's back (Wickizer & Donald, 1995); but now there are definitely a lot more definitions available today. Since the late 90th century, 'the utilization of the term community has stayed somewhat connected with the expectation and the desire of restoring again the closer, pleasant, increasingly agreeable kind of bonds between individuals' (Hoggett, 1997). It's only in 1915, the first definition about Community has appeared, before that there was very less literature on this. Various terminology of community immediately followed. Few considered community as a geological region; some considered it as a gathering of individuals living in a specific spot; and others which looked to community as a territory of normal life. (Gusfield J. R., 1975)Recognized two significant benefits of the term community. The first is the regional and topographical idea of community - neighborhood, town, and city. The second is related about "nature of character of human relationship, without reference to area". (Gusfield J. R., 1975)Noticed that the two modes are not fundamentally unrelated, in spite of the fact that, as (Durkheim, 1964)watched, present day society creates network around interests and aptitudes more than around region. (McMillan & Chavis, 1986)Suggested Community definition has 4 components: Membership, Influence, Integration and fulfilment of needs and shared emotional connection.

(Cole & Knowles, 2001)Characterize community as 'Groups of individual lives make-up community, social orders and societies. To see a portion of the complexities, difficulties, and disarrays inside the life of only one individual from a community is to pick up experiences into the system'.

(Lee, 1992)Characterizes community, essentially as a gathering of individuals who share something for all intents and purpose. For (Boothroyd & Davis, 1993) a community is 'A arrangement of multiple individuals wherein the individuals communicate with each other as time passes, where conduct and action are governed by unified judgement, and from which individuals may leave freely'. (Roberts, 1979) Considers it as 'a group of individuals who are familiar with some issue or some big objective, who have experienced a procedure of finding out about themselves and about their condition, and have figured a gathering objective'.

At the point when individuals are inquired about some information what 'network' signifies to them, it is such systems that are most normally referred to. 'For the majority of us, our most profound feeling of having a place is with our most close informal organizations, particularly loved ones.

(Rheingold, 1993)Describe Online Community as universal collections that came out of the web when vast amount of individuals communicate openly as it's a feeling of personal touch, to build individual contact on the internet. He described the Online Community as moving non-commercial and social inclined where individuals shares their desire, information, and their involvement with one another. (Fernback & Thompson, 1995)Characterized the online community as fake social connections on the internet through reiterated communications with in a certain limits. (Romm, Pliskin, & Clarke, 1997)Described Online Community as a cluster of individuals that connect with one another by means of digital media and having same passion without any constraint by topographical region, mutual collaboration, or any race. (Hagel & Armstrong, 1997)Defined Online Community as associations of individuals with regular benefits and specifications who comes as one to support each other. The majority of them are attracted to have a chance of the grouping with individual having similar mindset irrespective of wherever they live. Online Communities are not confined for social interaction but something beyond that can be possibly stretched out for professional purpose as well. Community which is formed with members having similar passion can be seen as potential business opportunity for the companies. Group members can also share data on the item's cost and quality. This implies Online Community is stronger than online networking. As per (Lazar & Preece, 1998)the program enabling Online Community are catalog server, newsgroup, announcement board, web transfer visit (IRC), or Multi-User Dungeon (MUD). These computer programs helps in transferring information within the group and confined within the limits of a group.

(Naruchitparames, Gunes, & Louis, 2011)Explained Online Community as an accumulation of individuals. Individuals are attracted to Online Community because they help in providing the platform of trust and understanding where individual can join in ongoing series of conversation. Virtual Publics, as mentioned by (Jones & Rafaeli, 2000)are human communication which happens through means of electronics devices which is clear and crystal clear which helps the users to interact with in a group. Online Community comprises of individuals, who connect socially as they endeavor to fulfill their own needs or perform extraordinary jobs, for example, driving or directing a common reason, information sharing, or administration that enables the community to follow some ethics, ceremonies, conventions (Preece J. , Online communities: designing usability, supporting sociability, 2000).Online Community composed of individuals that involve themselves in social networks that include both the creation and utilization of beliefs and suggestions. It shows the specified attributes; (1) a group of individuals, (2) normal individuals, (3) collaboration in the internet by not doing physical attachment, (4) communal exchange procedure (5) a mutual goal, character, enthusiasm among individuals (Balasubramanian & Mahajan, 2001). (Porter, 2004)Defined Online Community as a conglomeration of personalities or colleagues that communicate with each other with a common purpose, where the communication is either partially bolstered or potentially reconciled by technology and guided by sound standards. Consequently, we can describe that community individuals communicate with one another over interconnection, and their involvement is driven by basic reward, reason and essence.

3.3 Offline vs Online Communities

Despite the fact that the meanings of offline and online community are comparable, it is not necessarily the case that Offline and online community are subjectively the equivalent. One case of this distinction is that numerous online community appreciate scholarly actions more effectively as compared to offline community does. This aspect is likewise shared by (Hauben & Hauben, 1997) who note that online individuals are urged to be keen and to introduce their plans to the others. They should have some passion or interest. Another instance is that members of some online community may just communicate to one another through their PC screens or smartphones while members of some offline community may have more open doors for chance gatherings in the city, in broad daylight places, or through various correspondence advances. In Online communities, members have less chance to meet each other on the road in because they are physically separated sometimes by big distance and another reason is they don't know their physical appearance. The last instance that shows the differentiation factor between two types of community is that in case of offline community, individuals are able to know whether other members are taking interest in the discussion or not; it’s challenging in case of online. Besides, in case of online communication, discussion is not only more stretched as compared to offline but disintegrated as well (Kozinets, 2001).

Sproull & Kiesler (1991) expressed that offline communication has more advantage over online communication as it involves direct communication. For instance, tone in the voice, signals, dress, tone, pose and different hints are absent in online communication. Chidambaram and Jones (1993) acknowledged that in case of online community, communication depends majorly on the text based mode but nonverbal means of communication e.g. emoji’s and Internet language is utilized for communication (McLaughlin, Kerry, and Smith 1995).

(Sproull & Faraj, 1997)Expressed that geographical position not important for member’s presence in Online Communities. They expressed that most members in online community are generally unseen, hence it leads to a lower strategic and social expenses to take part in it. (Hiltz & Wellman, 1997) Indicated that Online Communities are nearly bigger, increasingly scattered in reality and individuals have progressively heterogeneous social qualities, for example, lifecycle stage, sex, ethnicity and financial status, however with increasingly homogeneous perspectives.

Some researchers has described Online Community as Virtual community.

3.4 Attributes of Virtual Communities

The literature proposed 5 characteristics used to define virtual networks: (1) Purpose, (2) Place, (3) Platform, (4) Population Interaction Structure, and (5) Profit Model. With different sequence of attributes, virtual communities have higher chance of success both in terms of members as well as stakeholders.

The suggested typology depends on both literature review as well as on the action taken place in actual virtual community. The methodology used to build up the proposed typology grasps (Hunt, 1991)description of an inductive gathering strategy to develop classifications. That implies virtual community belongs to certain category are going to have typical attributes. Yet, no virtual community possess all the attributes belongs to that particular class. So in a nutshell, to define any virtual community, attributes mentioned in following section would be required.

3.4.1 Attribute #1: Purpose (content of interaction)

Virtual community should work when and only when there is some sole purpose among the members (Gusfield J. R., 1978). It’s sure that formation of virtual community is possible by taking in to account some to the attributes from the limitless number of attributes. In fact it’s similar to the concept of “discourse focus" (Jones & Rafaeli, 2000) – basic guidelines for communication in virtual community (for e.g., playing cricket, staying with diabetes).

3.4.2 Attribute #2: Place (extent of technology mediation of interaction)

Since Virtual community is not confined to certain boundary so Attribute place play an important role (Jones Q. , 2004) as compared to offline communities which are restricted to some geographical location. Within a boundary, such interaction drives individuals to feel a perception of belongingness and sympathy. In this manner, the meaning of community signify both something basic (for example, a limited area) and something socio-mental (for instance a feeling of shared qualities created through connection with individuals).

(Harrison & Dourish, 1996)Outlined both the structural and socio-cultural properties of virtual community. Structural property deals with the space whereas Socio-cultural deals with the community place. They related virtual space to house and virtual place to home. In reality, house is nothing but a home having family members in it which are bound by some socio-cultural relationship. Radically, a home is a house in so much individual’s grasp that are socio-social obligations of associations between family section individuals. (Harrison & Dourish, 1996)Further suggested that house is a physical entity that helps in development of home. In similar manner, virtual space gave a chance for virtual place to flourish.

In spite of explanation on the terminology given by (Harrison & Dourish, 1996), different researchers ought to not completely embrace their view or terminology. In fact, the concept of space and place are generally consolidated in the record. For example, (Blanchard, 2004)implies that a society member's anticipate “sense of place” is determined by perceptual ideas in that virtual conditions (e. g. type of access, the timing of interaction, and membership boundaries). She recommends that community members use that type of hints to ascertain wherever community intercommunication happens, wherever people “are” in the flow of communications among different layers, and whether different members comprise present. She argued that individual society member has a “sense of place”, even if there is any difference in the opinion of among community members.

(Blanchard, 2004)Uses the other idea of "sense of place" than that utilized by (Harrison & Dourish, 1996). (Blanchard, 2004) Notion of "sense of place" is one that depends on a member cognitive familiarity with the area or availability of other members in a definite area. Hence, it is more logical as compared to suggested by (Harrison & Dourish, 1996) .

Other researchers queried the suitability of the difference amongst "space" and "place". Some advised that communication in virtual community isn't restrained to either virtual or just physical (Rothaermel & Sugiyama, 2001). These researchers advocate that the internet is thoughtfully encapsulated in physical space (Mitra & Schwartz, 2001) and interaction between individuals is confined by both physical and virtual space. Besides, the way that individuals in virtual community uses different methods of interaction, including eye to eye, phone, and mail. (Blanchard, 2004) Proposed that virtual and physical networks can exist together. In aggregate, (Wilson & Peterson, 2002) suggests that “the distinction of real and imagined or virtual community is not a useful one. “A virtual system is characterized, subsequently, by its “sense of place”, as mentioned by (Harrison & Dourish, 1996).

While that differentiation among "place" and "space" may demonstrate necessary for specific kinds of research, but for this type of classification, the "place" takes in to the account of "degree of virtualness" (Virnoche & Marx, 2007). (Virnoche & Marx, 2007)Proposed a location-based methodology by recommending that communities can be classified depending among degree that community individual’s take part in offline or online space. In offline community, where people communicate with each other in physical space, numerous virtual community consist of individuals that generally communicate in virtual space and occasionally in physical space as well.

(Virnoche & Marx, 2007) Characterize virtual expansions as genuine, truly based connections that are stretched out into virtual space. Subsequently, the location is being visualized as ought two levels: (1) Composition of both physical and virtual space (2) virtual space only.

3.4.3 Attribute #3: Platform (design of interaction)

Synchronicity is an important term used in the interaction in virtual community. Synchronicity is extent to which the system enable communication in real time (Hoffman & Novak, 1996). Interaction can be either synchronous or asynchronous. For instances, messaging applications comes under the category of synchronous communication whereas emails comes under the umbrella of asynchronous communication. However, email-based system permit individuals to view and reply to communications whenever it might suit them instead of in the right time (i.e., novel collaboration) (Preece J. , Online communities: designing usability, supporting sociability, 2000; Blanchard, 2004).

Since interaction plays important role in the synchronous technologies so it leads to the research in area of interaction, especially in electronic situations (Blanchard, 2004). With regards to PC based interaction, interactivity is formalized as "dependency among messages in threads" (Rafaeli & Sudweeks, 1997)calculated based on the length and depth of the interaction thread (Rafaeli & Sudweeks, 1997; Preece J. , Sociability and usability in online communities: Determining and measuring success, 2001; Whittaker, Terveen, Hill, & Cherny, 1998).

Synchronicity may be necessary for virtual community, given that individuals overwork the synchronized innovation arrangement by interacting (Blanchard, 2004). It furthermore can foster the advancement of social reality for individuals (Rafaeli & Sudweeks, 1997). In any case, in virtual communication, “interactivity is made possible, but not always exercised” (Rafaeli & Sudweeks, 1997).

3.4.4 Attribute #4: Population interaction structure (pattern of interaction)

Even though there is no specific classification which explains the communication patterns between virtual community members but three research related to this explained this issue: (1) virtual communities as computer-supported social networks (CSSNs), (2) virtual communities as small-groups or networks and (3) virtual publics versus virtual communities.

3.4.5 Attribute #5: Profit model (return on interaction)

As the name suggest, this attribute focus on whether the community makes some profit or not. Even though it is applicable to communities which are commercial in nature but there is a chance that other non-commercial communities will have economic aspect. For instance, non-commercial community may invite some advertising agency to put advertisement in their community. This is a win-win situation for both community and advertisement agency as community can earn some money where as advertisement agency can have access of community member's details.

3.5 . Principles for Online Community

The 12 Principles – Purpose, Identity, Reputation, Governance, Communication, Groups, Environment, Boundaries, Trust, Exchange, Expression, and History – were developed by Cynthia Typaldos of Real Communities (now part of Mongoose Technology, Inc.) (Typaldos, The 12 Principles of Collaboration, 2000). "Individuals usually come together to form communities in which they recognize common purposes, values, and visions." (Figallo, 1998)

3.5.1 The First Principle: Purpose

As per Figallo (Figallo, 1998), a "community should be a practical and useful thing for people to join." There should be some purpose or interest to join Community to let people achieve something and to contribute something. Community may fail if there is not enough adequate reason for members to come together. Individuals have to focus about the community purpose rather than individual purpose. Community will grow only when all the members will come together to accomplish purpose and achieve common goals. To have the community in existence, set of Application tools (for example, chat rooms) must be required to implement the same and they should have “application of purpose”.

3.5.2 The Second Principle: Identity

The challenge with the online community is to know individual identity because the base communication is online rather than face to face communication in case of offline community. There must be some ways in which members will be able to know each other in online community. One way is to have the user profile in which members will provide all the required information so that other members can able to see whenever required. Other way is to have dynamically generated profile based on the user behaviour with the system.

3.5.3 The Third Principle: Reputation

Members need ways to know how dependable or proficient another partner is. Reputation lies at the point among character and trust and impacts conduct in a few different ways. Reputation estimates give individuals an approach to assess one another, so they realize whom to trust, or whom not to trust. It assists individuals with framing the best unions to get the ideal data; the craving to have a decent notoriety debilitates terrible conduct and urges individuals to demand criticism from others to develop their reputation. Members create reputation dependent on their words and their activities. This turns into a basic component of character. As individuals invest more energy in the network, they too leave "trails" that record their conduct — where they go on the web, what they take a gender at, what exchanges they attempt. As databases get developed that store this information, it turns into another sort of reputation.

3.5.4 The Fourth Principle: Governance

Community governance is nothing but an extension of individual governance. Given the size of enormous and well known community networks, it is significant that individuals assume liability for their own conduct and have the authoritative instruments to govern themselves in various circumstances. Individual governance is increasingly effective, scalable, and network agreeable than describing violations to the community administrator who at that point goes about as the community police. Clearly, there is a key connection among reputation and administration: the better your reputation, the more state you have in how the community runs.

3.5.5 The Fifth Principle: Communication

For a community to remain, individuals must have the option to connect with one another. Any sort of communication, regardless of whether it happens online or offline, includes correspondence or something to that affect. (MacIver, 1937)In his investigation of real community, expressed that, “Without communication there can be no community, and the life of the community revolves around the points where communication is most intense”. In communication networks, how individuals connect is a main concern for the community administrator as well as individuals. The success of community is due to availability of variety of communication tools available which caters to different need related to synchronous or asynchronous communication. Picking the right tool depends on the context as well as the individual discussion mode.

3.5.6 The Sixth Principle: Groups

Based on the individual taste or interest, Individuals divide themselves into different group which caters to their needs. Bigger community leads to more number of groups which accomplish its behaviour. In Online community, this division of grouping and sub grouping should be regarded and encouraged. Technology that empowers people to make, join, and partake in respective purpose related group is vital to making community increasingly helpful and convincing. Online people group must empower an association of individuals to have a gathering personality, a gathering place with clear benefits, approaches to direct gathering rules, and access to instruments to execute joint reason.

3.5.7 The Seventh Principle: Environment

Even though Online Community doesn’t have the boundary constraints but like the offline community, the environment impact the members experience. Each community is unique and capacities best in a suitable situation, custom-made to its explicit needs and style. As in offline communication, meeting spaces share certain attributes that make all member’s more inviting and valuable — obviously entry doors and ways out, access to stopping furthermore, transportation, enough space for everybody — in a similar fashion effective sites must give a pertinent and predictable experience for their members. Few of the key components for an effective, synergistic communication on the online community includes seamless experience, flexible interface and lot of contents.

3.5.8 The Eighth Principle: Boundaries

There must be clarity on who can be or can’t be the member of online community. To make online community successful, there must be some boundaries or regulations like non-member access to community or make it password protected or some other access control to the members. It is not limited to determine who is member or non-member of the community but can be extended in various ways. For examples, members who have joined community long time back can be given some more privilege to them.

3.5.9 The Ninth Principle: Trust

Trust is very important factor for the success of community. Individuals must have the option to tell whether — and how much — they can confide in different individuals. What's more, everybody needs to believe that the individuals who run the community won't misuse or adventure their own data. Trust is worked after some time and should be earned both by individuals and by administrator. Researchers have found that with numerous positive communication, deeply understanding about individuals and agreeing assessments of other reliable individuals are critical to picking up trust in others. Group effectiveness increases once trust is build and it helps in resolving the conflicts as well. In case of online community, there are two types of trust, one is between the members themselves and other is between members and coordinator. In an online network, individuals need to believe that their private data is protected and that nobody can copy them. As far as possible, the act and outcomes of unveiling individual data ought to be under the individual control. This facilitate trust between the individuals and the site coordinators and facilitators.

3.5.10 The Tenth Principle: Exchange

The purpose of individual to join community is to exchange the information with other members and get relevant knowledge from other members who are expert in their domain. Information exchange can be in form of knowledge, wisdom, past experience. This exchange of information can be one to one or one to many.

3.5.11 The Eleventh Principle: Expression

Expression talks about what's occurring at some random time, who's hot, what themes are being talked about, which articles are being perused, who's arriving at his/her objectives. This is particularly significant for first-time individuals. Except if they see a significant level depiction of community action, they may presume that the community has nothing to offer them, that nothing's going on. In one to one confrontation, “expression refers to gestures, facial expressions and vocalizations,” on the other hand communication, which is “the use of language ... for the intentional transmission of a ‘message.” (Mark, 1998) Offline communication, both expression as well as communication takes in to consideration where as in online communication, only communication need to be considered.

3.5.12 The Twelfth Principle: History

A feeling of history is imperative for developing the community network. There must be some ways by which people remembers the community for a long time. Database of information exchange between the community members helps in information availability to other members as well. Best example of this one is the open source development forum whereas members develops some projects which can be used in future use as reference by some other members. A community also try to recollect its members, users, participants, even when they go out of touch. Identity of members can be firm; the society have found a way of resembling and remembering who that individual is and what action or play or role he or she did in the past.

Chapter Four: USER PROFILE CONSTRUCTION USING ONTOLOGY

This chapter will focus on the Ontology, what Ontology is, how ontology is represented, Ontology Description Language and Ontology Modelling. We introduce the Hobby Ontology and develop using Protégé software which is the being used in the recommendation system framework.

4.1 Introduction

Modelling the semantic knowledge is particularly one of the main purposes for the usage of ontologies. This section provides very detailed review of ontology including ontology definitions and a description of the ontology development process. It then discusses the application of ontology to data extraction and integration and the use of ontology in recommender systems.

4.1.1 Ontology Definition

As defined in the below section by (Noy & McGuniness, 2001), an Ontology is defined as a “formal explicit description of a domain, consisting of classes, which are the concepts found in the domain (also called entities).” Ontology is a data model that represents knowledge as a set of concepts within a domain and the relationship between these concepts. Ontology is a form of knowledge management. It capture the knowledge within organization as model. This model then can be created by users to enter complex questions and display relationship across enterprise. The word Ontology comes from two Greek words. One is “Onta” means existence or being real. Second one is “Logia” means

Science or study. The word Ontology is used both in philosophical and non-philosophical context. In philosophy, ontology is study of what exist, in general. Philosopher use the concept of ontology to discuss challenging questions to build theories and models, and to better understand the ontological status of the world. In Non-philosophical context, Ontology is description of what exist specifically within a determined field. This includes the relationship and hierarchy between these parts. Researchers are focused on naming parts and processes and grouping similar ones together in to categories.

Vast volumes of data can be found upon the internet and web (ICQ, n.d.). Consequently, proper information extraction and searching may not be a simple job. Hence, an efficient and robust method that seeks to establish and reclaim related data is very essential (Aula & Nordhausen, 2006). More information concerning user's requirements is comprised of the accelerated development of records accessible from the World Wide Web. Nevertheless, an immense measure of knowledge leads to finding a particular data is a tiring job. Due to low precision of web search engines, typically, remarkable important web pages are then delivered in addition to that a huge amount of unrelated pages are joined due over topic-specific articles that might happen in unusual circumstances. Accordingly, a relevant platform that can manipulate the large amount of reports on the internet is required.

Today people used to access more data in single day then most people used to access in their lifetime in previous decades. The problem is data is available in different forms. All these information captured in different format makes it impossible to understand existing relationship between data. Data needs to be represented in a format that allows relationship between them to be discovered. Ontology captures data in a way that allows relationship to become visible.

Two standards which governs the construction of ontology are Resource Description Framework (RDF) and Web Ontology Language (OWL).In accordance with RDF and OWL, an ontology is made up of two main components: Classes and relationship between them. Ontologies are easily extendable rather than writing new lines of codes means new relationship can be easily added to an existing ontology. Within Ontology, concepts are only defined in terms of their relationship to other concepts. The most popular aspects of ontology is in addition to capturing relationships is that using ontology for knowledge management is an alternative to source code. Ontology present a new method in managing knowledge and capturing relationship.

The very fundamental description of “ontology” in field of computer science was given by Gruber (Gruber, 1993)as an “explicit specification of a conceptualization”.Borst further describe ontology as a “formal specification of a shared conceptualization” (Gruber, 1993). (Coelho, Martins, & Almeida, 2010)Gave new interpretation of an ontology “as a knowledge domain conceptualization into a computer-process able format which models entities, attributes, and axioms”. Ontology is basically composed of vocabulary including connections among concept classes (FOAF, n.d.). As per (Antoniou & Van Harmelen, 2008), ontologies sub classes can include information such as of properties between the classes(X studies Y), limited values (only Engineering student can study this course), disjoint statements (Engineering and Art students are disjoint) specification of logical relationship between class objects. Various tools exist based on the ontology to design the system.

In Ontology, relationship is defined as how the objects of different classes are related to each other. It is the ability to describe relationship which makes the ontologies powerful and used to define the domain semantics. An essential form of relation is a “is – a” relationship which is used to describe which object belongs to which class. Many ontology division has been proposed by (Trajkova & Gauch, 2004) , for instance, ontology can specify to particular domain area that can further accommodate the various conceptual modelling of a specific area.

Ontologies can be classified into three categories - Domain Ontology, Upper Ontology and Application Ontology. The Domain Ontology as the name suggest define the vocabulary or concept belongs to some specific domain such as politics, physics, sports or automobiles. Domain specific ontology generally create terms specific to that domain. For instance, the word card can have different meanings. A domain ontology regarding poker game would model the "playing card" meaning of the word, while an domain ontology regarding computer hardware would model the "punched card" and "video card" meanings. An Application Ontology defines concepts of a particular domain and task. In the application domain, Upper Ontology, as well as Domain Ontology, can be integrated with Application Ontology. Efficient explanation about ontologies have presented their influence in several regards, as explained below:

• An ontology requires some factorization of information. In the similar manner as in the oriented objective approach, information is not replicated in each instance of a class (Rinku, Raihan, & Aravind, 2016).

• An ontology presents a collective framework that is used to overcome or eradicate uncertainties and confusion that are conceptual and terminological in nature (Daramola, Adigun, & Ayo, 2009).

• An ontology can significantly enhance the accuracy of various search engines that are being used in the approach. With the help of given semantics, an ontology can overcome problem related to lack of sensitivity and lack of specificity of the popular search engines (Ringe, Francis, & Palanawala Altaf, 2012).

• An ontology can thus support the flexible sharing and reuse of knowledge (Shvaiko & Euzenat, 2013) in this the researcher can further reuse the existing ontologies and, if they are adapted to meet with their need, will reduce the time of designing an ontology from scratch.

• An ontology then performs mechanisms of deductive argumentation, automated distribution, and knowledge retrieval moreover guarantees interoperability among various systems.

4.1.2 Ontology Representation

The essential fundamental components in ontologies are concepts classes, relations, rules, and class instances. The meaning of each and every element, as designated by (Noy & McGuniness, 2001), is further introduced underneath:

A **Concept** is the important element of ontology. In Ontology, classes are used to describe domain concept in hierarchical manner. Ontology graph can have different levels where as classes at upper level are known as parent classes whereas classes at lower level are known as child classes. For example, Hobby class can be the parent class whereas Sports and Cooking can be subclasses. Additionally, the concepts may have a wide range of recognizable properties.

A **relation** (otherwise called slot) is further utilized in the structure of ontology to provide information about the connections between concept classes in a particular domain area. To determine these two classes that are associated beside a specific relationship, super class will be depicted as a "Domain" whereas sub class is known as "Range"; for example, "Has brother" has the relationship between Domain Class “Person” and Range Class” Male”.

An **Axiom** (also called as rules) is further used in the ontology to constrain values for classes or instances. Axioms put the constraints on the intended meaning of a term by affirming required conditions for its use.

An **Instance** (otherwise called an individual) is the thing represented by a concept in the same way as an objects in object oriented language. Ontology is defined as a “conceptualisation of the domain “so generally it should not have any instance. The amalgamation of ontology and its corresponding instances is known as knowledge base.

4.2 Ontologies for domain knowledge representation

To add on the potential of modern technologies, ontologies has been incepted in late 90’s and used in the field of semantic technologies (Luke, Spector, & Rager, 1996).

Though Gruber's description (Gruber, 1993) is pervasively mentioned, that thought of philosophy has remained relatively accomplished in the application. In reality's (illustration., by the viewpoint concerning an IR scholar or researcher), ontologies remain generally organized as grouping of concepts with characteristics and relationships, that discover a specification to describe knowledge representation of ideas that are interrelated examples, representing domain particular information is managed in a knowledge base (KB). Even though ontology is comparable to dictionary but dictionary regularly consists of many pre-determined collection that are type's relationship, whereas on the other side, ontology is more flexible, can be related to some particular domain, extensible. In a nutshell, dictionary is considered as subset of ontology whereas ontology is considered as more powerful and rich.

Then again, ontological Knowledge Base has the capability (however not generally) to keep a lot of information, with a lot more immeasurable degree of feature which is not possible in case of dictionary. Hence we may state further these all Knowledge Base then considered among a transitional point between a database and a thesaurus. Ontologies are more formal as compared to Thesaurus and use to describe some particular domain concept using classes and their relations, properties and descriptive logic between them. Even though the amount of effort and money spend on development and to maintain it, it's worth considering the potential output. This outcomes in significant favorable condition for the improvement of impressive questioning and reasoning system. In return, the issues faced in this direction are notable, as the problem to formalize regular information, even in the little bits, is noticeable.

In the late nineties, the great innovative progress related to (standardization, principles, editors, APIs, reasoners, and so on) in the field of semantic based area helped in the development, misuse and upkeep of ontologies and Knowledge Base. Also regularisation of framework can be a significant advance towards the reuse of advances and assets as the ones identified with the utilization of thesaurus in IR.

It is very essential for consulting companies to sufficiently understand the requirements, preferences, and practices of their end-users to ensure that their services are very well conveyed to the correct individuals at the appropriate time. Nonetheless, accomplishing such an understanding of the end-user, in light of a broad scope of related aspects and incomplete information is a mind-boggling investigation activity. The end-user's present circumstance, preceding history, and social condition should be joined and organized. Information about the time and actions of end-users ought to be connected with the end-users' past data to understand their present behaviour; previous actions and information ought to be considered to interpret and comprehend current circumstances; relations with others and additional end-user conduct in same situation ought to be additionally considered and grab. With the development of the Social Web and interpersonal interaction territories, for example, Facebook, Twitter, Google+, and YouTube, an enormous measure of individual data is made regularly. The size of this individual and social setting information has a gigantic potential to advance the inclusion of end-user displaying approaches and upgrade the viability of versatile frameworks.

Multiple different attempts have been made by the Semantic Web (SW) group to focus on this issue. Vocabularies in well-defined form, for example, RDF and OWL, have been created to demonstrate end-users and their social ambience. Instances of these vocabularies incorporate FOAF – Friend of a Friend (FOAF, n.d.), OPO – Online Presence Ontology (Online Presence Ontology, n.d.). While certain ontologies do, for sure, seize end-user behaviour inside online networks, they just capture user static behaviour but don't capture end-user changing behaviour. The above mentioned vocabulary represent the very basic, uncooked information. However, important information introduces from percolating the dictionaries, picking useful information, and tapping the profile information to discover end user actions, requirement and inclination.

The advancement in the computation power and storage capacity has opened the door of innovation in terms of data digitalization and availability of large chunk of information available. The desire of having efficient information availability has driven the formulation of the ideas of the semantic web and PIMS. In order to understand the end user behaviour, user profiling is being required and modification is required to cater the user's requirement.

Ontologies have been certainly proves as a successful method to understand user behaviour. They are considered as an effective framework, as it helps in finding some domain area based on the end user activity and helps in query processing. At top level, Ontologies helps in modeling concept and connection between the domain classes whereas system do the processing and inferencing. To develop end user profile by utilizing an ontology has just been proposed in different applications like web search (Trajkova & Gauch, 2004; Lawrence, 2000)as well as in PIMS (Katifori, Poggi, Scannapieco, Catarci, & Ioannidis, 2005). Notwithstanding, as yet, ontologies demonstrating end user profiles are application-explicit, with each one having been made explicitly for a specific user interest. Considering the proceeding with consolidation of ontologies in new applications, there is a rising requirement for a standard ontology system that will demonstrate end user profiles; this standard philosophy will encourage the correspondence among applications and fill in as reference moment that profiling functionalities should be created. Ontologies as progressive systems of end-user interests have been proposed in (Trajkova & Gauch, 2004). (Gauch, Speretta, Chandramouli, & Micarelli, 2007) Additionally suggested a framework having end user profile based on weightage of different concepts which helps in browsing. The end-user has the privilege to create his/her own user profile. End-user profile can be portrayed based on ontology model. (Razmerita, Angehrn, & Maedche, 2003) Presented a framework based on common ontology to be used in the field of Knowledge Management System.

4.3 Ontology Creation Resources

For creation of the ontology, we have followed a top down approach rather than bottom-up approach during ontology creation process; firstly, by choosing only the very basic concepts that were subsequently enhanced and modified by creating child classes. During the creation process we followed design criteria given by Gruber’s (Boyd & Ellison, 2007) that include consistency, flexibility, simplicity, conceptualization at knowledge level, minimal ontological commitments.

4.4 Ontology Modelling

The OWL which is mainly known as the Web Ontology Language is concerned with schema of Ontology and vocabulary based on ontology. The design method recommended by (Noy & McGuniness, 2001) is used to generate Ontology model. To establish the ontology, (Noy & McGuniness, 2001)has suggested various seven steps those as follows:-

1. Define the "domain" and "scope" of the ontology: Purpose is give personalized suggestions.
2. Existing Ontology Reuse: As the name suggest, while creating the domain ontology, if required, can reuse existing ontology instead of developing the existing ontology from scratch to reduce the development effort and time. We have incorporated existing (vCard ontology) in our proposed ontology.

   For showing information related to individual's names, different address, knowledge, contact etc. of diverse, different business companies, vCard ontology is particularly used. Same has been used to keep the end-user data which he/she enters during account creation to provide personal information.
3. Enumerate terms: In this step, Identification towards the definition of concepts and associations between these have been used to construct this structure. In the proposed hobby ontology , the terms utilized as part of the vocabulary have been taken from website "www.discoverahobby.com" that further contains various type of interest, that are further divided through uncompromising examination and

investigation which is specifically based on that evaluation abundance among various consumers or users that visited particular web-page.

4. Define class and class hierarchy: This arrangement is further categorized in three main concept classes, namely - Hobby, DBPediaMapper and Hobby_Type (extra individual vCard ontology class). These concept classes are additionally separated among furthermore in 270 concept classes.
5. Define properties of classes: The particular set of concept classes that are described under the Hobby ontology are further connected among a list of attributes, those are 58 by number.
6. Define facets of slots: It is explained in more depth in subsequent parts.

7. Create instances: Class objects of the proposed Hobby ontology are formed for the purpose of end-user to for determining particular characteristics of user.

Hobby - This group includes of each and every hobbies that a user may be involved in on the frequently basis.

Hobby_Type – Specific groups are practiced for joint hobbies that possess some well-known similar features similar to badminton and dancing.

DBPediaMapper - This type of group is practiced for mapping groups among Hobby ontology into specific groups in the DBpedia ontology.

vCard - This type of group are particularly utilized toward classification among personalities, companies employing techniques that are using semantic web and consequently collecting their knowledge.

Structuring of concepts and connections among them is implemented by Graph. Furthermore, Graph is used for visualizing things in a better and more enhanced way. Nevertheless, there is a property of graphs that they do not preserve the interpretation of the concepts they are representing, and the knowledge acquired in these social networks are not formalized. They represent knowledge better. Attachments associated with a given domain and specification of concepts are the two use case for this type of Ontologies. Entities and relationships are the two primary components of Social networks. Modeling the information that is conflicting in nature and the efficacy of the knowledge encoded by rationalizing is not allowed by Ontologies. Besides, ontologies can presume new knowledge throughout the inferring.

4.5 Ontology Description Language

In recent days Ontology specification writings have gained significant recognition. The appearance of the Semantic Web boosts it. The layered technologies of the Semantic Web is shown in Figure 4.1, the panels generating by RDF Schema shown below systematized using World Wide Web Consortium (W3C).

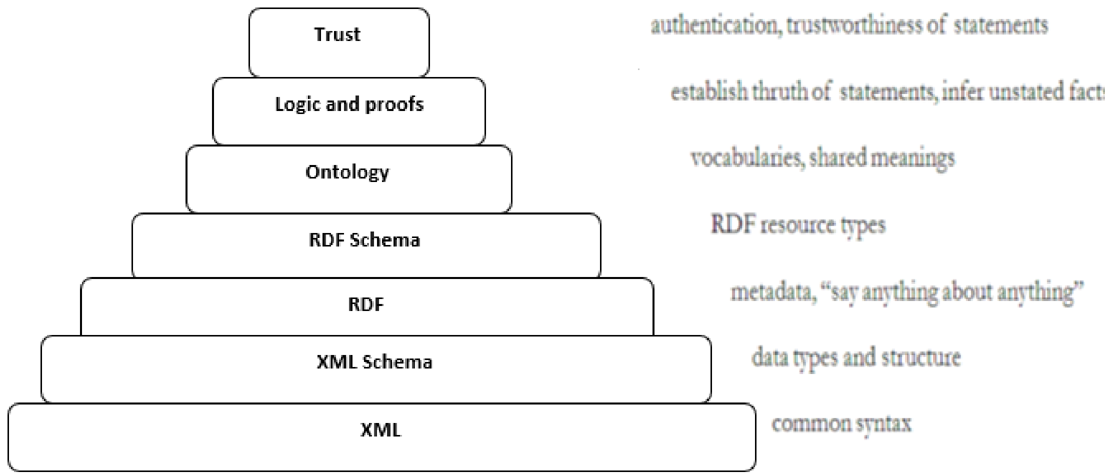

Figure 4.1 Layered technologies of the Semantic Web

The associations between various resources and ontology representation linguistics continue presented in given particular layered design. Every layer strengthening below layer. In XML formats, most information are required to be generated at base. Every layer tends to be furthermore complicated than the layers underneath it and is progressively enhanced specialized. There is no dependency of the lower layer on the higher layer. Consequently, these layers can be independently assembled with freedom.

XML (eXtensible Markup Language) is the communication structure that is in existence, from the late nineties, practiced for determining utmost recent semantics for reciprocation data overhead inside the Web.

XML framework defines type of communication that is used for structuring the arrangement of particular dictionaries based on XML. Resource Description Language is very adaptable way which includes information that is based on graph design maintaining, representing ontological meta-data, data about data in some sort of temporary expedients that are interlinked through semantic associations.

Ontology layer contains expressions for representing various dictionaries including limitations over practicing using various words including expressions inside connection particular vocabulary. Web Ontology Language that's also known as OWL is one of the before-mentioned languages.

Proof and Logic and is a layer wherever reasoning at the Logic level are in diverse used for checking the compatibility, accuracy of data-sets, furthermore for understanding current information not declared and needed by, compatible among, familiar assemblage using data.

Authentication Trust of identification and confirmation of the integrity of data, assistance, and instruments is presented by belief.

In the accompanying, we quickly depict XML, XML Schema, RDF and OWL, a few guides for mentioning associations, augmentations. Description behind interpreting communications are two-overlap. Initially, as soon as these languages are announced as W3C suggestions, these are by and large broadly used by legitimate and some semantic business applications (Rinku, Raihan, & Aravind, 2016).

XML addresses to a first way to deal with an ontology that is mainly based on the web support. XML permits organizing information and archives as structure same as trees consisting of labels including properties, Schema in XML is utilized for giving determination in trees, and the meaning by crude hence broadened types of information.

Since following the approach of XML in 1998, various measures have been distinguished for manifesting data in distinct areas, for example, business (Daramola, Adigun, & Ayo, 2009)news bulletin, or conventional insurance among numerous diverse fields. XML is an initial step to help an express information portrayal and very much characterized web structure substance, isolated (or installed among) introduction inside Hyper Text Markup Language. Notwithstanding, authentic help secured in Extensible Markup Language for the most part syntactic, by constrained expressiveness of semantic.

The Extensible Markup Language information design constitutes in form of tree construction and no requirement among articles and associations, nor is a relevant help supplied class regular systems. The primary draft of RDF was released in 1999. Moving to literature for the meaning of ontologies and data about data inside Web pages, RDF is now entirety numerous mainstream that is beyond board example inside Semantic Web people group. Required description in RDF is "triple" or sentence, which is represented by 2 nodes (item and subject) connected over a line or a link (known as predicate). Nodes depicts relationship, and the edge depicts to a property that relates the object and subject. For instance, a triple could explain the way that the “YY” which is of type student is the supervisor of “XX” which is of type Researcher, as appeared in below Figure 4.2.

Connecting some of those triples, semantic graphs or arrangements are developed.

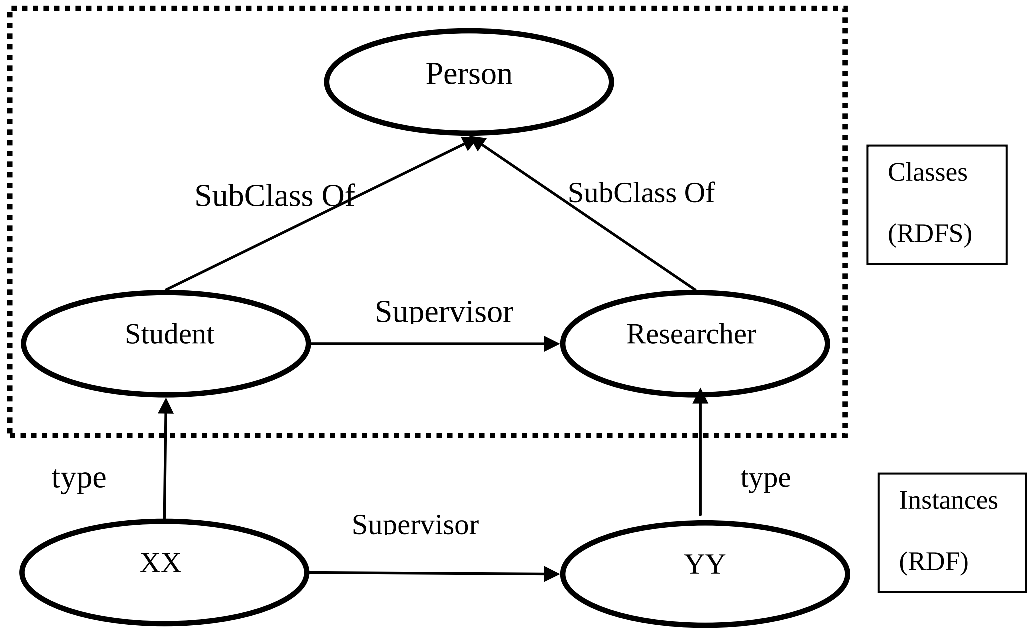

Figure 4.2 RDF Schema (RDFS)

RDF Schema (RDFS) is utilized for representing class hierarchies, including features and various relationships for possible resources. In RDF, object classes, relationship between them are additionally a type of resources, so others can be transferred as portion of that graph. To express RDF various syntactic ways have been already proposed, but conceivably XML-based is most widely used yet. Figure 4.3 depicts the instance of RDF(S) expression.

```
< Class about = “Student”>
        <subClassof resource = “Person”/>
</Class>
<Class about =”Researcher”>
        <subClassof resource = “Person”/>
</Class>
<Class about = “Person” />
```

Figure 4.3 RDF(S) Graph Example

RDF and RDFS along with usage of query language is used for maintaining the database. These languages bolster complex questions on a RDF diagram utilizing a primary decisive grammar. (Ringe, Francis, & Palanawala Altaf, 2012) The SPARQL query that is given underneath will surely return a list of confined classes. Figure 4.4 Returns Supper class of some particular hobby.

```
PREFIX ns: http://www.semanticweb.org/SachinPapneja/ontologies/2016/9/hobby.owl#
select ?superClass ?value
where{ns:" + hobby +" rdfs:subClassOf ?superClass,"
    + "[rdf:type "
    + "owl:Restriction;"
    + "owl:onProperty "
    + "ns:is_hobby_type;"
    + "owl:someValuesFrom ?value] ."
    + "filter(!isBlank(?superClass))}"
```

Figure 4.4 RDF(S) Query Example

The two ontology specification language recommendations were put ahead those are Ontology Inference Language that originated in Northern Hemisphere, and DARPA Agent Markup Language (DAML), known to be developed in the USA; they were given behind RDF and RDFS. Both were then joined and combined into a single one to form a new language that contains a feature of these two languages, as these two languages are more or less very similar: DAML+OIL. Endeavoring benefits of OIL+DAML that enhances, overcome its disadvantages, establishment of very new language known to as Web Ontology Language. OWL signify, expressed for constructing the same as RDF, regularly recognized by expanding following. OWL encompasses entire significant inclinations done by RDF(S), increases it by probability of handling discerning appearances. Further, OWL enables consigning characteristics for the semantic resources, and they are described as cardinality, inverse associations, or transitivity in nature relations.

Some of the few illustrations are described below. RDF(S) moreover OWL furthermore analyzed in various comprehensive ontology representation, a representation among different impressive actions continued to be produced they are as Web ODE or OCML. Figure 4.5 shows OWL expression of Cricket Class.

```
<owl:Class
rdf:about="http://www.semanticweb.org/SachinPapneja/ontologies/2016/9/hobby.owl#Cricket">
    <rdfs:subClassOf
rdf:resource="http://www.semanticweb.org/SachinPapneja/ontologies/2016/9/hobby.owl#Team_S
ports"/>
    <rdfs:subClassOf>
      <owl:Restriction>
        <owl:onProperty
```

```
rdf:resource="http://www.semanticweb.org/SachinPapneja/ontologies/2016/9/hobby.owl#is_hobb
y_type"/>
            <owl:someValuesFrom
rdf:resource="http://www.semanticweb.org/SachinPapneja/ontologies/2016/9/hobby.owl#Medium
_Ball_Related"/>
         </owl:Restriction>
      </rdfs:subClassOf>
   </owl:Class>
```

Figure 4.5 Example of OWL

4.6 Hobby Ontology Development

Protégé 5.0.0 is used to create and visualize the Hobby Ontology. Jena API 2.6.4 is a Java API (Application Programming Interface) framework (Jena 2011) that provides classes and interfaces to construct ontologies using the set of extracted semantic concepts and their corresponding relationships. The constructed ontology is represented in the form of semantic markup language called Web Ontology Language (OWL). In Figure 4.6, hobby ontology is constructed manually for different type of hobbies defined in website www.discoverahobby.com. The root class is Thing from where different nodes in the graph are derived. There are total 6 levels in the ontology graph and having 269 leaf nodes whereas each leaf node represent some particular hobby. There are 2 main concept classes in the Ontology graph one is Hobby and other is Hobby Type.

Figure 4.6 Hobby Ontology

Figure 4.7 illustrate the visualization of Hobby Ontology using OntoGraf Option in the Protégé tool. Hobby Class is derived from main class Thing where as it has 10 different sub classes, each represent different type of hobby category. Dotted line from Hobby to Hobby_Type class represent that Hobby is Hobby Type. Figure 4.8 represent visualization of Hobby_Type Class which is also derived from root class Thing and having 26 subclasses.

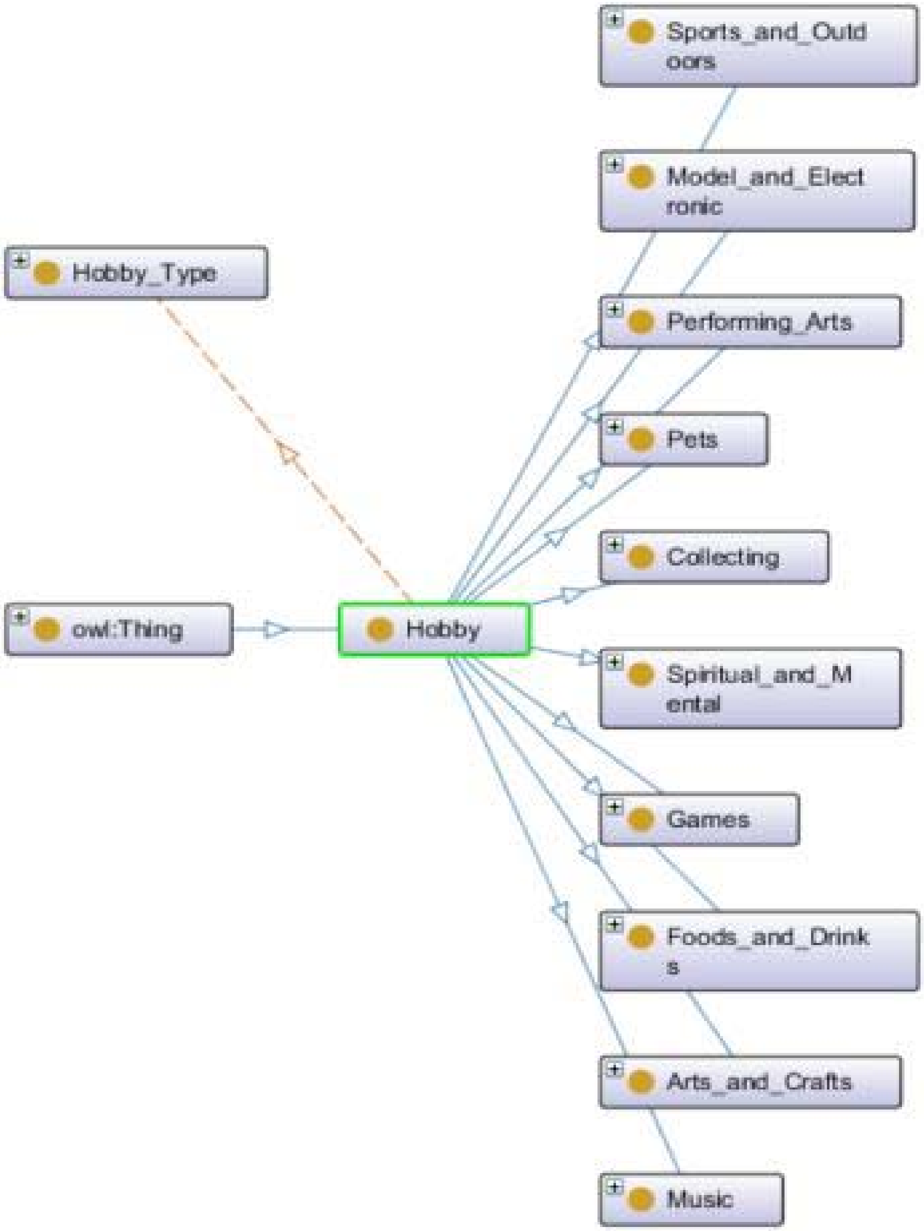

Figure 4.7 Hobby Ontology Visualization using OntoGraf

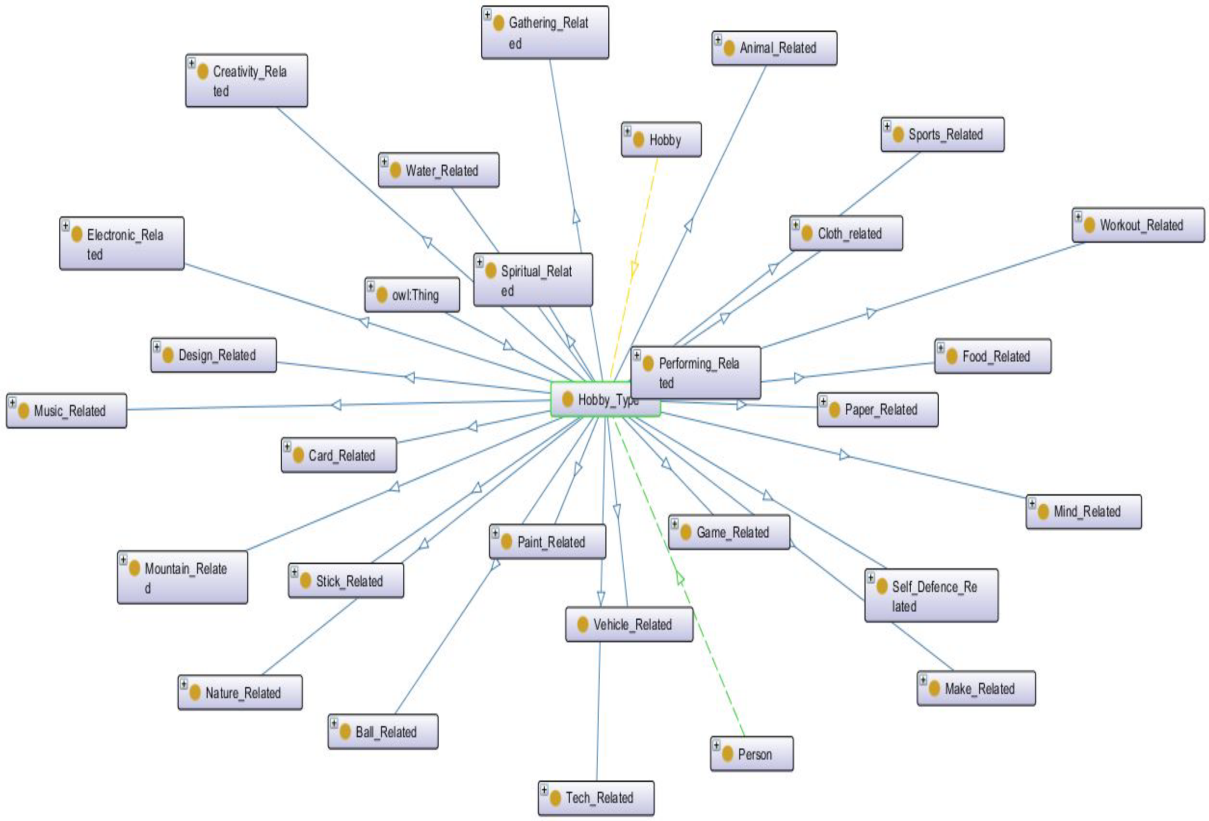

Figure 4.8 Hobby Type Visualization using OntoGraf

Figure 4.9 depicts Team Sports Class which is subclass of Sports_and_Outdoors class and has 14 subclasses related to outdoor sports. All the subclasses are also derived from subclasses of Hobby_Type Class. For example, Football class as shown in Figure 4.10 which is subclass of Team Sport, is also a subclass of Big Ball Related which is further a subclass of Hobby_Type Class. Idea behind is that person who is interested in Football may be interested in other sports related to Big Ball. Other sports related to Big Ball are American Football, Baseball, Basketball, Rugby, Throw ball and Volleyball as shown in Figure 4.11. Similarly, Polo, Tennis, Squash, Hockey, Cricket, Croquet and Lacrosse comes under Medium Ball Category as shown in Figure 4.12

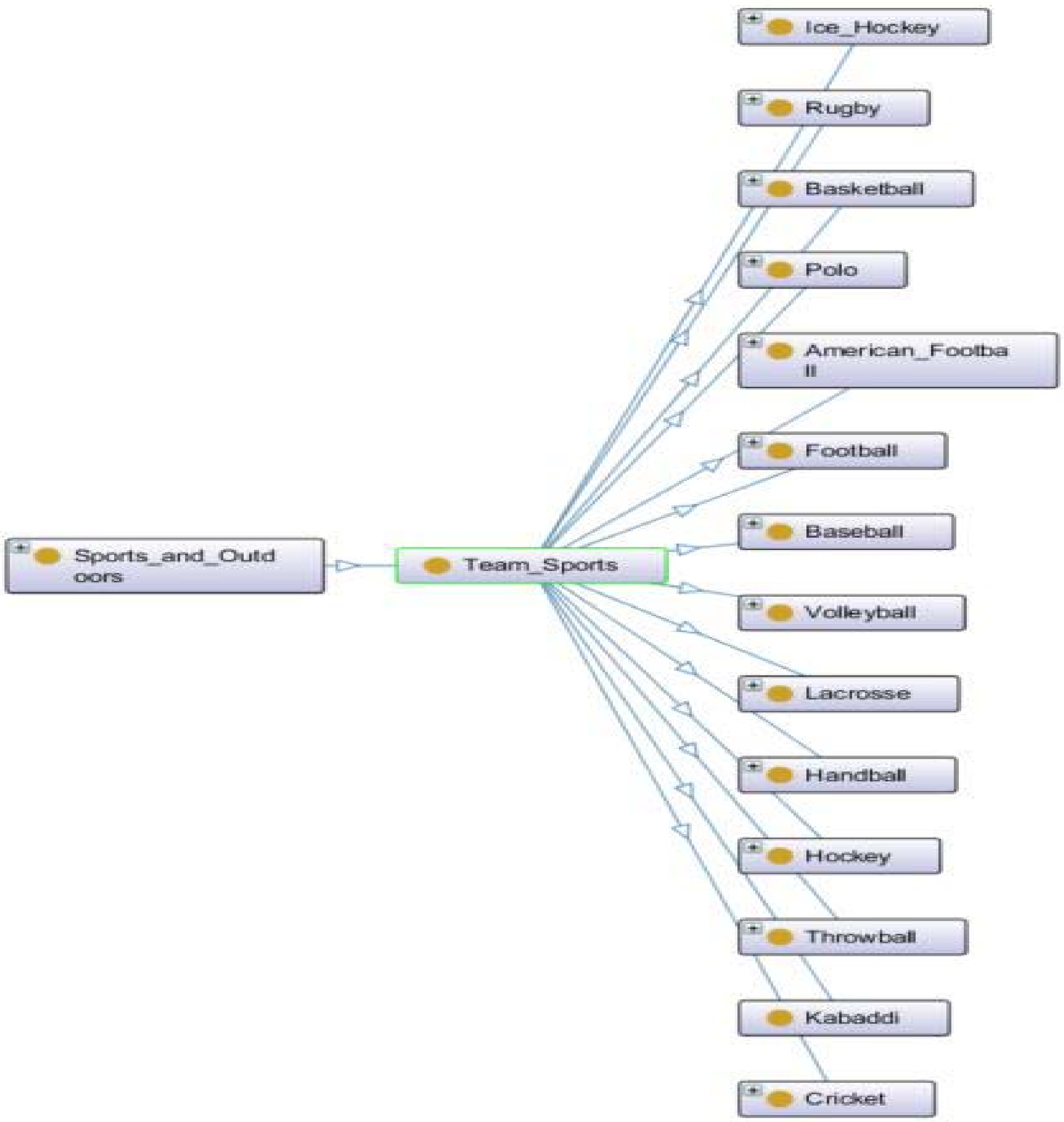

Figure 4.9 Team Sports Visualization using OntoGraf

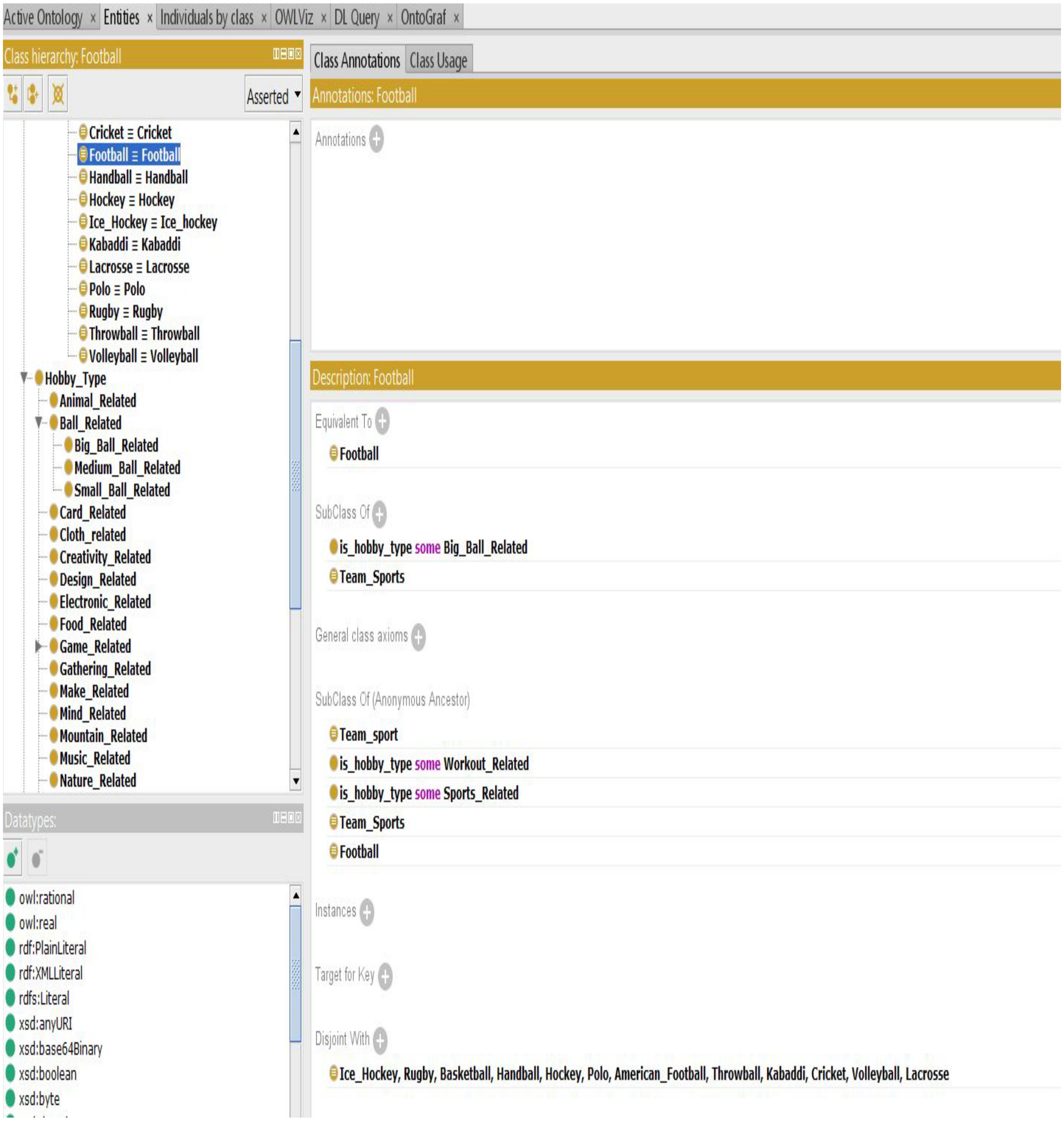

Figure 4.10 Football Class and related Subclasses

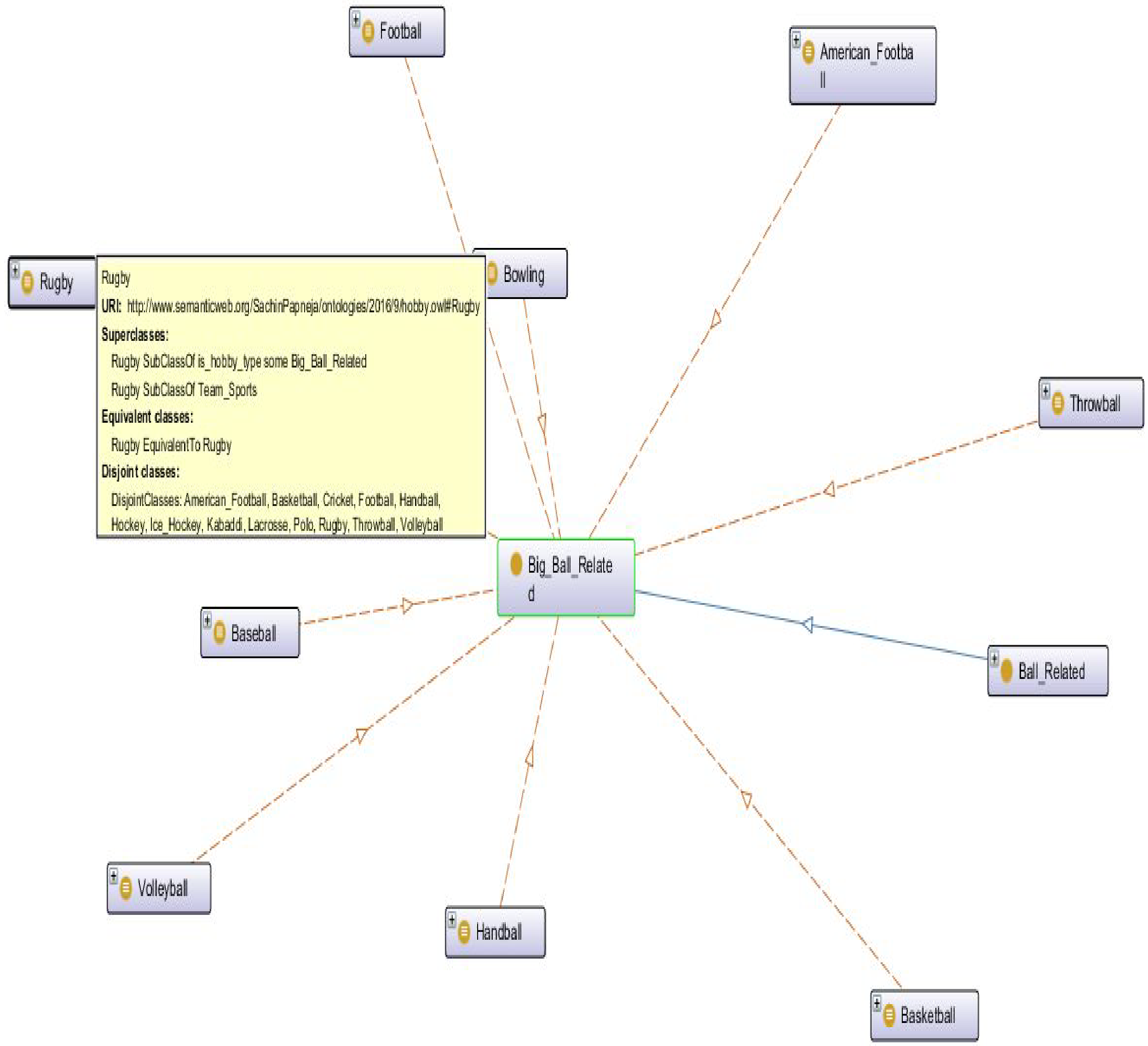

Figure 4.11 Big Ball Ontology Visualization using OntoGraf

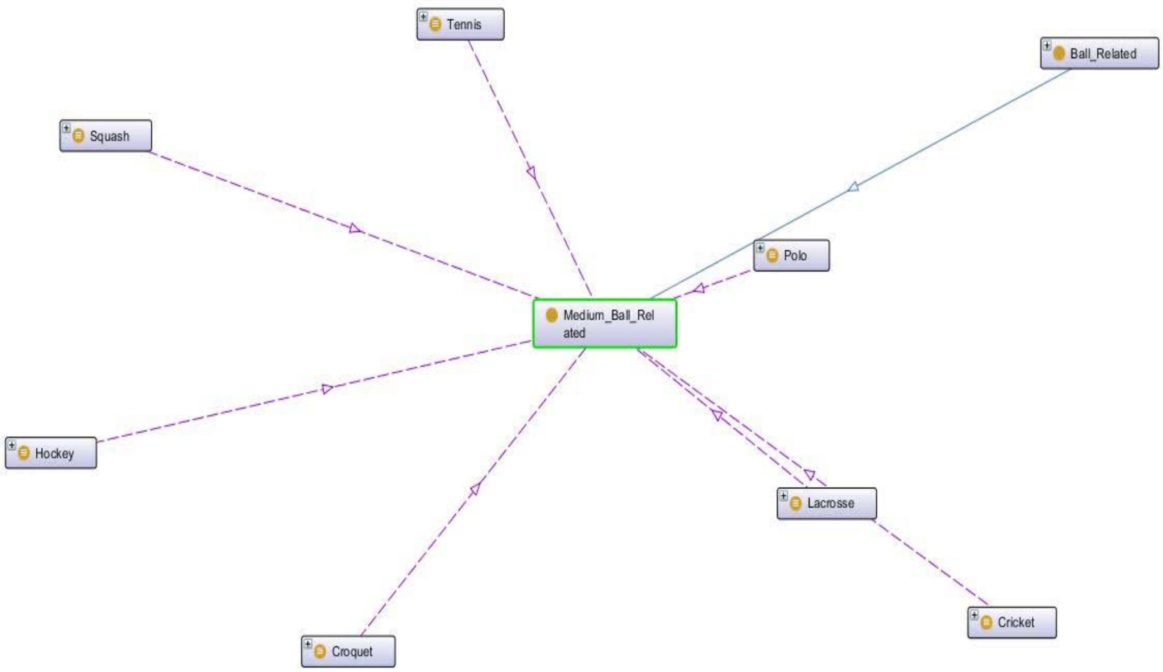

Figure 4.12 Medium Ball Related Sports Ontology Visualization using OntoGraf

Once hobby.owl file is generated using Protégé, the same will be uploaded in to Apache Jena Fuseki Server firstly by executing the following command from the command prompt

Fuseki-server --update --mem /ds

The Jena reserves knowledge as RDF triples in the form of directed graphs, and further permits to manipulate, add, store and publish, add such knowledge. Fuseki is a type of server that is used for publishing data, that can update, and present, RDF representations across some web sources that are practicing SPARQL and HTTP.

After executing the command, hobby.owl file is uploaded to Apache Jena Fuseki Server using the local host, http://loalhost:3030/dataset.html?tab=upload&ds=/ds,as shown in Figure 4.13. Total 2913 triples are generated. Once Ontology is uploaded to Server, Recommendation system can send the SPARQL query over HTTP to get the required Information form the hobby ontology.

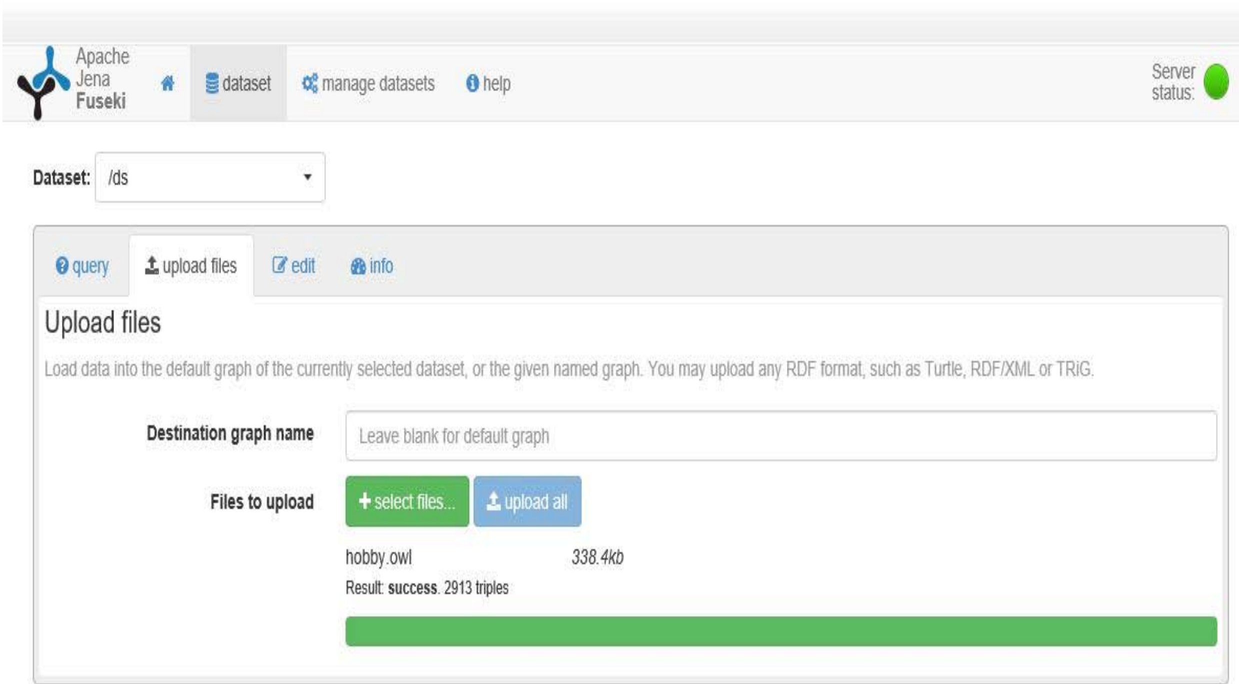

Figure 4.13 Ontology upload to Apache Jena Fuseki Server

Chapter Five: EFFICIENT RECOMMENDATION SYSTEM

This chapter will explain about the proposed recommendation system framework, the underlying Spreading Activation algorithm which is used on top of Ontology graph to recommend content to end users and predict the user behaviour. Based on the predicted user interest, recommendation system will further recommend the friends with similar interest to end user.

5.1 Introduction

In this chapter, the research work has been extended by use of already constructed Hobby Ontology in efficient Content Recommendation by one of the graph based algorithm that is known as Spreading Activation algorithm. In Spreading Algorithm recommendation approach, contents that are recommended for the particular consumer or user are entirely and purely predicted from the past behavior of user. Then the activation score of various nodes are determined using spreading activation algorithm depending upon which the operation performed by the user on the recommended content. Based on the activated leaf nodes where each leaf node represent one hobby of the ontology graph and the threshold value, content related to other related hobby domains are recommended to end user. This proposed system is useful in recommendation where the interest of user keeps on changing as it helps in finding out the user interest domain. Once the system predict the user interest domain, it will start recommending user with the Friend Recommendation.

In the proposed recommendation approach, news articles are suggested for the users based on the user dynamically changing interest. Earlier the user is suggested various news articles and other activities based on his/her information consisted in his/her profile entered by him/her in the initial phase but later on as his/her interests keeps ongoing, the changes interest is being captured through recommendation system, and the corresponding articles are recommended accordingly.

Recommendation system performs the following set of tasks:

Content Recommendation: Learns the user's dynamically changing interest and recommend articles accordingly.

Friend Recommendation: Recommend Friend to user based on the common interest and a results form a social community of user's based on various Interest.

5.2 Content Recommendation Proposed System

User profile is constructed based on the hobby Ontology graph described in previous chapter. Interest that is related with person or consumer or human or user is encapsulated in implicit or explicit manner. Knowledge or Information that is explicit in nature form contains and specifically includes age of the person or consumer or human or user, gender of the person or consumer or human or user and location of the person or consumer or human or user, and particular interest of the person or consumer or human or user. Whereas knowledge or information that is Implicit in natures is captured on various time to time that is specifically based on behavior of person or consumer or human or user observed in near past. Then the interest or choices of person or consumer or human or user is further mapped to the concept that is based on domain Ontology. Initially, majority of contents are recommended to end person or consumer or human or user that is further based on the person or consumer or human or user explicit choices or interest and as per the information contains in person or consumer or human or user profile and least number of random articles. Based on the user action on the recommended articles, whether he likes, dislike, share or bookmark the article, recommendation system will learn the user behavior and subsequently, new articles will be recommended to the end user. Hobby Ontology graph is an input to Spreading Activation that is the strength and power of the various recommendation system will then learn from the person or consumer or human or user behavior based on the person or consumer or human or user feedback and is being explained particularly in the very detail in the following subdivisions sections.

5.2.1 Spreading Activation

The hypothesis of Spreading Activation was first proposed by (Quillian, 1966), in a model of human semantic memory. The way human brain works, Quillian proposed a theoretical model of human memory by which a human's mind may process and comprehend the semantics of Natural language. This model was upgraded by (Collins & Quillian, Retrieval time from semantic memory, 1969) for activities related to information retrieval, and further altered by (Collins & Loftus, A spreading-activation theory of semantic processing, 1975). (Collins & Loftus, A spreading-activation theory of semantic processing, 1975) gave motivation to inquire about in numerous other related fields, from psychological brain science to neuroscience, to NLP, among others (Pace-Sigge, 2013).The fundamental reason of spreading activation is identified with that of relationship in AI, which utilizes comparative models for neural systems to mirror the fan-out impact of electrical sign in the human cerebrum. On account of neural systems, a vertex in the graph relate to an individual neuron, and edges depicts connections between these neurons. In data recovery (Crestani, Application of spreading activation techniques in information retrieval, 1997)and word-sense disambiguation (Tsatsaronis, Vazirgiannis, & Androutsopoul, 2007), by and large vertices will depicts words and edges will depicts some type of relationship, either lexical or semantic linkage, between these words.

Spreading activation implied as a paradigm of knowledge; however, it is not a very new approach in the analysis of semantic networks linked research, as in the area of knowledge retrieval there are many number of method and approaches applications of spreading activation. Crestani suggested to utilize the typical usage of constrained SA on network having hyperlinks to create them without any intervention so as to provide the way and method of performing and doing browsing between these particular networks. The SA theory (Anderson J. R., A spreading activation theory of memory, 1983)has been demonstrated to give a model a high level of logical capability in mental psychology (Sharifian & Samani, 1997). The primary benefit of this model is that it acquire both the manner in which information is acquired and furthermore the manner in which it is handled. As indicated by this hypothesis, information is stored in the nodes and associated connections between nodes, which structure a semantic system of concepts. The intensity of the connection and the separation between the nodes are dictated by the semantic connections between the connected nodes. This model expect that activation spreads from one applied node to those around it, with more prominent activation to the closer ones (Gazzaniga, Ivry, & Mangun, April 1998). This transmission of activation helps in achieving more area of network to be available for better psychological processing. Acceleration and probability of getting to a memory is dictated by its degree of initiation, which thusly is controlled by how much of the time and how as of late it has been utilized (Anderson J. R., Cognitive Psychology and its Implications, 1995), as communicated by the strengths of nodes and the connections. This intensity decreases as the time passes (Anderson J. R., A spreading activation theory of memory, 1983).

Experimental proof has indicated that the value of activation of a node to be activated depends upon the intensity of the connection between the destination node and the source activation node (Lorch, 1982). Also, the measure of activation spreading from a given node along a path is related to the connection of that path comparative with the total of the qualities of all ways originated from that particular node (Reder & Anderson, 1980). This point out to the spreading of the activation originated from one node between all its related nodes, reducing the value of activation other nodes obtain (Anderson J. R., A spreading activation theory of memory, 1983).

(Yang, 2010)is associated through “some sort of activation decay” that probably can be involved in some (arbitrary) phases of adjustment done on pre part or adjustment done on post part ; (Chang, Lin, & Chen, 2016)completely covers a decay element; and further (Agre & Dongre, 2015)covers an activation holding level for the similar objective.

Nevertheless, those terms only figure wherewith significance of things is missed, and they do not really apprehend the assumption of these “task that are current”. Neural systems and especially Hopfield (Daoud, Tamine, Boughanem, & Chebaro, 2007)neural network tried to access and replicate the associative memory repeatedly by applying weight of nodes at distinct position. For this situation, the individual system nodes are not discrete concepts oneself, yet rather, in their entire, are utilized to depict to memory states. This methodology relates to the neuron elements of the human mind and mostly centers around the capacity of recollections, though our own endeavors to re-enact the human memory applied system works and concentrate on the illustration of initiation of individual concepts.

Recently, theory that is based on spreading activation has been identified as a proposal for establishing individual intercommunication among the various system, in the recently developing fields of personal information management (PIM) and the other is Task Information Management (TIM).

5.2.2 Spreading Activation Model

Spreading Activation is composed up of a simple processing method that is dependent upon on a system construction over the data. The system information construction completely consists of nodes that are further attached by various links, as portrayed inside accompanying Figure 5.1.

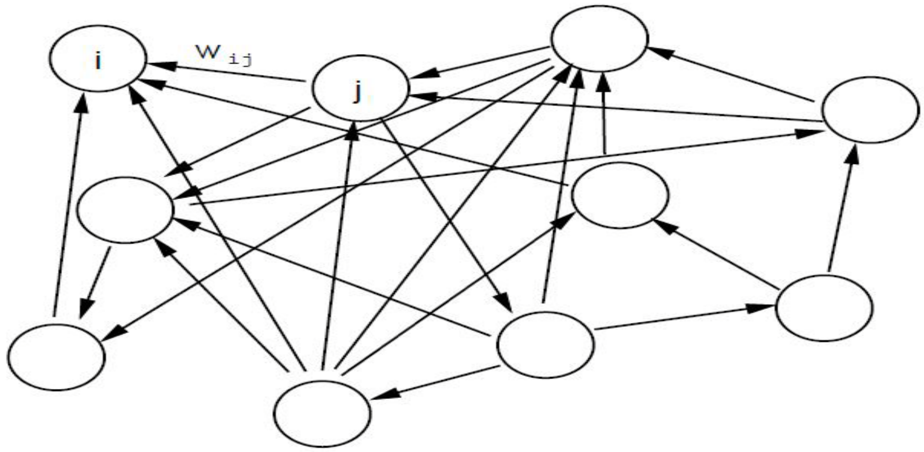

Figure 5.1 Spreading Activation Model Structure

Nodes of the graph depicts the “physical world” entities. They are generally classified with the name of the entities that are designate to express. Relationships among various nodes are links model; moreover they can hold weighted and/or labeled. The relationship model moreover exhibits the associations among various specific features and/or objectives of the “present physical world” that are further designated. Each connection is a type of an edge or a line joining two points or nodes that normally has a fixed direction, a description, and a weight that is allotted to it in accordance to a particular direction.

The method that is technically characterized by a series of repetitions similar to that example schematically illustrated in Figure 5.2. Where each repetition is come up by different repetition continuously is stopped by that user or through the triggering of any finish state. Repetition consists of particularly two things that are further followed as:

1. Pulses that are one or more

2. Checking of termination condition.

The simple SA prototype from another new complicated prototype is distinguished by the series of operations that composed of the pulse. Basically, a pulse consists of the three phases that are as follows below:

1. Pre adjustment of pulse
2. Spreading of pulse
3. Post adjustment of pulse

Pre-adjustment of pulse and post settlement of pulse states are not compulsory, it is not necessary to perform those adjustments, but an unusual type of activation reduction can continue implemented to the connections that remain completely active in nature. The purpose of these pulses is to prevent holding of activation from preceding pulses, which further enables us to take control of both the overall activation of the network and activation of single nodes. They perform a pattern of “lack of interest” within connections that exist and remains not activated continuously.

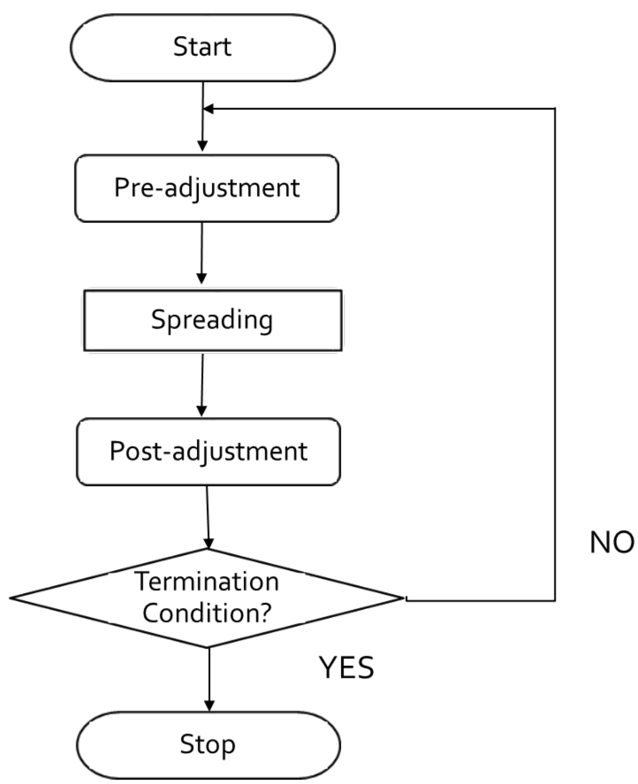

Figure 5.2 Spreading Activation Model

From one node to other additional nodes associated to it, the growing phase consists of several iterations of activation patterns. The activation over a network could be expanded in plenty of methods (for instance, look at (Jussupova-Mariethoz & Probst, 2007)). In the more simplistic fashion, toward a single entity layer, SA consists beginning within each individual calculation about each length information by using this formula:

$$\boldsymbol{Ij = \sum_i Oj\, wij} \qquad (5.1)$$

Wherever:

Ij is that cumulative input about connection j;

Oi holds a result of entity i combined to node j;

Wij holds a weight related to that connection combining connection i through connection j.

Specific data to this is usually the weight that is real numbers, though their statistical representation is defined by the particular conditions of the importance to be modeled. Inappropriate, both can contain the binary states (0 or 1), excitatory/inhibitory states (+1 or -1), or all can obtain notably conditions that further indicate the power of the relationship among various nodes. Output value must be determined by following a link that has measured its input value, its. The requirements of the application also determine the numerical type of the output of a node. The two common survived cases comprising the paired active or non-active type (Zero or One), including this extremely preference standard. In SA algorithm, there exists no differentiation among “activation” or “output”. This output value is the activation level of a system. This is being calculated as a function f of Input value as follows:

$$O_j = f\,(I_j) \qquad (5.2)$$

Function f can be linear, step or sigmoid function based on the requirement or implementation as shown in Figure 5.3.The threshold function is a majorly function that is used in pure SA models. In case of output in form of binary value, above function can be modified:

$$\boldsymbol{Oj} = \begin{cases} 0\ Ij < kj \\ \mathbf{1}\ \boldsymbol{Ij} > kj \end{cases} \qquad (5.3)$$

Wherever k_j means that threshold amount toward link j. The threshold amount about each mathematical activation function is utilization conditioned that may change from connection to one connection. Consequently, the representation k_j for single threshold value is being practiced.

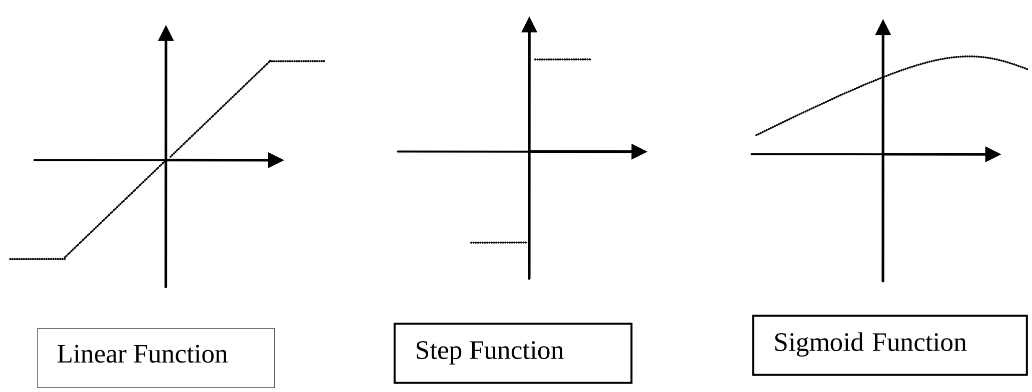

Figure 5.3 Commonly used Spreading Functions

Following each node has estimated its calculated value that propagate this over every nodes that are correlated to it, normally giving the equivalent amount to several of them. After going through multiple pulses, that activation settles across that system moving links that remain far away from the activated ones initially. Post certain number of pulses got triggered, a terminating circumstance is being verified. If the situation remains satisfied, then the individual Spreading Activation process stops, unless it runs on toward different sequence concerning pulses. Spreading Activation does, consequently, iterative in nature, and consists of a sequence of pulses and a condition that is a termination condition for stopping of this method.

5.2.3 Spreading Activation over Hobby Ontologies Graph

The profile or account of the user is built by collecting the interests of the person or consumer or user through their interactions with the provided interface, which recommends content in the form of news articles. The user actions such as clicking, liking and sharing of any news article tells us of their interest in the tags associated with that particular news article, which is already processed and is already present in our database, and thus the interest the user might have in a particular tag or topic based on their actions can be calculated. This information that keep collecting between intervals of time is used to run Spreading Activation cycles and populate the various nodes and thus form a general and overall understanding of the person in question and thus construct their user profile. So, the user profile is nothing but the Spreading Activation values of the different nodes at that instance of time.

All nodes have an initial activation value of 0 before any spreading activation cycle begins. On action executed by a person or human or consumer or user on a specific collection of the various articles, the method performs Spreading Activation utilizing every node in the particular Activation Set since beginning Activation connections, including measures Activation charges concerning every connections. Later on, the node values start getting populated. With each new activation cycle, changes to the previous value of a node are made, which in turn was determined by the previous activation cycles. After the current cycle has stabilized, the activation values that the nodes possess at that moment in time represent their pertinence in that context. Activation values are nothing but a number which denotes how relevant one topic is to another connected topic. Super notion, sub-notion, and sibling notion connections are stimulated by employing Spreading Activation, nearby notion connections that is., and this activated concepts, when they further provide the particularized guidelines, shall give us information regarding that consumer or user grew enthusiastic, also that longing also assists in making of credentials and holding this consumer or user existing choices and specific interest in brain. A depreciation factor is calculated that further diminishes activation conditions of links as interval improves, and that proceeds sure past human or consumer or people or user choices and interests and then show user choices or benefits both are carried into thought. Each and every primary activation series seems not to originate of a clean slate exactly, including the amounts measured during that initial activation cycles that happen further decreased based upon the decay circumstance in the implementation of Spreading Activation. The technique and method that is designed would particularly work on the following manner, where let’s say 3 nodes (Sports, Team Sports, Hockey) are connected, where Team Sports is the parent node of Hockey and Sports is the parent node of Team Sports. Now assuming that Hockey is the initially activated node from where the activation cycle begins, the Hockey node is assigned an activation value of say, 0.8, and this value propagates to the nearby attached, adjacent nodes, in this case that would be Team Sports. The activation cycle passes on to it, with a reduced value and in order to simulate that reduction the activation value is divided by 2, i.e., 0.4 received by Team Sports. Now this activation cycle propagates further on to Sports from Team Sports, its value further decreasing by a factor of 2. And this process continues until the Spreading Activation cycle dies out naturally due to the threshold value that is set initially.

We perceive our most advanced materials that provide through each user or consumer upon those issues where the activation rate remains higher than that of an exceptional inception amount. While all of us receive that stimulating benefits behind a growing activation pass across the philosophy. To mimic the disappearance of the excitement regarding that human or user that is under some problems are undecidedly obtained by Spreading Activation, We also tried to reduce these states' activation regarding those various connections. It is being observed that the advantages particularized through each person or human or user that then diminish about some faraway more secondary degree as opposed to those consequences discovered completely within the consideration. This sense for this observation comprises a time-based inside each activation state 'A' from each particular subject(game philosophy department) declines by some specific value 'D' following some designated number concerning experience 'T' produces moved. That goes off 'D' signifies utterly changed during particular enthusiasm and the absolute user choices.

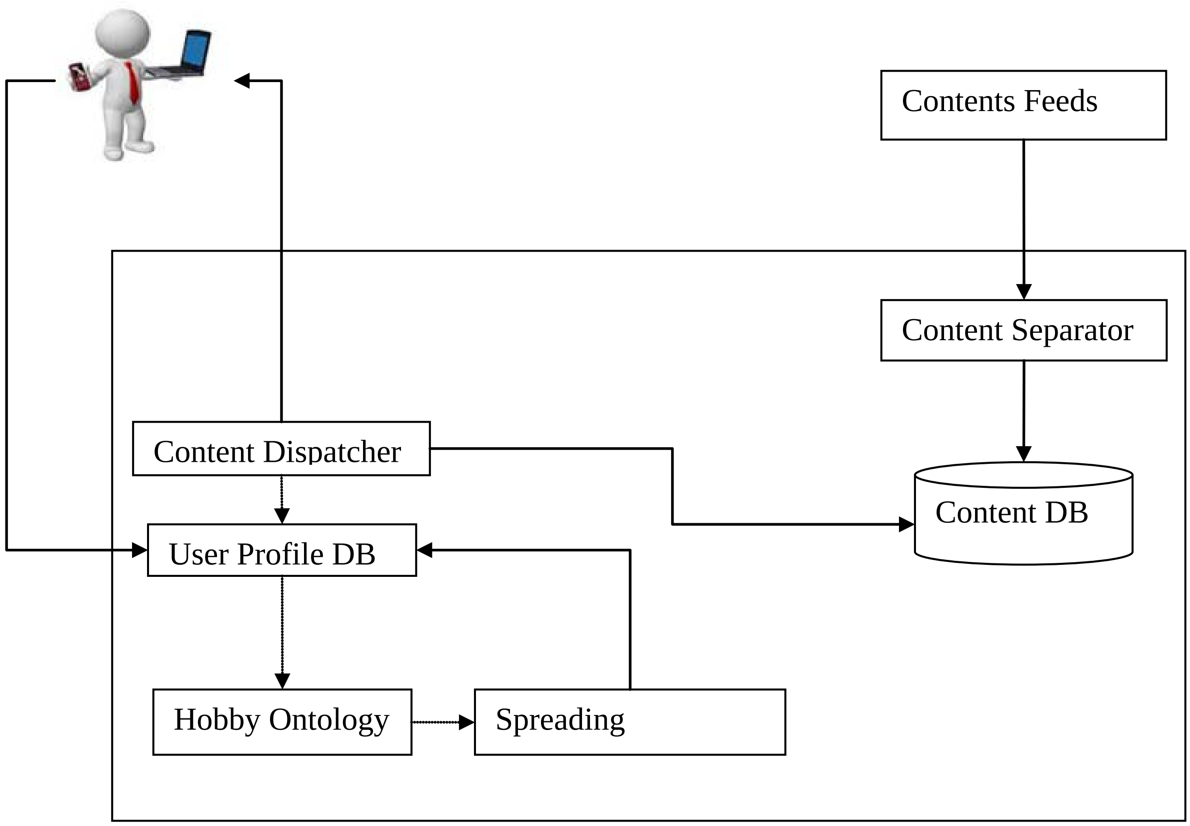

Figure 5.4 Content Recommendation Framework

5.3 Friend Recommendation Proposed System

Based on the spreading activation cycle which runs on the ontology, certain nodes get activated and thus get values assigned to them. As time progresses, the node values keep changing due to activation and decay of their values. Now if two people are to be recommended to each other then it is not only the % similarity of the activation values of their interest nodes such as cricket or football etc. but also the variance of these nodes which matters. The variance of these nodes is nothing but the value which represents how frequently the activation value of that particular node changes. Like for example: If the activation value of the node cricket varies in time as follows - 10,3,14,1,19,4 then the variance of the above dataset would be larger as compared to the variance of the following dataset - 5,6,5,4,5,5. A low variance indicates stability in the data whereas a large variance indicates instability in the data. The dataset for the users is generated randomly. Firstly the variance for a node is generated randomly and a normal distribution is constructed with this variance. Now the variation of the node values over time is drawn from this normal distribution. This is done for all nodes of a user, for all users. All experiments are to be performed on this dataset generated, in which the importance or role of variance in friend recommendation shall be gleaned.

The proposed architecture (Figure 5.4) consists of different steps of which the first step is the building of the person or consumer or human or user account and the collection his/her interaction knowledge. The person or consumer or human or user profile is constructed with gathering choices or interests of the person or consumer or human or user through their interactions with the provided interface which recommends content in the form of news articles. The user actions such as clicking, liking and sharing of any news article tells us of their interest in the tags associated with that particular news article, which is already processed and is already present in our database, and thus the interest the user might have in a particular tag or topic based on their actions can be calculated. This information that keep collecting between intervals of time is used to run

Spreading Activation cycles and populate the various nodes and thus form a general and overall understanding of the person in question and thus construct their user profile. So, the user profile is nothing but the Spreading Activation values of the different nodes at that instance of time.

Now in the next step the variance of each active non-zero node for the user is being determined. The values of each node that has been assigned over time and the changes that have been reflected in them is being tracked. This allows to calculate the variance of the node which shall be used further while recommending friends to the user. In the final steps the similarity between users based on how close their individual node activation values is calculated, which tells how similar their current interests are, are and also based on the respective variance of the nodes, which signifies how consistent their interests are. This similarity calculation can be done using any similarity measures like jaccard similarity, Euclidean distance or cosine similarity. In this particular paper likeness is to be calculated by the degree of likeness of two person or consumer or human or users.

Thus person or consumer or human or user shall be suggested to each other based on the maximum similarity scores i.e. the top n people that possess a very large likeness score from the person or consumer or human or user may be shown to him/her and additional constraints can be applied to the suggestions to further refine and filter them so that they are more relevant for the person or consumer or human or user, Let’s say for particular illustration if the location supposed to be a deciding criteria for friends being suggested then this can be used to filter out the recommendations who live in close proximity of the person or consumer or human or user, leading to better quality of suggestions. Also, other criteria like the genders and age can be considered. This will increase the chances of the user accepting the friend suggestions and a connection between them being made.

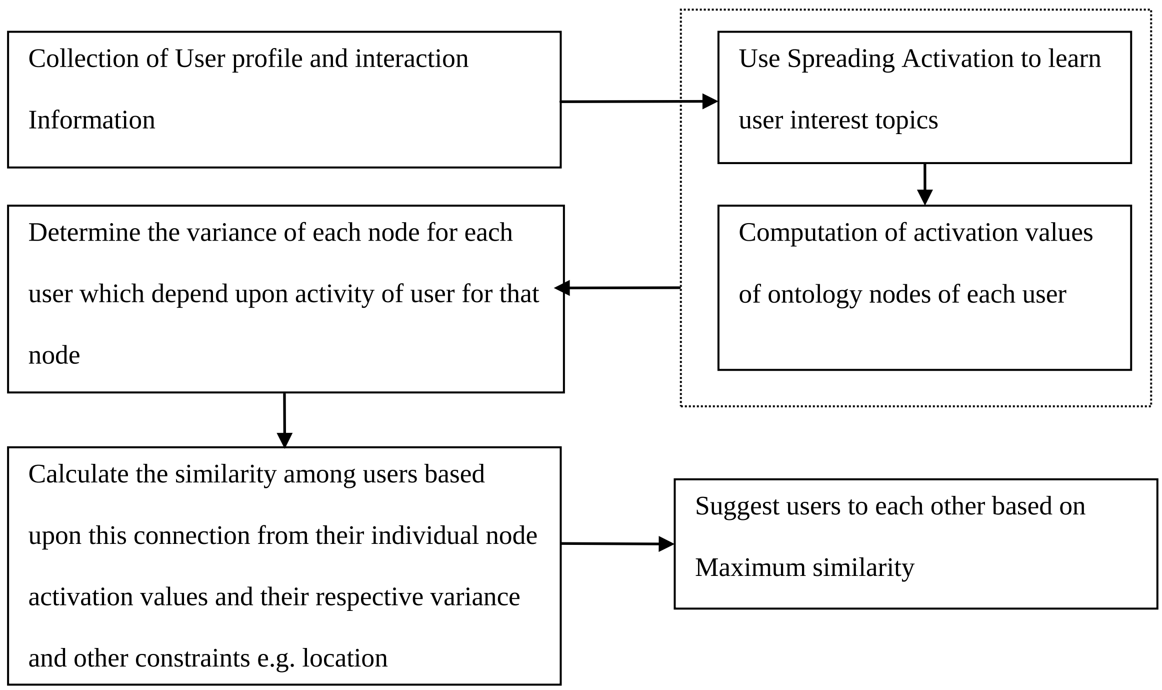

Figure 5.5 Friend Recommendation Architecture

Algorithm1: Distance Calculation without Variance

Input:

Num: Total Number of nodes in Graph

T: Number of Iterations

G [T][Num]: Spreading Activation Graph

Ni: Node in Spreading Activation Graph

Map [Num] [Num]: Node distance matrix

Data:

I: index

Output:

dist: distance between current and previous top 5 nodes

```
begin
    foreach I in T do
    sort(NI1,NINum)
    return dist (map[NI[1:5]],map[NI-1[1:5]])
end
```

Algorithm2: Distance Calculation with Variance Square

Input:

Num: Total Number of nodes in Graph

T: Number of Iterations

G [T][Num]: Spreading Activation Graph

Map [Num] [Num]: Node distance matrix

Ni: Node in Spreading Activation Graph

Data:

I: index

α, β: Tuning parameter

Output:

dist: distance between current and previous top 5 nodes

begin

foreach I in T do

foreach Ni in G do

$\overline{N_I} = \text{mean}\ (N_1, N_{i-1})$

$\text{Var} = (\sum ((Ni - \overline{N_I})\ (Ni - \overline{N_I})))/I$

$N_i = N_i * \alpha + \beta/\text{Var}$

end

end

sort (N_1, N_i)

return dist (map $[N_I\ [1:5]]$, map $[N_{I-1}[1:5]]$)

end

Chapter Six: EXPERIMENTAL EVALUATION

In this Chapter, Implementation part of the Recommendation system is explained, which Software language and other tools are used in the development phase along with the results in the form of tables and graphs considering different user inputs.

6.1 EXPERIMENTAL SETUP

In this chapter data set along with the results for both Content Recommendation and Friend Recommendation is explained in details. Simulation is done in JAVA language for Content Recommendation whereas Python language is used for Friend Recommendation. In both cases, random data set is being generated. The similarity that is based on Cosine Similarity is practiced to examine the profile of the user to provide Friend Recommendation.

In this chapter, firstly Content Recommendation is explained followed by Friend Recommendation.

6.1.1 CONTENT RECOMMENDATION

6.1.1.1 LANGUAGE USED FOR IMPLEMENTATION

Java is a very easy, portable, object-oriented type, categorized, reliable, understood, which have robustness, having neutral architecture, allows entirely multi-threaded, and is a very dynamic programming language. There are various and many advantages of Java programming language over other different many languages and particular environments that further make it more comfortable and suitable for various many programming tasks. Java becomes a language of choice to implement the concepts for providing worldwide internet solutions. IDE used is IntelliJ IDEA.

6.1.1.2 User Input

All the User inputs are entered through IDE console. User have to enter the following information.

Hobby names in which he/she is interested

Probability of reading Implicit and Explicit Articles

Number of iteration of Spreading Activation algorithm

Percentage of recommending random articles

6.1.1.3 Recommendation System Evaluation

Recommendation System is evaluated by simulating User behavior and considering different set of input values. The graph is plotted based on the output value to figure out the user behavior. Experiment results shows that the slope of the graph is according to the user behavior. Slope is high if the user behavior is of reading the random articles instead of reading the explicit articles. Simulation is performed considering behavior of 1 user on the basis of multiple Spreading Activation algorithm iterations. Table 6.1 shows the sum of the Spreading Activation nodes over a set of 30 iterations with different set of probability of reading Implicit and Explicit articles.

Case 1: As mentioned in Table 6.1, User has hobby in Animation and Cricket. System Recommend 50 Articles out of which 20% of the total recommended articles are random and total 30 iteration of Spreading Activation runs.

Input Type	**Input Value**
Explicit Interest Hobby	Animation, Cricket
Total Number of Recommended Articles	50
Percentage of Random Articles Recommended	20%
Total Number of Spreading Activation Algorithm Iterations	30

Table 6.1 User Hobby as Animation and Cricket

Apart from the above input values, each column in the Table 6.1 represent the values related to Probability of Reading Explicit Articles vs Probability of Reading Implicit Articles. For e.g. In Column 1, value 0.1: 0.9, 0.1 represent Probability of Reading Explicit Articles and 0.9 means Probability of reading implicit articles. Similarly, in 2^{nd} Column, value 0.2:0.8 depicts 0.2 Probability of Reading Explicit Articles and 0.8 means Probability of reading implicit articles and so on.

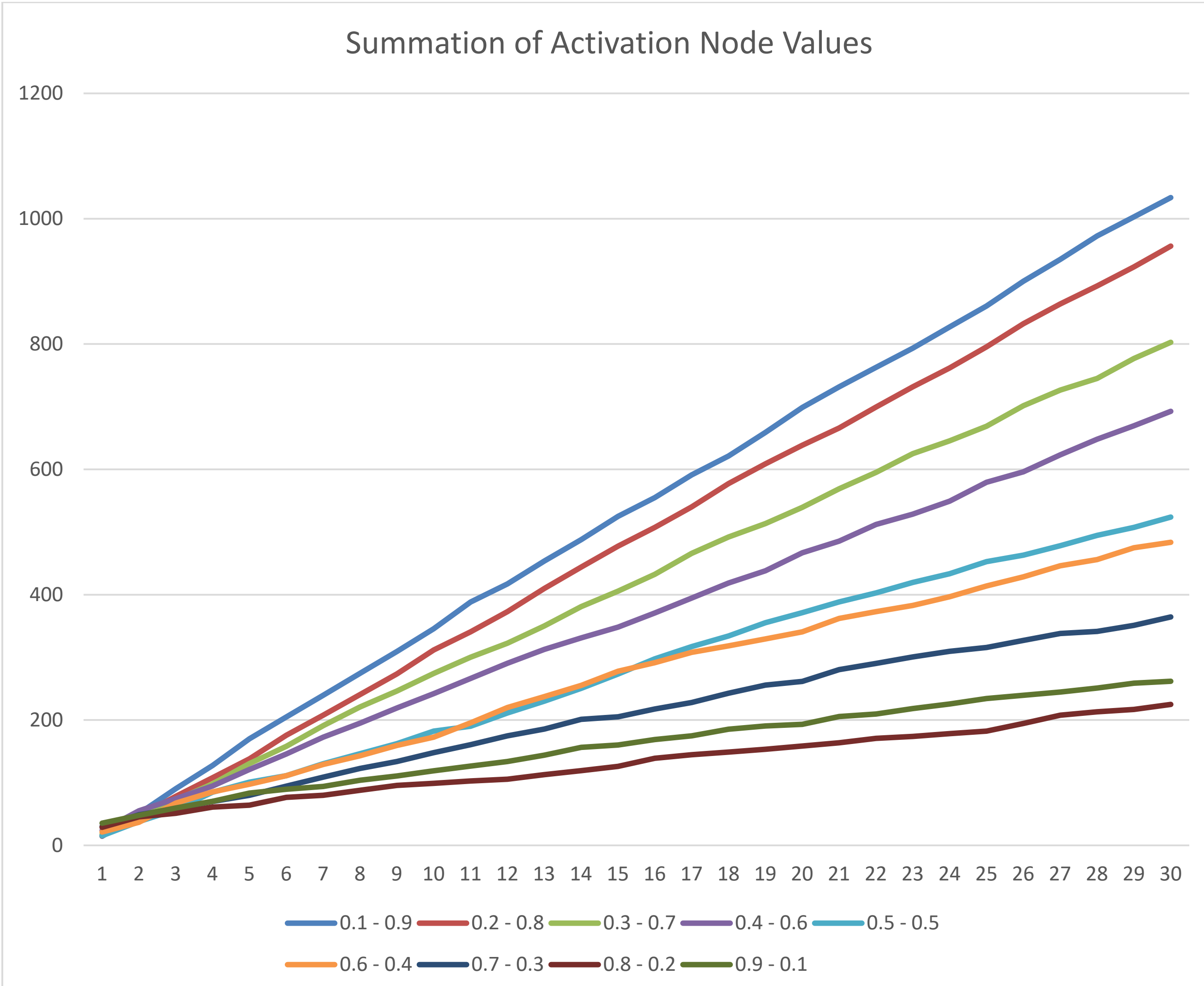

Figure 6.1 Summation of Activation Nodes vs 30 Iteration

In figure 6.1, Y axis represent Sum of Activation node values in the Ontology graph whereas X axis represent the iteration number. Slope of value 0.1 vs 0.9 is the highest one, reason being is that probability of reading implicit articles is more as compared to reading explicit articles. That means that more number of spreading activation nodes in the ontology graph will be activated, resulting in the increase in the total sum of Activation node values.

Table 6.2 Summation of Activation Nodes vs Iterations

Iteration #	Probability of Reading Explicit: Implicit Articles								
	0.1 - 0.9	0.2 - 0.8	0.3 - 0.7	0.4 - 0.6	0.5 - 0.5	0.6 - 0.4	0.7 - 0.3	0.8 - 0.2	0.9 - 0.1
1	14.34375	15.03125	17.15625	23.53125	15.5	21.25	29.90625	28.875	35.40625
2	50.6625	46.18125	43.16875	55.20625	38.075	36.88125	48.80625	45.44375	47.93125
3	90.55	78.34375	72.66875	74.375	56.91875	67.30625	58.7375	51.19375	59.58125
4	127.3875	107.95	97.9875	94.76875	84.8375	85.36875	69.925	61.1375	70.1875
5	169.975	137.9938	130.4625	121.0688	100.7313	97.44375	79.9625	64.13125	83.33125
6	204.8438	175.7563	158	145.8875	111.1563	111.1125	94.375	76.51875	89.45625
7	239.5875	207.6438	191.0375	172.5313	130.35	129.125	108.8187	79.99375	94.1
8	274.55	240.625	220.9813	195.05	146.3563	142.9188	122.7312	87.90625	103.8688
9	309.45	273.6063	246.3	219.15	162.175	159.45	133.8938	95.6	110.6063
10	345.6625	311.6813	274.4312	242.1875	182.3375	172.575	147.9937	98.91875	118.9563
11	388.2812	340.725	300.1875	266.2562	190	195.5438	160.6563	102.675	126.4
12	417.3063	373.05	322.6938	290.4938	211.1625	219.95	174.6938	105.65	133.7688
13	453.7375	410	350.2313	312.6688	230.0438	237.4813	185.6063	112.75	143.8875
14	487.825	444.1063	380.925	331.0187	250.4875	255.2313	201.0938	119.1625	156.3813
15	524.9437	477.4	405.775	348.3062	273.275	277.7625	205.0813	126.1375	160.175
16	555.0313	507.4125	432.5312	370.75	297.5937	291.4813	217.65	138.9875	168.8313
17	591.2438	540.0188	465.85	394.5063	317.2563	307.9188	227.875	144.5563	174.8625
18	621.3313	577.1875	492.1375	418.575	334.1063	318.2938	242.6938	148.8125	185.2375
19	659.075	608.7937	513.5813	438.0812	355.3938	329.4188	255.825	153.3188	190.5188
20	698.7875	638.4	539.3687	466.8375	371.3375	340.6375	261.7063	158.4188	193.1313
21	731.6563	666.1	568.9688	485.5	388.5937	362.1062	280.4313	163.7688	205.6813
22	762.775	699.4563	595.2562	512.0375	402.85	373.075	290.375	170.65	209.6563
23	793.3625	731.7188	625.1375	528.575	419.6375	382.825	300.6	173.8438	218.2875
24	827.2937	761.7625	645.5188	549.3313	433.3625	396.6375	309.6375	178.35	225.6375
25	860.6312	795.525	668.9	579.3687	452.7437	413.9812	315.7687	182.3563	234.3313
26	900.2188	832.35	701.5938	596.0312	463	428.4187	327.15	194.425	239.4625
27	934.9938	864.1125	726.475	623.225	478.1938	446.2813	338.2188	207.7125	244.5938
28	972.5188	892.6562	745.0437	648.0437	494.7313	456.1438	341.4125	213.0125	251.1
29	1002.825	922.9812	776.8625	669.5188	507.3	474.9437	351.1375	216.9875	258.7312
30	1033.631	956.3688	802.7125	692.5562	523.9	483.5562	364.5187	224.8375	261.7375

Case 2:

As mentioned in Table 6.3, User has hobby in Cooking and Reading. System Recommend 50 Articles out of which 20% of the total recommended articles are random and total 30 iteration of Spreading Activation runs. Slope of the graph is represented in Figure 6.2

Input Type	**Input Value**
Explicit Interest Hobby	Cooking, Reading
Total Number of Recommended Articles	50
Percentage of Random Articles Recommended	20%
Total Number of Spreading Activation Algorithm Iterations	1000

Table 6.3 User Hobby as Cooking and Reading

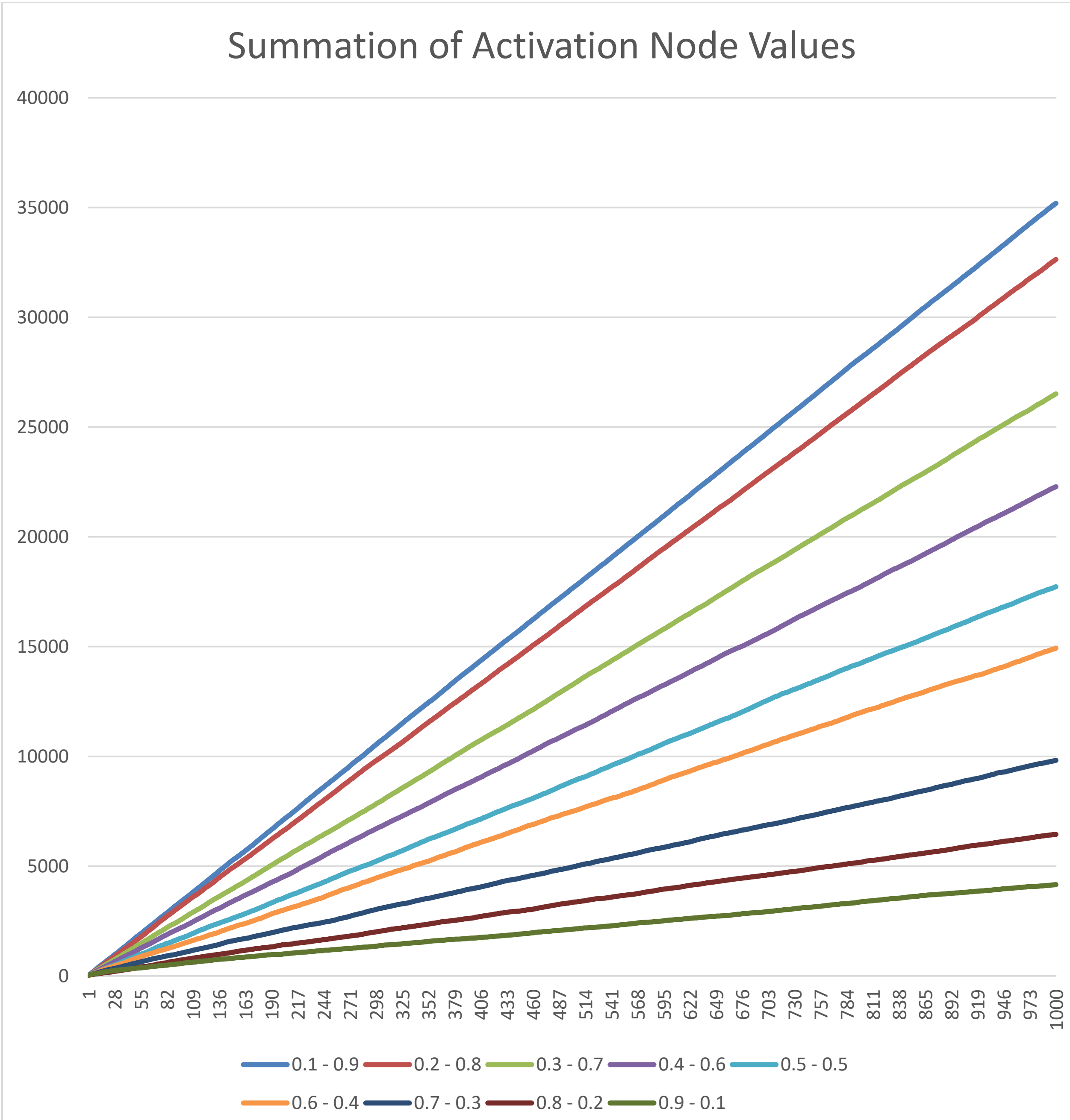

Figure 6.2 Summation of Activation Nodes vs. 1000 Iterations

Case 3:

As mentioned in Table 6.4, User has hobby in Cooking, Painting and Cricket. System Recommend 50 Articles out of which 20% of the total recommended articles are random and total 1000 iteration of Spreading Activation runs. Slope of the graph is shown in Figure 6.3

Input Type	Input Value
Explicit Interest Hobby	Cooking, Painting, Cricket
Total Number of Recommended Articles	50
Percentage of Random Articles Recommended	20%
Total Number of Spreading Activation Algorithm Iterations	1000

Table 6.4 User Hobby as Cooking, Painting and Cricket

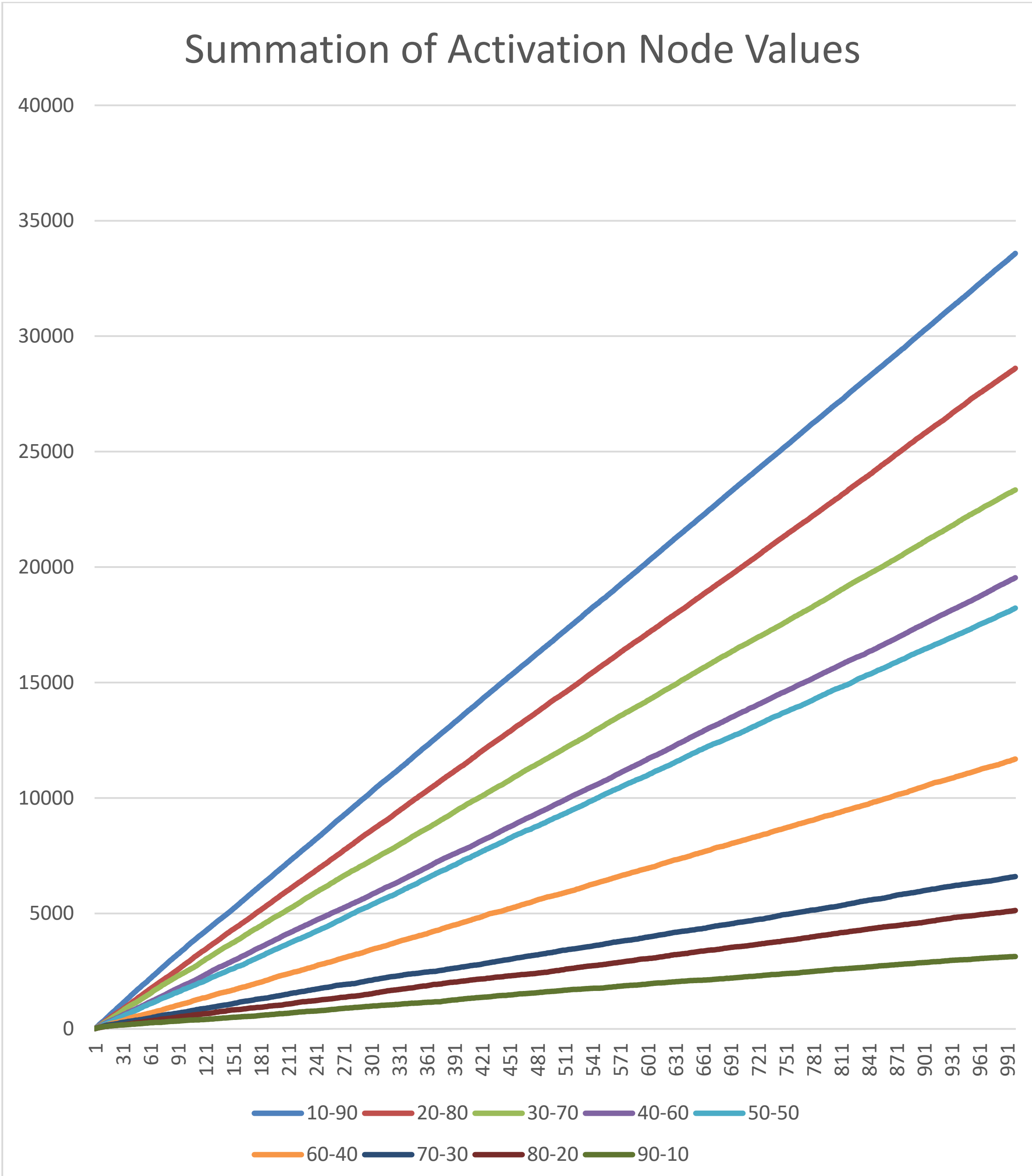

Figure 6.3 Summation of Activation Nodes vs. 1000 Iteration

Case 4:

As mentioned in Table 6.5, User has hobby in Boxing and Football. System Recommend 20 Articles out of which 20% of the total recommended articles are random and total 30 iteration of Spreading Activation runs. Slope of the graph is shown in Figure 6.4

Input Type	Input Value
Explicit Interest Hobby	Boxing, Football
Total Number of Recommended Articles	20
Percentage of Random Articles Recommended	20%
Total Number of Spreading Activation Algorithm Iterations	30

Table 6.5 User Hobby as Boxing and Football

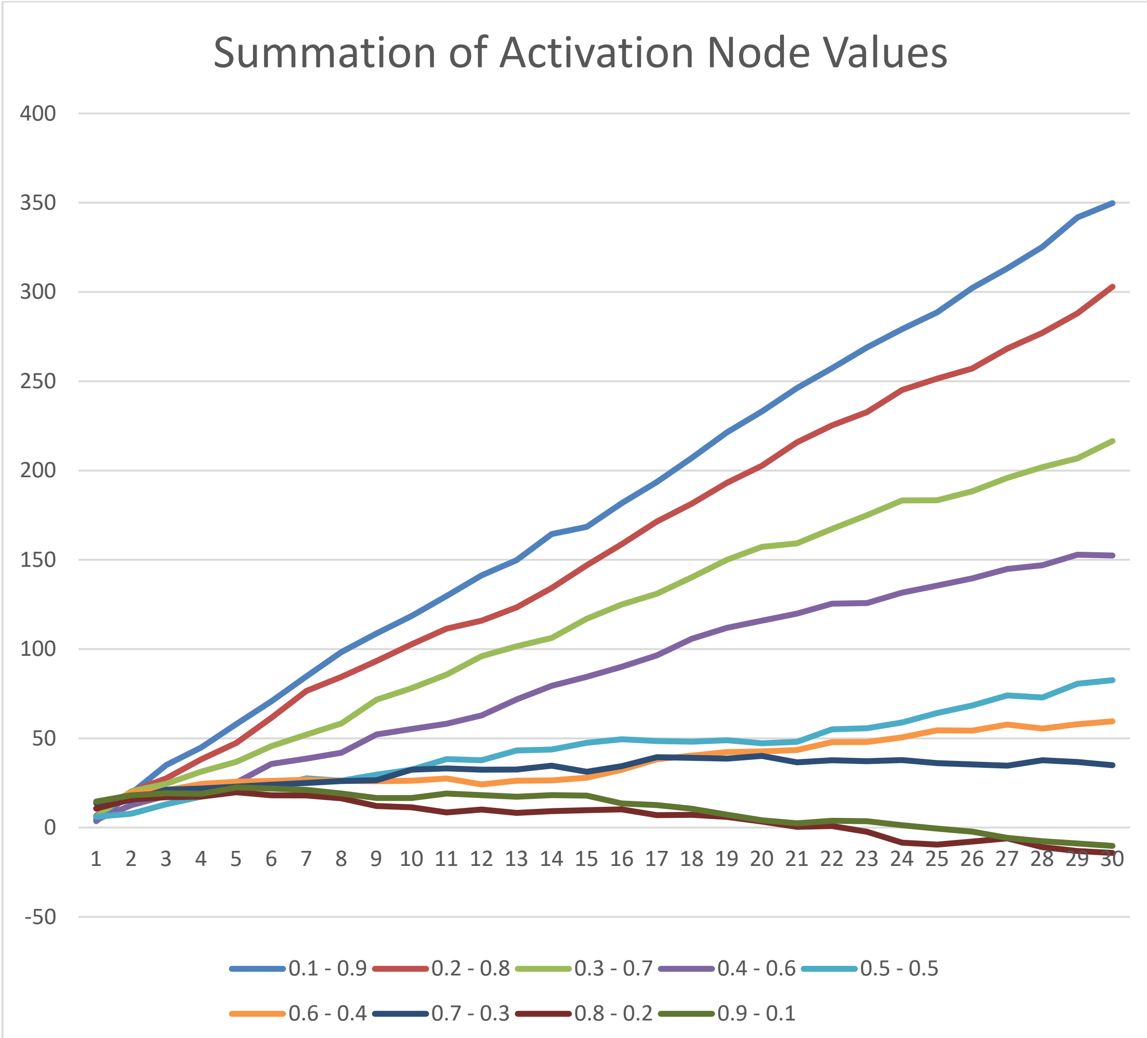

Figure 6.4 Summation of Activation Nodes vs. 30 Iterations

6.1.2 FRIEND RECOMMENDATION

In the above Content Recommendation section, all the results are shown based on experiment performed on 1 user with having different set of inputs. Over a period of time, nodes values gets changed based on the user interest. On the same front, we can have the Spreading Activation values for different users over the period of time. Now 2 people cannot be recommended as friend just by comparing the top spreading activation values but what matters is the variance between the nodes as well. Variance is nothing but the value which represents how frequently the activation value of that particular node changes. A low variance indicates stability in the data whereas a large variance indicates instability in the data. The dataset for the users is generated randomly. Firstly the variance for a node is generated randomly and a normal distribution is constructed with this variance. Now the variation of the node values over time is drawn from this normal distribution. This is done for all nodes of a user, for all users. All experiments are to be performed on this dataset generated, in which the importance or role of variance in friend recommendation shall be gleaned.

6.2 LANGUAGE USED FOR IMPLEMENTATION

6.2.1 Python

Python is interpreted, and one of the multi-paradigm that uses or conforming to more than one paradigm programming language. In python Object-oriented programming is supported and also the python programming language completely supports structured programming. Python programming language offers its users a feature for some support for functional programming in the tradition of Lisp. It provide it with map, filter, and reduce functions; list comprehensions, sets, dictionaries, and generator expressions.

6.2.2 Input Data

For experimental purpose, random data is generated for nodes users over a period of 1000 iterations. Considering that there can be 2 types of users, one type of user whose interest gets changed rapidly and another set of users whose interest don't gets changes so quickly. Based on this, nodes values are being initialized accordingly. Another data being generated is related to distance between the nodes. The least distance among two leaf nodes in each ontology graph is observed to be 2 whereas greatest distance between ontology graphs will be 6.

6.2.3 Recommendation System Evaluation

In our domain ontology, there are total 270 concept classes spread across 5 different hierarchical levels and each class represent one node in the ontology graph. Experiment is performed to simulate the end user behavior and generated the data for 1 end user .In the experiment, there are 100 nodes and stored data of 100 nodes in 1000 iterations. First the individual distances between pairs of all 100 nodes in our ontology hierarchy is stored. In our simulation, 2 types of nodes are being generated, one which are changed rapidly (30%) and the remaining one's (70%) which are changed slowly. For 30 % nodes, data is being incremented by probability of 0.2 and decremented by probability of 0.1. For remaining 70% nodes, value is incremented with probability of .001 to 0.099. After simulating data of 1 user in 1000 iterations, distance is calculated between ith and (i-1)th iteration in two experiments. In 1st experiment (Algorithm 1), top 5 nodes on the basis of spreading activation score are considered and in second experiment (Algorithm 2), top five nodes on the basis of spreading activation and variance are being considered.

The reason for taking the variance square factor in to account to find out the top 5 nodes which are being activated regularly on account of article read by the user related to that node. In both the cases, distance is plotted to show difference in top nodes in different iterations. In case of experiment without variance, distance between top nodes is not stable (Figure 6.5). Whereas , in case of experiment with spreading activation and variance square(Figure 6.6,6.7,6.8), distance between top nodes of different iteration is getting stable with time which is showing the normalization of the high effect to randomly chosen articles. Higher the variance square factor, earlier the distance between the nodes is getting stabilized (Figure 6.6). Once the interest of 1 user is known, in similar way, the interest of other users can be find out as well and finally provide friend recommendation to users with similar interest.

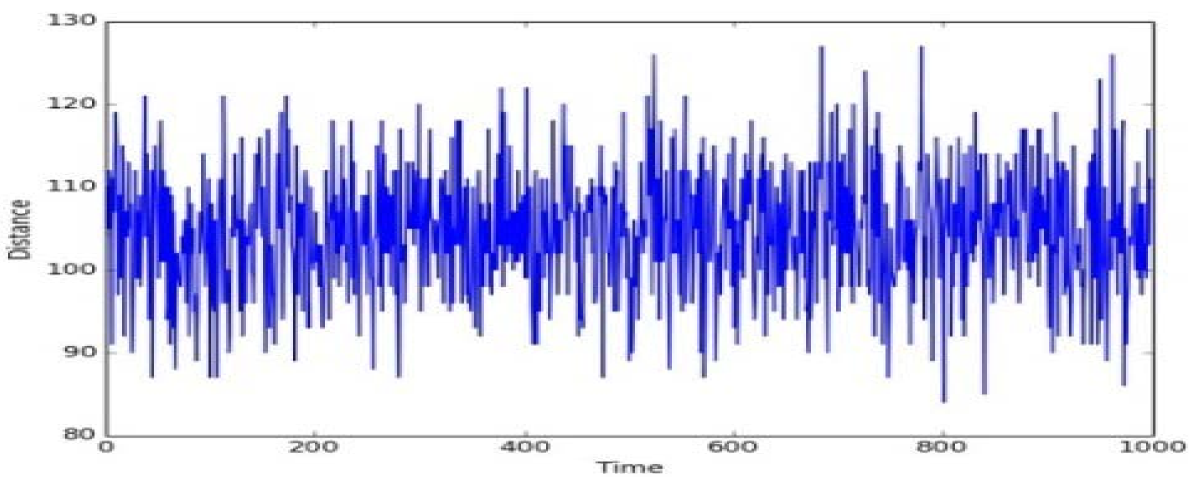

Figure 6.5 Distance between top 5 nodes without Variance

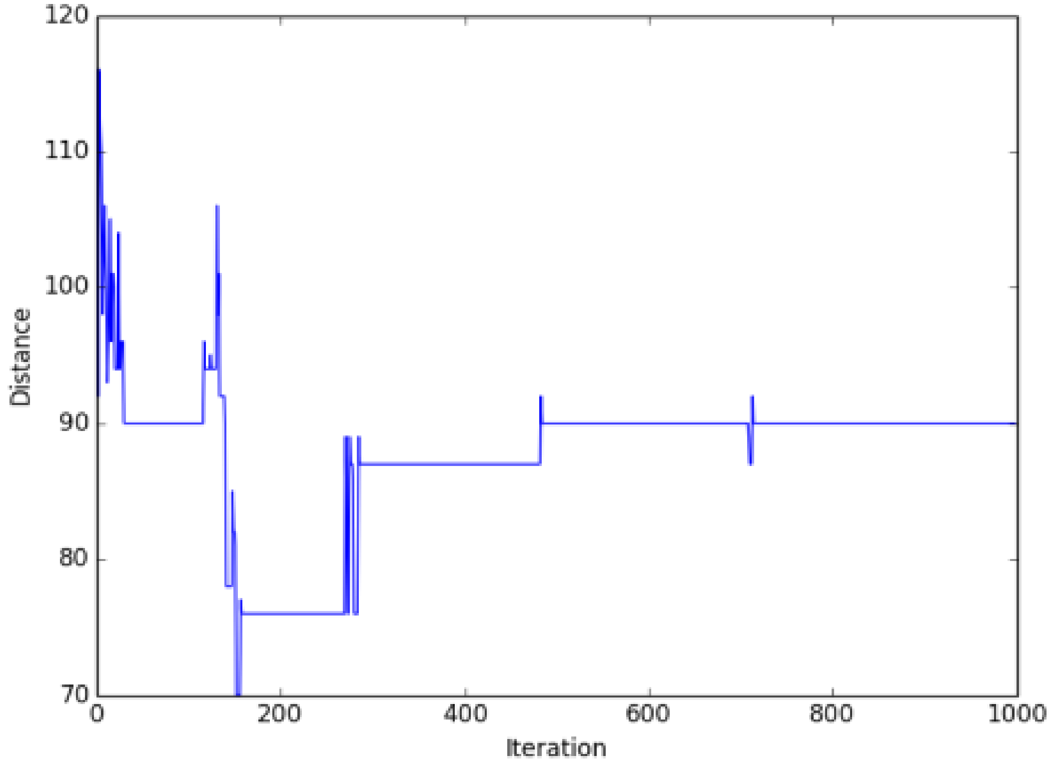

Figure 6.6 Distance between top 5 nodes with 90% Variance Square

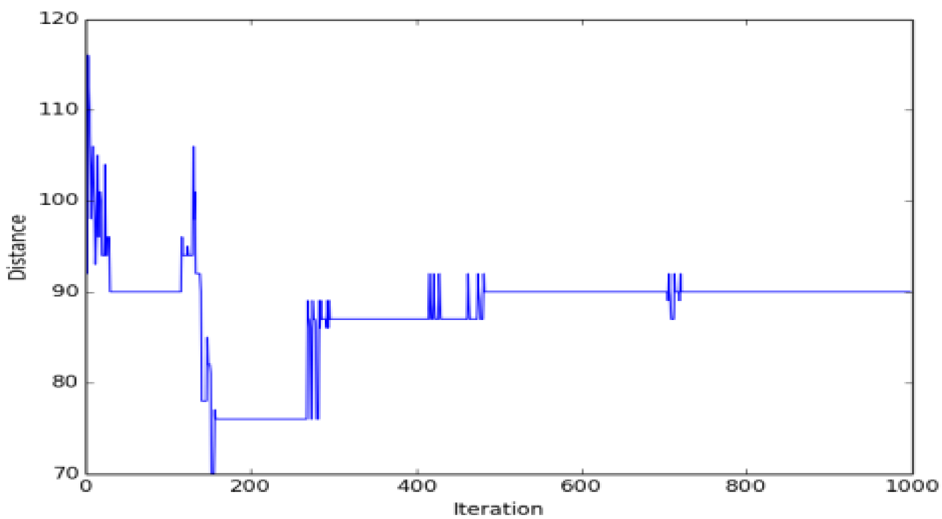

Figure 6.7 Distance between top 5 nodes with 40 % Variance Square

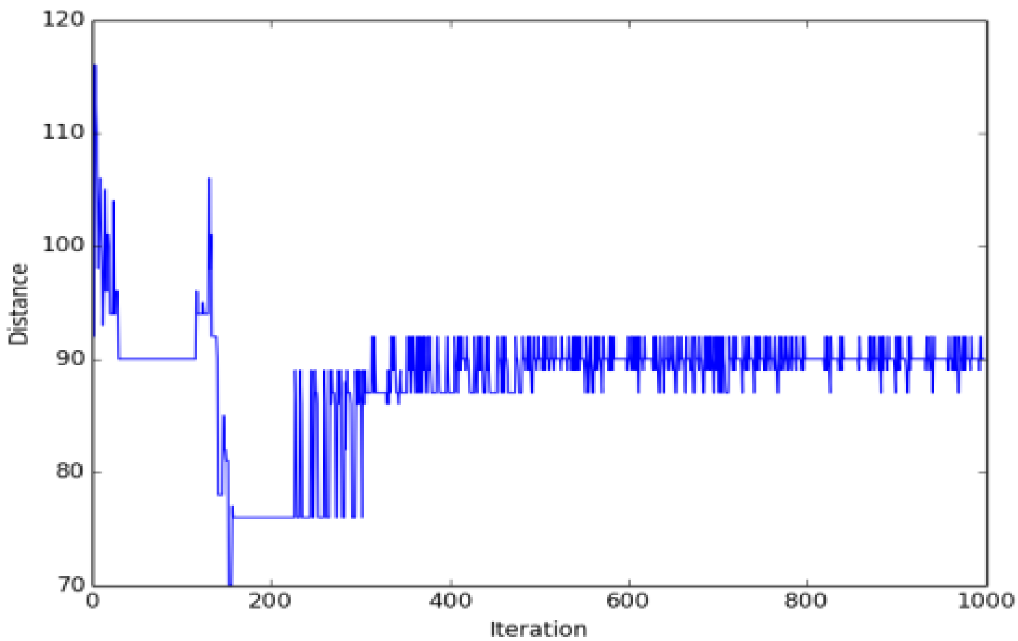

Figure 6.8 Distance between top 5 nodes with 10 % Variance Square

In the simulation, Friend Recommendation for a user is done by us utilizing the Spreading Activation Graph generated for respective users over time. For each user a Spreading Activation Graph will be generated. This graph is nothing but the Ontology classes being populated with values signifying the level of interest the user might hold in them. These interests might be explicit i.e. specified by the user at the start, or these might be implicit i.e. inferred through the Spreading Activation cycles which run at regular intervals based on the user's activity. These interests accumulate values over time (or might decay, based on the user's interest increasing or declining) over multiple Spreading Activation cycles. Hence end up with zero and non-zero values for the particularly that elements inside the Ontology that are present. The interests of the user is being represented using these values, rather than the user profile itself.

Now once user profile of a user is updated based on his current interest area according to Spreading Activation graph, Friend Recommendation is done by comparing his/her user profile by comparing profiles of various other users. Let us take example of two users U_1 and U_2 having user profiles UP_1 and UP_2 respectively. In order to determine that U_1 and U_2 can be recommended to each other as friends, the User profiles UP_1 and UP_2 are compared and how similar they are based on their profiles conclusion can be made that they should be recommended to each other or not. The metric to measure the similarity of two user profiles is Cosine Similarity. As done in the simulation of one end user to generate use profile based on user current interest, same has been extended by considering ten Users having interest in six domains (Table 6.6).Now based on ten users having interest in six domains (I_1 to I_6), similarity between user is calculated by considering the cosine similarity between six interest areas.(Table 6.7).User U_4 and U_5 have similar interest in all the Items from I_1 to I_6 so the cosine similarity between them is almost same(Table 6.7).

The similarity in Cosine angle is measured as the relationship among two various vectors through computing the angle of cosine between them. The cosine of 90° is 0 and cosine of 0° is 1. It is thus an indication of orientation and not magnitude. It means two vectors having a cosine similarity of 1 will have the same orientation whereas two vectors will have a similarity of 0 which are having orientation of 90° relative to each other, and two vectors have a similarity of -1 which are opposite to each other.

By utilizing the formula provided by Euclidean which is Euclidean dot product formula the cosine among two vector or two non-zero vectors can be obtained as:

$$\mathbf{U_1 . U_2} = \|\mathbf{U_1}\|\|\mathbf{U_2}\| \cos(\boldsymbol{\Theta}) \tag{6.1}$$

When we are given two vectors with attributes, U1 and U2, the angle of cosine and similarity, among them is calculated as cos (Θ), and is expressed utilizing a dot product furthermore where magnitude is as

$$\mathbf{similarity} = \mathbf{cos}(\boldsymbol{\Theta}) = \frac{\mathbf{U_1 . U_2}}{\|\mathbf{U_1}\|\|\mathbf{U_2}\|} = \frac{\sum_{i=1}^{n} \mathbf{U_{1i}} \, \mathbf{U_{2i}}}{\sqrt{\sum_{i=1}^{n} \mathbf{U_{1i}^2}} \sqrt{\sum_{i=1}^{n} \mathbf{U_{2i}^2}}} \tag{6.2}$$

Whereas U_{1i} is component of vector U_1 and U_{2i} are components of vector U_2 respectively. Thus angle of cosine relationship gives us the strength of association among two outlines hence on the foundation based above it can decide if two users should be recommended to each other or not.

	I1	I2	I3	I4	I5	I6
U1	2.369	2.320	2.799	2.556	2.938	2.482
U2	2.45	2.073	2.052	2.506	2.779	2.777
U3	2.512	2.328	2.570	2.840	2.004	2.845
U4	0.973	4.934	4.294	0.857	0.582	0.691
U5	0.760	4.418	4.701	0.922	0.734	0.761
U6	4.752	0.830	0.803	0.823	0.739	4.207
U7	0.329	0.174	0.488	0.294	4.848	0.452
U8	0.355	0.119	0.476	4.755	0.118	0.385
U9	0.452	0.384	0.234	0.200	0.059	4.860
U10	4.802	0.224	0.200	0.426	0.085	0.488

Table 6.6 Users Interest Matrix

	U1	U2	U3	U4	U5	U6	U7	U8	U9	U10
U1	1	0.99	0.98	0.73	0.75	0.72	0.59	0.52	0.49	0.48
U2	0.99	1	0.98	0.66	0.67	0.78	0.59	0.53	0.56	0.52
U3	0.98	0.98	1	0.73	0.74	0.78	0.46	0.57	0.56	0.52
U4	0.73	0.66	0.73	1	0.99	0.36	0.2	0.22	0.2	0.22
U5	0.75	0.67	0.74	0.99	1	0.35	0.22	0.24	0.21	0.19
U6	0.72	0.78	0.78	0.36	0.35	1	0.24	0.24	0.72	0.8
U7	0.59	0.59	0.46	0.2	0.22	0.24	1	0.1	0.11	0.1
U8	0.52	0.53	0.57	0.22	0.24	0.24	0.1	1	0.13	0.17
U9	0.49	0.56	0.56	0.2	0.21	0.72	0.11	0.13	1	0.2
U10	0.48	0.52	0.52	0.22	0.19	0.8	0.1	0.17	0.2	1

Table 6.7 User Cosine Similarity Matrix

Chapter Seven: Movie Recommendation to Friends using Whale Optimization Algorithm

This Chapter will explain about the Whale Optimization Algorithm used to recommend movies to friend once our proposed recommendation system will provide the friend recommendation to end users as explained in previous chapter. MAE, Recall and Precision of WOA is compared with other existing algorithm like, K-Means, SOM, PCA and ABC. Results shows that WOA is better than existing meta-heuristic algorithms.

7.1 Introduction

Recommendation systems produce and is being presented in the discussion with this intent of knowledge accumulated in these social media sites. In the recent years, the advances in the wireless devices and GPS have created new opportunities for the friends' recommendation system. In the social media sites, users shares their interest, locations, and outings. Some other users may share their contents to expand it. The users get a number of movie recommendations from the social media friends and other platforms. However, the task of recommendation system is challenging since the users interest is dynamic and changes frequently with the time.

Two decades ago, generally the recommendation was done by the neighborhood or with the persons work together. These friends are called as G-friends, which mean the friends influenced based on geographical distance. With the advancement in the social media services such as, LinkedIn, Myspace or Twitter, a revolution has come in the worldwide in the recommendation system. As per the Facebook data, one person has approximately 130 friends, which is much larger figure than two decades ago data (Netflix india U watch tv programmes online, n.d.)and they used to recommend the items such as movies or other things.

One of the promising challenges with the recommendation using social media is that it may not be as per the user's interest. Often, people lie on the existing online platforms for getting the idea of some product. A number of recommendation systems are available which recommends the products on the basis of rating. However, these methods of recommendation are not suitable to fulfill the contemporary demands (Giddens, 1991; Røpke, 1999; Huynh, Fritz, & Schiele, 2008; Tomlinson, 2003)to overcome the above mentioned problem, A proposal for a novel movie recommendation system using the whale optimization algorithm is being discussed in this paper.

WOA (Mirjalili & Lewis, 2016)is a population based algorithm developed in 2017. It mimics the hunting mechanism of humpback whales. Recently, WOA has shown competitive performance on a number of benchmark and outperformed other popular meta-heuristics for wide application areas, data clustering, image, and optimization and computer vision. Swarm of whales mathematically simulates the hunting behavior WOA. The WOA explores and exploits the search space to find the optimal solution. It performs encircling and spiral movements to search the prey. This paper utilizes the ability of WOA to optimize the recommendation process. The similarity among the users has been maximized with WOA to optimize the recommendation process.

7.2 Whale Optimization Algorithm for Friend Recommendation

Whale Optimization Algorithm also is known as WOA is freshly generated meta-heuristic algorithm impersonated by (Mirjalili & Lewis, 2016). It is inspired from the hunting behavior of humpback whales which is called as bubble-net style hunting. Generally, whales encircle the prey using the bubbles, which makes a ’circular’ path. The humpback whales hunts small fishes or which are found in the seas. The whales used to plunge deep in the water and come back to catch the prey which forms a spiral shape. This process of whales is simulated as exploitation phase; however the exploration phase is modeled by the random walks. The mathematical simulation is explained as follows:-

7.2.1 Exploitation phase

In WOA, the current best solution is considered as prey near to optimum solution. Each whale represents a search agent and the position of each agent is updated in two ways: i) encircling and ii) Spiral shaped. Let the $\vec{W}(i)$ and $\vec{W}b(i)$ represents the position i^{th} search agent and best search agent respectively. Then the position in of agent in the next iteration is defined by Eq. (7.1).

$$\vec{\boldsymbol{X}}(\boldsymbol{t}+\mathbf{1}) = \overrightarrow{\boldsymbol{X}^{*}}(\boldsymbol{t}) - \vec{\boldsymbol{A}}.\vec{\boldsymbol{D}} \tag{7.1}$$

Where, $\vec{A}$ denotes coincident vector defined by Eq. (7. 2) and D represents the absolute distance between the best and current solution defined by Eq. (7.3).

$$\vec{\boldsymbol{A}} = \mathbf{2}.\vec{\boldsymbol{a}}.\vec{\boldsymbol{r}}.-\vec{\boldsymbol{a}} \tag{7.2}$$

$$\vec{\boldsymbol{C}} = \mathbf{2}.\vec{\boldsymbol{r}} \tag{7.3}$$

Where, $r \in (0, 1)$ denotes a random number, $a \in (2, 0)$, and $\vec{C}$ is known as adjustment vector.

Thereafter, the distance (D´) of prey with search agent is computed as D´ = $\vec{X}$ b(i) − $\vec{X}$ (t), after that the spiral movement is defined using Eq. (7.5).

$$\vec{\boldsymbol{X}}(\boldsymbol{t}+\mathbf{1}) = \vec{\boldsymbol{D'}}.\boldsymbol{e}^{\boldsymbol{bl}}.\mathbf{cos}(\mathbf{2}\boldsymbol{\pi l}) + \vec{\boldsymbol{X^*}}(\boldsymbol{t}) \quad (7.4)$$

$$\vec{\boldsymbol{D'}} = |\vec{\boldsymbol{X^*}}(\mathbf{t}) - \vec{\boldsymbol{X}}(\boldsymbol{t})| \quad (7.5)$$

Where, m ∈ (−1, 1) represents a random generated number, b is basically a constant which defines spiral shape.

The encircling and spiral phases defines the exploitation of WOA with the similar probabilities as given in Eq. (7.6).

$$\vec{\boldsymbol{X}}(\boldsymbol{t}+\mathbf{1}) = \begin{cases} \vec{\boldsymbol{X^*}}(\boldsymbol{t}) - \boldsymbol{A}.\vec{\boldsymbol{D}} & , \ p < 0.5 \\ \vec{\boldsymbol{D'}}.\boldsymbol{e}^{\boldsymbol{bl}}.\mathbf{cos}(\mathbf{2}\boldsymbol{\pi l}) + \vec{\boldsymbol{X^*}}(\boldsymbol{t}), & p \geq 0.5 \end{cases} \quad (7.6)$$

Where, $p \in (0, 1)$ is random number.

7.2.2 Exploration phase

The WOA defines the positions of each search agent by best and random search agent. A is used to make the decision that which agent will be used to defining new position. If A> 1than global search is performed whiles in rest cases whales does local search. The updated positions are mathematically defined using Eq (7.7) and (7.8).

$$\vec{\boldsymbol{D'}} = |\vec{\boldsymbol{X^*}}(\mathbf{t}) - \vec{\boldsymbol{X}}(\boldsymbol{t})| \quad (7.7)$$

$$\vec{\boldsymbol{X}}(\boldsymbol{t}+\mathbf{1}) = \vec{\boldsymbol{X^*}}(\boldsymbol{t}) - \vec{\boldsymbol{A}}.\vec{\boldsymbol{D}} \quad (7.8)$$

Here, $\vec{W}$ rand demonstrate any randomly chosen whale from the search space. The complete procedure of the WOA is presented in Algorithm 1.

The calculation time of the recommended approach is based on that clustering is directly proportionate to the quantity of groups and data subjects. In the particular paper, as mentioned, WOA generates the optimal cluster centroids with $O(N \times K \times t)$ operations for t iterations, where N is used to denotes the number of information points and K is used to mean the quantity of clusters that are required. Therefore, for the population of size P, the complete time complexity regarding the suggested clustering method is BigO('P' × 'N' × 'K' ×'t').

7.2.3 WOA Based Movie Recommendation

The proposed recommendation system leverages the strengths of WOA algorithm to avoid local optima. K-means is a popular clustering technique; however it often trap into local optima.

To rectify this concern WOA algorithm is used in which gas ability to bypass local optimum solution. The WOA is particularly being used for grouping and clustering the users based upon some choices or interest that is quite similar in nature or similarity. Initially, these random cluster centers are initialized which are optimized with the course of iterations using the WOA. In the WOA based clustering, the position X of each whale denotes a group of centroids ($X_1, X_2, X_3, \cdots, X_K$), wherever "K" is used to represents the quantity from clusters. The fitness functions provided by the WOA based clustering is further defined in Eq. (7.9), which represents the intra-cluster distance.

$$\boldsymbol{f(Z, C) = \sum_{i=1}^{K} \sum_{Z_i \in C_i} (d(Z_i, C_i))^2} \quad (7.9)$$

In the end of the algorithm the final cluster centroids are represented by the position of the whale with the best fitness value. The pseudo-code of the WOA based clustering method same as described in Algorithm 1. Further, to test the validity of proposed system the precision, recall and mean absolute error are calculated using the Eq. (7.9) and Eq. (7.10 -7.12) respectively.

$$\boldsymbol{Precision = |INTERSECTING \cap TopN|/N} \quad (7.10)$$

$$\boldsymbol{RECALL = |INTERSECTING \cap TopN|/Intersecting} \quad (7.11)$$

$$\boldsymbol{MAE = \sum |P_{ij} \cap r_{ij}|/M} \quad (7.12)$$

7.3 Performance Analysis

The proposed WOA based method has been tested on Movie-lens dataset which is publicly available on https://grouplens.org/datasets/movielens/.

Algorithm 1 WOA

1: **Input:** Randomly distributed swarm (W*i)* of whales where i: = 1, 2..., n

2: **Output:** The position of whale with best fitness W *

3: Define the fitness of each whale and find the position of (W *)

4: **while** (Max Iterations are not reached) **do**

5: **for** each whale **do**

6: Update m, p, $\vec{A}$, $\vec{C}$, and $\vec{a}$

if ($p < 0.5$) **then**

if $|A| < 1$ **then**

perform encircling and redefine the location of each whale using Eq. (1)

else if $|A| \geq 1$ **then**

Find random whale (W_{rand})

Perform exploration and redefine the location of each whale using Eq. (8)

end if

else if ($p \geq 0.5$) **then**

Perform spiral movements and redefine the location of the whales using Eq. (5)

end if

end for

Perform bound checks

Calculate the fitness of each whale find (W^*).

$i = i+1$

end while

Return W^*

The movie-lens data set contains 100,000 data points along with the ratings given by users of various movies. There are 1000 users in the dataset and 1700 movies in the dataset. Each user has given rating to at least 20 movies. The experimental results are performed, and these outputs are then matched among those five other methods that are present within that particular research namely, K-mean, PCA-K- mean, SOM, PSO, and ABC during the duration of Mean absolute error(MAE) it is also known as MAE, Precision and Recall. Table 7.1 presents the MAE, Precision, and Recall value of each method on the deferent number of clusters. As it is depicted from the table, the MEA value of WOA based method is minimum on all the groups except 5. However, when the number of clusters is set as 5 the ABC has outperformed all the considered methods. Further, it is concluded that the proposed method has outperformed other algorithms on 88.5% cases. Moreover, the precision value of the WOA as depicted in table 7.3 is also maximum on all the number of clusters except 5. For 5 numbers of clusters PSO has outperformed other methods. Additionally, the performance of WOA is further analyzed in terms of recall rate, as bestowed inside table 7.2. While that signifies and clearly being observed by that table, where the WOA has given competitive results in terms of recall, whereas ABC as performed well, when the number of clusters was set as 4. Thus it is concluded that the WOA method can be served as an effective recommendation system to solve many world problems.

No. Clusters	5	10	15	20	25	30	35	40
K-means	0.824	0.824	0.820	0.820	0.817	0.815	0.813	0.800
PCA-K-means	0.852	0.843	0.842	0.838	0.838	0.820	0.819	0.830
SOM	0.821	0.820	0.820	0.817	0.815	0.813	0.815	0.843
GA	0.752	0.751	0.742	0.742	0.732	0.732	0.730	0.732
PSO	**0.742**	0.750	0.741	0.740	0.738	0.732	0.730	0.734
ABC	0.775	0.765	0.765	0.738	0.760	0.782	0.780	0.786
WOA	**0.742**	**0.734**	**0.731**	**0.730**	**0.725**	**0.720**	**0.710**	**0.730**

Table 7.1 MAE of the WOA and other considered methods

No. Clusters	5	10	15	20	25	30	35	40

K-means	0.111	0.142	0.123	0.112	0.113	0.117	0.116	0.118
PCA-K-means	0.111	0.142	0.123	0.112	0.113	0.117	0.116	0.118
SOM	0.345	0.342	0.348	0.335	0.321	0.328	0.324	0.314
GA	0.355	0.362	0.368	0.372	0.361	0.352	0.370	0.375
PSO 0.358	0.376	0.376	0.381	0.384	0.390	0.387	0.390	0.370
ABC	0.378	0.389	0.398	0.410	0.412	0.416	0.423	0.420
WOA	**0.459**	**0.480**	**0.450**	**0.425**	**0.490**	**0.420**	**0.430**	**0.440**

Table 7.2 Recall of the WOA and other considered methods

No. Clusters	3	4	5	6	7	8	9	10
K-means	0.050	0.050	0.06	0.072	0.078	0.082	0.087	0.097
PCA-K-means	0.070	0.075	0.078	0.082	0.088	0.090	0.093	0.098
SOM	0.105	0.110	0.130	0.182	0.205	0.250	0.300	0.400
GA	0.150	0.160	0.162	0.272	0.280	0.305	0.450	0.500
PSO 0.105	0.110	0.180	0.282	0.310	0.450	0.520	0.610	0.620
ABC	**0.312**	0.315	0.370	0.482	0.410	0.450	0.520	0.630
WOA	0.310	**0.340**	**0.470**	**0.490**	**0.515**	**0.550**	**0.570**	**0.650**

Table 7.3 Precision of the WOA and other considered methods

Chapter Eight: CONCLUSION AND FUTURE SCOPE

Finally, in this last chapter, conclusion about the research done in the thesis is presented. Also, it gives a short summary of the important points which are focused in the thesis and sum up the important contributions given in this thesis. Absolutely, no research can't be granted as complete as there is always an enhancement scope in the research, hence the closing chapter summarize few potential research directions worth exploring later on.

8.1 Conclusion

This thesis based on the issue related to the friend recommendation based on the user current interest. As mentioned in previous chapters, finding a friend based on the user interest can be difficult as well as extremely complicated task for end users. Recommending friends to end users from a list of available users in order to meet user individual needs is a real challenge job. The existing algorithms are majorly based on static profile based comparison or location based comparison are not accurate enough to provide correct recommendation about the friends.

This thesis targets on the below three essential perspectives:

1. Propose and developed a hobby ontology which will be the base for recommendation system framework.

2. Developed a novel Content Recommendation system using Hobby Ontology and Spreading Activation on top of it to recommend news content to end user based on user changing interest.

3. Proposed and implement a Friend recommendation system based on the user interest which is captured using the content recommendation system.

The major contribution of the thesis is that it help to understand user changing interest based on the news content reading habit of end user. The contents are being provided by the content recommendation system which learns the user behaviour or interest continuously and based on the user interest it recommends the content to the end users.

Especially, the task done in the thesis has the below contributions:

1. It helps in adding the expertise in the area of ontology, semantic networks and recommendation systems. To know about the different type of recommendation systems and their shortcoming to address the research problem.

2. Hobby ontology considering different area of user interest has been created from the scratch as it is not developed and available till now. Same ontology can be extended based on the research requirement and can be used by any other researcher in this field.

3. To develop a unique recommendation system framework by bringing together the capability of ontology, semantic network and recommendation system techniques. The recommendation system framework comprises of understanding user behaviour and feeding contents to user accordingly and based on the user interest giving friend recommendation to the end user.

Understanding the user behaviour to know about the user current interest is the most challenging task of the recommendation system. In the research work, new ontology has been developed related to user interest. To know about the user interest, news content has been provided to end user. To know the user dynamically changing interest, spreading activation algorithm runs on top of ontology graph. Based on the user habit of reading, liking and disliking of content, corresponding nodes and neighbouring nodes of the ontology graph gets activated using spreading activation algorithm. The combination of ontology graph and Spreading activation algorithm helps to overcome the shortcomings of traditional recommendation system used in friend recommendation. Instead of using the static profile of user to provide content and friend recommendation, dynamically generated profile is being used to provide efficient and accurate recommendations to users.

8.2 Future Work

Even though with the experimental results it's proven that the proposed recommendation outcome is positive but still there is scope of improvement in following different aspects.

1. As of now the current framework is not fully automated as there is no GUI for client side. User behaviour is predicted using implicit and explicit probability. With the help of Client side GUI, we can ask user to enter his profile details initially. Based on the initial profile information , will start recommending content and based on user click and feedback about the recommended article can recommend new content and able to predict his interest area.

2. Currently hobby ontology contains 269 different hobbies segregated in to different type of interest. In future, it can be extended further by adding other category of hobbies which are not covered yet.
3. In the current recommendation system, content recommended to users are only in the English language. In future, we can extend it to multi-lingual as well.

REFERNCES

Abbassi, Z., Yahia, S. A., Lakshmanan, L., Vassilvitskii, S., & Yu, C. (2009). Getting recommender systems to think outside the box. *Proceedings of the third ACM conference on Recommender systems.* New York,USA.

Adiele, C., & Penner, W. (2006). Monitoring interactivity in a web-hased community. *Proceedings of the 7th International Conference on Web Information Systems Engineering.*

Adomavicius, G., & Tuzhilin, A. (2005). Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions. *IEEE Transactions on Knowledge and Data Engineering, 17*(6), 734–749.

Agre, G., & Dongre, S. (2015). A Keyword Focused Web Crawler Using Domain Engineering and Ontology. *International Journal of Advanced Research in Computer and Communication Engineering, 4*(3), 463–465.

Ahmed, A., Kanagal, B., Pandey, S., Josifovski, V., Pueyo, L., & Yuan, J. (2013). Latent factor models with additive and hierarchically-smoothed user preferences. *Proceedings of the Sixth ACM International Conference on Web Search and Data Mining, WSDM '13.* NY, USA.

Almohammadi, K., Hagras, H., Alghazzawi, D., & Aldabbagh, G. (2017). A survey of artificial intelligence techniques employed for adaptive educational systems within e-learning platforms. *Journal of Artificial Intelligence and Soft Computing Research, 7*(1), 47–64.

Amini, B., Ibrahim, R., & Othman, M. S. (2011). Discovering the impact of knowledge in recommender systems: A comparative study. *International Journal of Computer Science & Engineering Survey, 2*(3).

Anderson, J. R. (1983). A spreading activation theory of memory. *Journal of Verbal Learning and Verbal Behaviour, 22*, 261-295.

Anderson, J. R. (1995). *Cognitive Psychology and its Implications.* UK: Worth Publishers.

Anderson, T. C. (2000). The body and communities in cyberspace: A Marcellian analysis. *Ethics and Information Technology, 2*(3), 153-158.

Antoniou, G., & Van Harmelen, F. (2008). *A semantic Web primer.* USA: Choice Reviews Online.

Aula, A., & Nordhausen, K. (2006). Modeling successful performance in Web searching. *Journal of the American Society for Information Science and Technology, 57*(12), 1678-1693.

Bakardjieva, M. (2003). Virtual Togetherness: an Every-day Life Perspective. *Media, Culture & Society, 25*(3), 291-313.

Balasubramanian, S., & Mahajan, V. (2001). The Economic Leverage Of The Virtual Community. *International Journal of Electronic Commerce, 5*(3), 103-110.

Barab, S., Makinster, J., Moore, J., & Cunningham, D. (n.d.). Designing 8c Building an On-line Community: The Struggle to Support Sociability in the Inquiry Learning Forum. *Educational Technology Research and Development, 4*(49), 71-96.

Barrag´ans-Mart´ınez, A., Costa-Montenegro, E., Burguillo, J., Rey-L´opez, M., Mikic-Fonte, F., & Peleteiro, A. (2010). A hybrid content-based and item-based collaborative filtering approach to recommend tv programs enhanced with singular value decomposition. *Information Sciences, 180*(22), 4290–4311.

Bell, R. M., & Koren, Y. (2007). Improved neighborhood-based collaborative filtering. *KDD Cup and Workshop at the 13th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining.*

Bell, R. M., & Koren, Y. (2007). Scalable collaborative filtering with jointly derived neighborhood interpolation weights. *Data Mining, 2007. ICDM 2007. Seventh IEEE International Conference.* Omaha, NE.

Benjamins, V., Davies, J., Baeza-Yates, R., Mika, P., Zaragoza, H., & Greaves, M. (2008). Near-Term Prospects for Semantic Technologies. *IEEE Intelligent Systems, 23*(1), 76-88.

Bhowmick, P. K., Sarkar, S., & Basu, A. (2010). Ontology based user modeling for personalized information access. *IJCSA*, 1–22.

Bian, L., & Holtzman, H. (2011). Online Friend Recommendation through Personality Matching and Collaborative Filtering. *The Fifth International Conference on Mobile Ubiquitous Computing, Systems, Services and Technologies.* Lisbon, Portugal.

Blanchard, A. (2004). Virtual behavior settings: An application of behavior setting theories to virtual communities. *Journal of Computer Mediated Communication, 9*(2).

Boothroyd, P., & Davis, H. C. (1993). Community Economic Development: Three Approaches. *Journal of Planning Education and Research, 12*, 230-240.

Borst, W. N. (1997). *Construction of engineeriing ontologies for knowledge sharing and reuse.* The Netherlands: SIKS, the Dutch Graduate School for Information and Knowledge Systems.

Boyd, D., & Ellison, N. (2007). Social Network Sites; Definition, History, and Scholarship. *Journal of Computer-Mediated Communication, 13*(1), 210-230.

Breese, J. S., Heckerman, D., & Kadie, C. (1998). Empirical analysis of predictive algorithms for collaborative filtering. *Proceedings of the Fourteenth conference on Uncertainty in artificial intelligence, UAI'98.* San Francisco,CA,USA.

Burke, R. (2002). Hybrid recommender systems: Survey and experiments. *User modeling and user-adapted interaction, 12*(4), 331-370.

Burke, R. (2007). *Hybrid web recommender systems.* The adaptive web,Springer.

Burnett, G. (n.d.). Information Exchange in virtual communities: a typology. *Information Research : An International Electronic Journal, 5*(4).

Cacheda, F., Carneiro, V., Fern´andez, D., & Formoso, V. (Feb. 2011). Comparison of collaborative filtering algorithms: Limitations of current techniques and proposals for scalable, high-performance recommender systems. *ACM Transactions on the Web, 5*(1), 2-33.

Cantador, I., Bellogín, A., & Vallet, D. (2010). Content-Based Recommendation in Social Tagging Systems. *Proceedings of the Fourth ACM Conference on Recommender systems.*

Carlen, U. (2002). Typology of OLC. *Proceeding of the first International Conference of NetLearning2002.* Ronneby, Sweden.

Carver, C. (1999). Building a virtual community for a tele-Learning environment. *IEEE Communications Magazine, 37*(3), 114-118.

Case, S., Azarmi, N., Thint, M., & Ohtani, T. (2001). Enhancing E-communities with Agent-based systems. *Computer, 34*(7), 64-69.

Chang, E., Scharl, A., Weichselbraun, A., & Liu, W. (2005). Semi-Automatic Ontology Extension Using Spreading Activation. *Journal of Universal Knowledge Management, 0*(1), 50-58.

Chang, P., Lin, C., & Chen, M. (2016). A hybrid course recommendation system by integrating collaborative filtering and artificial immune systems. *Algorithms, 9*(3), 18.

Chen, J., Huang, J., & Min, H. (Dec 2008). Easy recommendation based on probability model. *Semantics, Knowledge and Grid, 2008. SKG '08. Fourth International Conference.* Beijing, China.

Chen, Y., Wu, C., Xie, M., & Guo., X. (2011). Solving the sparsity problem in recommender systems using association retrieval. *Journal of computers, 6*(9), 1896–1902.

Chu, C.-H., Wu, W.-C., Wang, C.-C., Chen, T.-S., & Chen, J.-J. (2013). Friend Recommendation for Location-Based Mobile Social Networks. *Seventh International Conference on Innovative Mobile and Internet Services in Ubiquitous Computing.* Taichung.

Coates, A. (2001). The Role of XML in Finance. *Proceedings of the XML Conference and Exposition 2001.* Orlando, FL, USA.

Coelho, B., Martins, C., & Almeida, A. (2010). Web intelligence in tourism: User modeling and recommender system. *Proceedings - 2010 IEEE/WIC/ACM International Conference on Web Intelligence.* Toronto,Canada.

Cole, A. L., & Knowles, J. G. (2001). *Lives in Context: The Art of Life History Research.* CA: AltaMira Press.

Collins, A. M., & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological review, 82*(6), 407.

Collins, A. M., & Quillian, M. R. (1969). Retrieval time from semantic memory. *Journal of verbal learning and verbal behavior, 8*(2), 240-247.

Cremonesi, P., Koren, Y., & Turrin, R. (2010). Performance of recommender algorithms on top-n recommendation tasks. *Proceedings of the fourth ACM conference on Recommender systems.* Barcelona, Spain.

Crestani, F. (1997). Application of spreading activation techniques in information retrieval. *Intelligence Review, 11*(6), 453–482.

Crestani, F. (1997). Retrieving documents by constrained spreading activation on automatically constructed hypertexts. *Proceeding of EU- FIT 97- Fifth International Congress on Intelligent Techniques and Soft Computing.* Aachen, Germany.

Daoud, M., Tamine, L., Boughanem, M., & Chebaro, B. (2007). Learning implicit user interests using ontology and search history for personalization. *Proceedings of the 2007 international conference on Web information systems engineering,WISE'07.* Berlin, Heidelberg.

Daramola, O., Adigun, M., & Ayo, C. (2009). Building an Ontology-Based Framework for Tourism Recommendation Services. *Information and Communication Technologies in Tourism.* The Netherlands.

Deshpande, M., & Karypis, G. (2004). Item-based top-n recommendation algorithms. *ACM Transactions on Information Systems, 22*(1), 143–177.

Di Bono , M., Pieri , G., & Salvetti, O. (n.d.). *WP9: A Review of Data and Metadata Standards and Techniques for Representation of Multimedia Content.* MUSCLE ProjectDeliverable, IST-CNR.

(n.d.). *Digital Market overview India Understanding the scale of change of online audiences and digital media in India.* Retrieved from http://www.http://www.digitalstrategyconsulting.com/documents/DSC-Research-Market_Overview-India_with_intro.pdf, Digital Strategy Consulting & Partners (2009).

Djuana, E., Xu, Y., Li, Y., & Josang, A. (2011). Ontology learning from user tagging for tag recommendation making. *Proceedings of the 2011 IEEE/WIC/ACM International Conferences on Web Intelligence and Intelligent Agent Technology.*

Durkheim, E. (1964). *The division of labor in society.* New York: Free Press of Glencoe.

Etzioni, A., & Etzioni, O. (1999). Face-to-Face and Computer-Mediated Communities, A Comparative Analysis. *The Information Society, 15*, 241-248.

Eysenbach, G., & Till, J. (2001). Ethical issues in qualitative research on internet communities. *British Medical Journal*, 1103-1105.

Facebook becomes bigger hit than Google. (2010). (Financial Times) Retrieved from http://www.ft.eom/intl/cms/s/2/67e89ae8-30f7-l ldfb05700144feabdc0.html#axzz23C8 olcjM

Fernback, J., & Thompson, B. (1995). Virtual Communities: Abort, Retry, Failure? *Computer Mediated Communication and the American Collectivity*.

Figallo, C. (1998). *Hosting Web Communities.* New York: John Wiley & Sons.

FOAF. (n.d.). Retrieved from www.foaf-project.org

Fortuna, B., Fortuna, C., & Mladenic, D. (2010). Realtime news recommender system. *Proceedings of the 2010 European conference on Machine learning and knowledge discovery in databases:Part III.* Athens, Greece.

Fox, N., & Roberts, C. (1999). GPs in Cyberspace: the Sociology of 'virtual Community. *The Sociological Review, 47*(4), 643-671.

Franz, R., & Wolkinger, T. (2003). Customer Integration with Virtual Communities. Case study: The online community of the largest regional newspaper in Austria. *Proceedings of the Hawai International Conference on System Sciences.* Big Island,Hawaii.

Gantner, Z., Drumond, L., Freudenthaler, C., Rendle, S., & Schmidt-Thieme, L. (2010). Learning attribute-to-feature mappings for cold-start recommendations. *Data Mining (ICDM), 2010 IEEE 10th International Conference.* Sydney, NSW.

Gauch, S., Speretta, M., Chandramouli, A., & Micarelli, A. (2007). User profiles for personalized information access. *The adaptive web ,Springer*, 54–89.

Gazzaniga, M. S., Ivry, R. B., & Mangun, G. R. (April 1998). *Cognitive Neuroscience, The Biology of the Mind.* W. W. Norton & Company.

Giddens, A. (1991). *Modernity and Self-Identity: Self And Society in the Late Modern Age.* US: Standford University.

Graham, S. (2007). Disagreeing to agree: Conflict, (im) politeness and identity in a computer-mediated community. *Journal of Pragmatics, 39*(4), 742-759.

Gruber, T. (1993). A Translation Approach to Portable Ontologies. *Knowledge Acquisition, 5*(2), 199-220.

Guo, G. (2012). Resolving data sparsity and cold start in recommender systems. *Proceedings of the 20th International Conference on User Modeling, Adaptation, and Personalization, UMAP'12.* Berlin, Heidelberg.

Guo, G. (2013). Improving the performance of recommender systems by alleviating the data sparsity and cold start problems. *Proceedings of the Twenty-Third international joint conference on Artificial Intelligence.* Beijing, China.

Gusfield, J. R. (1975). *The community: A critical response. .* New York: Harper Colophon.

Gusfield, J. R. (1978). *Community: A Critical Response.* USA: Harper & Row.

Hagel, J., & Armstrong, A. (1997). Net Gain: Expanding Markets through Virtual Communities. *Harvard Business School Press, Cambridge*.

Hahsler, M. (2011). *Developing and testing top-n recommendation algorithms for 0-1 data using recommenderlab.* NSF Industry University Cooperative Research Center for Net-Centric Software and System.

Hang, Y., Guiran, C., & Xingwei, W. (2009). A cold-start recommendation algorithm based on new user's implicit information and multi-attribute rating matrix. *Hybrid Intelligent Systems, 2009. HIS'09. Ninth International Conference.* Shenyang, China.

Hanson, W. (2000). *The Principles of Internet Marketing.* Cincinnati, Ohio: South-Western College Publishing.

Harrison, S., & Dourish, P. (1996). Replace-in space: The roles of place and space in collaborative systems. *Proceedings of the 1996 ACM Conference On Computer Supported Cooperative Work.* Boston, MA, USA.

Hasan, M. (2003). A Spreading Activation Framework for Ontology-enhanced Adaptive Information Access within Organisations. *Proceeding of the Spring Symposium on Agent Mediated Knowledge Management (AMKM 2003).* Stanford University, California, USA.

Hauben, M., & Hauben, R. (1997). *Netizens: On the History and Impact of Usenet and the Internet.* Washington: IEEE Computer Society Press.

Herlocker, J. L., Konstan, J. A., & Riedl, J. (2000). Explaining collaborative filtering recommendations. *Proceedings of the 2000 ACM Conference on Computer Supported Cooperative Work, CSCW '00.* New York, NY, USA.

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

Hill, W., Stead, L., Rosenstein, M., & Furnas, G. (1995). Recommending and evaluating choices in a virtual community of use. *CHI '95: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 194–201.

Hiltz, S. R., & Wellman, B. (1997). Asynchronous Learning Networks as a Virtual Classroom. *Communications of the ACM, 40*(9), 44-49.

Hoffman, D., & Novak, T. (1996). Marketing in hypermedia computer-mediated environments: Conceptual foundations. *Journal of Marketing, 60*(3), 50–68.

Hoggett, P. (1997). *Contested communities.* United Kingdom: Bristol: Policy Press.

Hunt, S. D. (1991). *Modern marketing theory: Critical issues in the philosophy of marketing science.* cincinnati,Ohio ,USA: South-Western Publishing Co.

Hurley, N., & Zhang, M. (2011). Novelty and diversity in top-n recommendation – analysis and evaluation. *ACM Transactions on Internet Technology, 10*(1), 14-30.

Huynh, T., Fritz, M., & Schiele, B. (2008). Discovery of Activity Patterns using Topic Models. *Proceedings of the 10th international conference on Ubiquitous computing.* Seoul,Korea.

ICQ. (n.d.). Retrieved from http://www.icq.com/

Jang, H., Olfman, L., Ko, I., Koh, J., & Kim, K. (2008). The influence of online brand community characteristics on community commitment and brand loyalty. *International Journal of Electronic Commerce, 12*(3), 57-80.

Jannach, D., Zanker, M., Felfernig, A., & Friedrich, G. (2010). *Recommender systems: an introduction.* Cambridge University Press.

Jawaheer, G., Weller, P., & Kostkova, P. (2014). Modeling user preferences in recommender systems: A classification framework for explicit and implicit user feedback. *ACM Transactions on Interactive Intelligent Systems, 4*(2), 1-26.

Jones, Q. (2004). Virtual-communities, virtual settlements & cyber-archaeology: A theoretical outline. *Journal of Computer Mediated Communication, 3*(3).

Jones, Q., & Rafaeli, S. (2000). Time To Split, Virtually: 'Discourse Architecture' And 'Community Building' Create Vibrant Virtual Publics. *Electronic Markets, 10*(4), 214-223.

Jussupova-Mariethoz, Y., & Probst, A. (2007). Business concepts ontology for an enterprise performance and competences monitoring. *Computers in Industry, 58*(2), 118-129.

Kanagal, B., Ahmed, A., Pandey, S., Josifovski, V., Yuan, J., & Pueyo, L. (2012). Supercharging recommender systems using taxonomies for learning user purchase behavior. *Proceedings of the VLDB Endowment, 5*(10), 956–967.

Karvounarakis, G., Alexaki, S., Christophides, V., Plexousakis, D., & Scholl, M. (2002). A Declarative Query Language for RDF. *Proceedings of the 11th International World Wide Web Conference (WWW 2002).* Honolulu, HI,USA.

Katifori, V., Poggi, A., Scannapieco, M., Catarci, T., & Ioannidis, Y. (2005). OntoPIM: how to rely on a personal ontology for Personal Information Management. *Proceeding of the 1st Workshop on The Semantic Desktop.*

Kawtrakul, A., Suktarachan, M., & Imsombut, A. (2003). Automatic Thai Ontology Construction and Maintenance System.

Khandelwal, D., Shanbhag, D., Shriyan, A., Thorve, R., & Borse, Y. (2018). LeMeNo: Personalised News using Machine Learning. *Fourth International Conference on Computing Communication Control and Automation (ICCUBEA).* Pune, India.

Kim, H. R., & Chan, P. K. (2008). Learning Implicit User Interest Hierarchy for Context in Personalization. *Applied Intelligence, 28*(2), 153-166.

Kim, Y. S. (2013). Recommender system based on product taxonomy in e-commerce sites. *Journal of Information Science and Engineering, 29*(1), 63-78.

Koenigstein, N., Dror, G., & Koren, Y. (2011). Yahoo! music recommendations: Modeling music ratings with temporal dynamics and item taxonomy. *Proceedings of the Fifth ACM Conference on Recommender Systems, RecSys '11.* New York, NY, USA.

Koren, Y. (2008). Factorization meets the neighborhood: multifaceted collaborative filtering model. *Proceedings of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD '08.* New York, NY, USA.

Koren, Y. (2010). Factor in the neighbors: Scalable and accurate collaborative filtering. *ACM Transactions on Knowledge Discovery from Data (TKDD), 4*(1).

Koren, Y., Bell, R., & Volinsky, C. (2009). Matrix factorization techniques for recommender systems. *Computer, 42*(8), 30–37.

Kozinets, R. V. (2001). Utopian Enterprise: Articulating the Meanings of Star Trek's Culture of Consumption. *Journal of Consumer Research, 28*(1), 67-88.

Krulwich, B. (1997). Lifestyle Finder Intelligent User Profiling Using Large-Scale Demographic Data. *AI Magazine, 18*(2), 37 - 45.

Kumar, R., Raghavan, P., Rajagopalan, S., & Tomkins, A. (1999). Trawling the Web for Emerging Cybercommunities. *Computer Networks*, 1481-1493.

Kuo, & Ying-Feng. (n.d.). A Study on Service Quality of Virtual Community Websites. *Total Quality Management, 14*(4), 461-473.

Kwon, J., & Kim, S. (2010). Friend Recommendation Method using Physical and Social Context. *International Journal of Computer Science and Network Security, 10*(11), 116-120.

Lawrence, S. (2000). Context in web search. *IEEE Data Engineering Bulletin, 23*(3), 25-32.

Lazar, J., & Preece, J. (1998). Classification Schema for Online Communities. *Proceedings of the 1998 of Association for Information Systems Americas Conference.* Baltimore,USA.

Lechner, U., & Hummel, J. (2002). Business Models and System Architectures of VirtualCommunities: From a Sociological Phenomenon to Peer-to-Peer Architectures. *International Journal of Electronic Commerce, 6*(3), 41.

Leckenby, J. D., & Li, H. (n.d.). Why We Need the Journal of Interactive Advertising. *Journal of Interactive Advertising*.

Lee, B. (1992). Colonialization and community: Implications for first nations development. *Community Development Journal, 27*(3), 211–219.

Li, L., Wang, D., Li, T., Knox, D., & Padmanabha. (2011). Scene : A scalable two-stage personalized news recommendation system. *In SIGIR '11 Proceedings of the 34th international ACM SIGIR conference on Research and development in Information Retrieval* (pp. 125–134). China: ACM.

Li, L., Wang, D., Li, T., Knox, D., & Padmanabhan, B. (2011). SCENE : A Scalable Two-Stage Personalized News Recommendation System. *Proceedings of the 34th international ACM SIGIR conference on Research and development in Information Retrieval.* Beijing, China.

Li, X., & Chang, S.-K. (2005). A Personalized E-Learning System Based on User Profile Constructed Using Information Fusion. *Citeseer*, 109–114.

Liang, H., Xu, Y., Li, Y., & Nayak, R. (2010). Personalized recommender system based on item taxonomy and folksonomy. *Proceedings of the 19th ACM international conference on Information and knowledge management*, (pp. 1641–1644).

Lieberman, H. (1997). Autonomous Interface Agents. *CHI '97: Proceedings of the ACM SIGCHI Conference on Human factors in computing systems.* Atlanta GA.

Lin, C., Xie, R., Li, L., Huang, Z., & Li, T. (2012). Premise: personalized news recommendation via implicit social experts. *In Proceedings of the 21st ACM international conference on Information and knowledge management* (pp. 1607–1611). Hawaii , USA: ACM.

Liu, J., Dolan, P., & Pedersen, E. (2010). Personalized news recommendation based on click behavior. *Proceedings of the 15th international conference on Intelligent user interfaces.* Hong Kong, China.

Lops, P., Gemmis, M. d., & Semeraro, G. (2011). *Content-based recommender systems: State of the art and trends.* Springer Science.

Lops, P., Gemmis, M. d., Semeraro, G., & Musto, C. (2009). *A Semantic Content-Based Recommender System Integrating Folksonomies for Personalized Access.* Web Personalization in Intelligent Environments,Springer.

Lorch, R. F. (1982). Priming and search processes in semantic memory: A test of tree models of spreading activation. *Journal of Verbal Learning and Verbal Behavior, 8*, 240-248.

Luke, S., Spector, L., & Rager, D. (1996). Ontology-based Knowledge Discovery on the World-Wide Web. *AAAI-96 Workshop on Internet-Based Information Systems.* Menlo Park, CA, USA.

Macchiette, B., & Roy, A. (1992). Affinity marketing; What is it and how does it work? *Journal of Services Marketing, 6*(3), 47-57.

MacIver, R. M. (1937). *Society, a textbook of sociology.* New York: Farrar & Rinehart, Inc.

Maes, P. (1994). Agents that reduce work and information overload. *Communications of the ACM, 37*(7), 33-41.

Mark, G. (1998). *Self Without Body: Textual Self-Representation in an Electronic Community*. Retrieved from http://131.193.153.231/issues/issue3_4/giese/index.html

Markoff, J. (2002). High-speed Wireless Internet Network is Planned. USA: New York Times.

Markus, U. (2002). *Characterizing the virtual community (5th edition). SAP Design Guild.* Retrieved from http://www.sapdesignguild.org/editions/edition5/communities

McLaughlin, M. R., & Herlocker, J. L. (information retrieval). A collaborative filtering algorithm and evaluation metric that accurately model the user experience. *Proceedings of the 27th annual international ACM SIGIR conference on Research and development in information retrieval,ACM.* New York, USA.

McMillan, D. W., & Chavis, D. M. (1986). Sense of Community: A Definition and Theory. *Journal of Community Psychology, 14*, 6-23.

Melville, P., & Sindhwani, V. (1997). Recommender Systems. *Proceedings of the ACM Magazine on Communications.* USA.

Melville, P., Mooney, R., & Nagarajan, R. (2002). Content-boosted collaborative filtering for improved recommendations. *Eighteenth national conference on Artificial intelligence.* Menlo Park, CA, USA.

Middleton, S. E., Shadbolt, N. R., & De Rour, D. C. (Jan. 2004). Ontological user profiling in recommender systems. *ACM Trans. Inf. Syst, 22*(1), 54–88.

Mirjalili, S., & Lewis, A. (2016). The Whale Optimization Algorithm. *ELSEVIER, 95*, 51-67.

Mitra, A., & Schwartz, R. L. (2001). From cyber space to cybernetic space: Rethinking the relationship between real and virtual spaces. *Journal of Computer Mediated Communication, 7*(1).

Montaner, M., Lpez, B., & Rosa, J. d. (2003). A taxonomy of recommender agents on the internet. *Artificial Intelligence Review*, 285–330.

Mowbray, M. (2001). Philosophically Based Limitations to Freedom of Speech in Virtual Communities. *Information Systems Frontier, 3*(1), 123-131.

Nadee, W., Li, Y., & Xu., Y. (2013). Acquiring user information needs for recommender systems. *Web Intelligence (WI) and Intelligent Agent Technologies (IAT), 2013 IEEE/WIC/ACM International Joint Conferences.* Atlanta, GA, USA.

Nanas, N., Roeck, A. D., & Vavalis, M. (2009). What happened to content-based information filtering? *Springer.* Cambridge, UK.

Nanas, N., Uren, V., & Roeck, A. d. (2003). Building and applying a concept hierarchy representation of a user profile. *Proceedings of the 26th annual international ACM SIGIR conference on Research and development in informaion retrieval*, (pp. 198–204).

Naruchitparames, J., Gunes, M., & Louis, S. (2011). Friend Recommendations in Social Networks using Genetic Algorithms and Network Topology. *IEEE Congress on Evolutionary Computation (CEC).* USA.

Netflix india U watch tv programmes online. (n.d.). (Netflix) Retrieved 03 12, 2018, from https://www.netflix.com/ in/

Noy, N. F., & McGuniness, D. L. (2001). *Ontology Development 101: A guide to creating your first ontology.* Standford University. California: Stanford Knowledge Systems Laboratory Technical Report KSL-01-05.

Online Presence Ontology. (n.d.). Retrieved from www.online-presence.net/ontology.php

Pace-Sigge, M. (2013). *Lexical Priming in Spoken English Usage.* UK: Palgrave Macmillan.

Papagelis, M., Plexousakis, D., & Kutsuras, T. (2005). Alleviating the sparsity problem of collaborative filtering using trust inferences. *Trust management,Springer*, 224-239.

Pathak, D., Matharia, S., & Murthy, C. (2013). Nova: Hybrid book recommendation engine. *Advance Computing Conference (IACC), 2013 IEEE 3rd International Conference.* UP, India.

Pazzani, M. J. (1999). A Framework for Collaborative, Content-Based and Demographic Filtering. *Artificial Intelligence Review, 13*, 393–408.

Pazzani, M. J., & Billsus, D. (2007). *Content-based recommendation systems.* The adaptive web,Springer.

Philip, W. (2009). *Major firms ban Twitter from the workplace as social networking sites cost businesses £1.38bn a year.* Retrieved from http://www.dailymail.co.uk/ news/article- 1223034/Major-firms-ban-Twitter-workplace-employees-tweets-costfirms

Pil´aszy, I., & Tikk, D. (2009). Recommending new movies: even a few ratings are more valuable than metadata. *Proceedings of the third ACM conference on Recommender systems.* New York.

Plant, R. (2004). Online communities. *Technology in Society, 26*(1), 51-65.

Porter. (2004). A Typology of Virtual Communities: A Multi-Disciplinary Foundation for Future Research. *Journal of Computer-Mediated Communication, 10*(1).

Preece, J. (2000). *Online communities: designing usability, supporting sociability.* Chichester: John Wiley & Sons, Ltd.

Preece, J. (2001). Sociability and usability in online communities: Determining and measuring success. *Behaviour & Information Technology, 20*(5), 347–356.

Preece, S. (1981). *A spreading activation model for Information Retrieval.* University of Illinois, Urbana-Champaign, USA: PhD thesis.

Quillian, M. R. (1966). *Semantic Memory.* USA: Carnegie Institute of Technology.

Rafaeli, S., & Sudweeks, F. (1997). Networked interactivity. *Journal of Computer Mediated Communication, 2*(4).

Razmerita, L., Angehrn, A., & Maedche, A. (2003). Ontology based user modeling for Knowledge Management Systems. *Proceedings of the User Modeling Conference,Springer Verlag.* Pittsburgh, USA.

Recommender systems. (2010). Encyclopedia of machine learning,Springer.

Reder, L. M., & Anderson, J. R. (1980). A partial solution of the paradox of interference: The role of integrating knowledge. *Cognitive Psychology, 12*, 447-472.

Rendle, S., Freudenthaler, C., Gantner, Z., & Thieme, L. S. (2009). Bpr: Bayesian personalized ranking from implicit feedback. *In Proceedings of the Twenty-Fifth Conference on Uncertainty in Artificial Intelligence.*

Resnick, P., Iacovou, N., Suchak, M., & Bergstrom, P. (1994). Grouplens: an open architecture for collaborative filtering of netnews. *In Proceedings of the 1994 ACM conference on Computer supported cooperative work* (pp. 175-186). NY: ACM.

Resnick, P., Iacovou, N., Suchak, M., Bergstrom, P., & Riedl, J. (1994). GroupLens: an open architecture for collaborative filtering of netnews. *CSCW '94: Proceedings of the 1994 ACM conference on Computer supported cooperative work*, 175-186.

Rheingold, H. (1993). *The Virtual Community: Homesteading on the Electronic Frontier.* Addison-Wesley.

Ricci, F., Rokach, L., & Shapira, B. (2011). *Recommender Systems Handbook, chapter Introduction to Recommender Systems Handbook.* Boston, MA: Springer US.

Ringe, S., Francis, N., & Palanawala Altaf, H. (2012). Ontology Based Web Crawler. *International Journal of Computer Applications in Engineering Sciences, 2*(3), 194-198.

Rinku, Raihan, M., & Aravind, A. (2016). Children’s mind: Cultural values in Chhota Bheem. *IUP Journal of English Studies, 11*(1), 28-31.

Roberts, H. (1979). *Community Development: Learning and Action.* Canada: University of Toronto Press.

Romm, C., Pliskin, N., & Clarke, R. (1997). Virtual Communities and Society; Toward an Integrative Three Phase Model. *International Journal of Information Management, 17*(4), 261-270.

Røpke, I. (1999). The dynamics of willingness to consume . *ELSEVIER , 28*(3), 399-420.

Rothaermel, F. T., & Sugiyama, S. (2001). Virtual internet communities and commercial success: Individual and community-level theory grounded in the atypical case of TimeZone.com. *Journal of Management, 27*(3), 297–312.

Sadasivam, G. S., & Praveen, E. M. (2014). Personalisation of News Recommendation using Genetic Algorithm. *3rd International Conference on Eco-friendly Computing and Communication Systems.* Mangalore, India.

Sarwar, B., Karypis, G., Konstan, J., & Riedl, J. (2001). Item-based collaborative filtering recommendation algorithms. *Proceedings of the 10th international conference on World Wide Web.* Hong Kong: Proceedings of the 10th international conference on World Wide Web.

Schafer, J. B., Frankowski, D., Herlocker, J., & Sen, S. (2007). *Collaborative filtering recommender systems.* The adaptive web,Springer.

Schein, A. I., Popescul, A., Ungar, L., & Pennock, D. M. (2002). Methods and metrics for cold-start recommendations. *Proceedings of the 25th annual international ACM SIGIR conference on Research and development in information retrieval.* Tampere, Finland.

Shah, D., Schmierbach, M., Hawkins, J., Espino, R., & Donavan, J. (2002). Nonrecursive Models of Internet Use and Community Engagement: Questioning whether Time Spent Online Erode Social Capital. *Journalism & Mass Communication, 79*(4), 964-987.

Shardanand, U., & Maes, P. (1995). Social information filtering: algorithms for automating "word of mouth". *CHI '95: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 210-217.

Sharifian, F., & Samani, R. (1997). Hierarchical Spreading of activation. *Proc. of the Conference on Language, Cognition and Interpretation*, 1-10.

Shvaiko, P., & Euzenat, J. (2013). Ontology Matching: State of the Art and Future Challenges. *IEEE Transactions on Knowledge and Data Engineering, 25*(10), 158-176.

Silva, N. B., Tsang, I.-R., Cavalcanti, G. D., & Tsang, I.-J. (2010). A graph-based friend recommendation system using Genetic Algorithm. *IEEE Congress on Evolutionary Computation.* Barcelona, Spain.

Singh, S., Shepherd, M., Duffy, J., & Watters, C. (2006). An Adaptive User Profile for Filtering News Based on a User Interest Hierarchy. *Proceedings of the American Society for Information Science and Technology.*

Sproull, L., & Faraj, S. (1997). *Atheism, sex and databases: the net as a social technology, in Culture of the Internet .* NJ: Lawrence Erlbaum Associates.

Stanoevska-Slabeva, K., & Schmid, B. (2001). A Typology of Online Communities and Community Supporting Platforms. *Proceedings of the 34th Hawaii International Conference on System Sciences.*

(2011). *State of the Media: Social Media Report Q3.* http://blog.nielsen.com/nielsenwire/social/, Nielsen.

Steinmueller, W. (2002). *Virtual communities and the new economy.* Inside the communication revolution: Evolving patterns of social and technical interaction. Oxford: Oxford University Press.

Su, X., & Khoshgoftaar, T. M. (2009). A survey of collaborative filtering techniques. *Advances in Artificial Intelligence, Volume 2009*, 19.

Tak´acs, G., Pil´aszy, I., N´emeth, B., & Tikk, D. (2008). Matrix factorization and neighbor based algorithms for the netflix prize problem. *Proceedings of the 2008 ACM Conference on Recommender Systems, RecSys '08.* New York, USA.

Tang, J., Yao, L., Zhang, D., & Zhang, J. (2010). A Combination Approach to Web User Profiling. *ACM Transactions on Knowledge Discovery from Data (TKDD), 5*(1), 1-39.

Tomlinson, M. (2003). Lifestyle and social class. *European Sociological Review, 19*(1), 97-111.

Tonnies, F. (1887). *Gemeinschaft and Gesellschaft.* Retrieved from Accessed Community - Wikipedia, the free encyclopedia.mht

Trajkova, J., & Gauch, S. (2004). Improving Ontology-based User Profiles. *Proceeding of RIAO 2004.* Avignon, France.

Tsatsaronis, G., Vazirgiannis, M., & Androutsopoul, I. (2007). Word Sense Disambiguation with Spreading Activation Networks Generated from Thesauri. *IJCAI, 7*, 1725-1730.

Turow, J., & Kavanaugh, A. L. (2003). *The Wired Homestead.* Cambridge, MA: MIT.

Typaldos, C. (2000). *The 12 Principles of Collaboration.* Retrieved from www.typaldos.com/12 principles.htm

Typaldos, C. (2001). The 12 Principles of Collaboration. USA: Mongoose Technology.

Vallet, D., Castells, P., Fernández, M., Mylonas, P., & Avrithis, Y. (2007). Personalized content retrieval in context using ontological knowledge. *IEEE Transactions on Circuits and Systems for Video Technology, 17*(3), 2007.

Vargas, S. (2014). Novelty and diversity enhancement and evaluation in recommender systems and information retrieval. *Proceedings of the 37th International ACM SIGIR Conference on Research; Development in Information Retrieval, SIGIR '14.* New York ,USA.

Virnoche, M., & Marx, G. T. (2007). “Only Connect”—E. M. Forster in an Age of Electronic Communication: Computer-Mediated Association and Community Networks. *Sociological Inquiry, 67*(1), 85-100.

Wang, Z., Liao, J., Cao, Q., Qi, H., & Wang, Z. (2015). Friendbook: A Semantic-Based Friend Recommendation System for Social Networks. *IEEE TRANSACTIONS ON MOBILE COMPUTING, 14*(3), 538-551.

Wellman, B., Salaff, J., Dimitrova, D., Graton, L., Gulia, M., & Haythornthwaite, C. (1996). Computer Networks as Social Networks: Collaborative Work, Telework, and Virtual Community. *Annual Review of Sociology, 22*(1), 213-238.

Weng, L.-T., Xu, Y., Li, Y., & Nayak, R. (2008). Exploiting Item Taxonomy for Solving Cold-Start Problem in Recommendation Making. *20th IEEE International Conference on Tools with Artificial Intelligence.* USA.

Whittaker, S., Terveen, L., Hill, W., & Cherny, L. (1998). The dynamics of mass interaction. *Proceedings of the ACM Conference On Computer Supported Cooperative Work.* Seattle, Washington, USA.

Wickizer, T., & Donald, P. (1995). *Community and Health.* New York: Oxford University Press.

Wilson, S. M., & Peterson, L. C. (2002). The anthropology of online communities. *Annual Review of Anthropology, 31*(1), 449–467.

Woerndl, W., Schueller, C., & Wojtech, R. (2007). *A hybrid recommender system for context-aware recommendations of mobile applications.* Istanbul, Turkey: Data Engineering Workshop, 2007 IEEE 23rd International Conference.

WORDNET. (n.d.). Retrieved from http://wordnet.princeton.edu/

Xu, B., Bu, J., Chen, C., & Cai, D. (2012). An exploration of improving collaborative recommender systems via user-item subgroups. *Proceedings of the 21st international conference on World Wide Web.* Lyon, France.

Xue, R., Zeng, J., Chen, Z., Ma, Y., Xi, W., Fan, W., & Yu, Y. (2004). Optimizing Web Search Using Web Click-through Data. *Proceeding of ACM Thirteenth Conference on Information and Knowledge Management (CIKM).* Washington D.C., U.S.A.

Yang, S. Y. (2010). Developing an ontology-supported information integration and recommendation system for scholars. *Expert Systems with Applications, 37*(10), 7065–7079.

Yeung, K. F., & Yang, Y. (2010). A Proactive Personalized Mobile News Recommendation System. *Developments in E-systems Engineering.* London, UK.

Yu, J., Liu, F., & Zhao, H. (2012). Building user profile based on concept and relation for web personalized services. *n International Conference on Innovation and Information Management.*

Yu, X., Pan, A., Tang, L.-A., Li, Z., & Han, J. (2011). Geo-Friends Recommendation in GPS-based Cyber-physical Social Network. *International Conference on Advances in Social Networks Analysis and Mining.* Kaohsiung, Taiwan.

Zack, M. H. (n.d.). Interactivity and communication mode choice in ongoing management groups. *Information Systems Research, 4*(3), 207–239.

Zhang, Y., & Koren, J. (2007). Efficient bayesian hierarchical user modeling for recommendation system. *Proceedings of the 30th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval.* New York, NY, USA.

Zhang, Y., & Koren, J. (2007). Efficient bayesian hierarchical user modeling for recommendation system. *Proceedings of the 30th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, SIGIR '07.* New York ,USA.

Zhang, Y., Ahmed, A., Josifovski, V., & Smola, A. (2014). Taxonomy discovery for personalized recommendation. *Proceedings of the 7th ACM International Conference on Web Search and Data Mining.* New York, USA.

Zhang, Z., Liu, Y., Ding, W., & Huang, W. (2015). A Friend Recommendation System Using Users' Information of Total Attributes. *International Conference on Data Science.* Sydney, Australia.

Ziegler, C.-N., Lausen, G., & Schmidt-Thieme, L. (2004). Taxonomy-driven computation of product recommendations. *Proceedings of the thirteenth ACM international conference on Information and knowledge management.* Washington DC, USA.

Ziegler, C.-N., McNee, S., Konstan, J., & Lausen, G. (2005). Improving recommendation lists through topic diversification. *Proceedings of the 14th ACM international conference on World Wide Web.* Chiba, Japan.

PUBLICATIONS

INTERNATIONAL JOURNALS

- Papneja, S., Sharma, K., Khilwani, N.," Friend Discovery Based on Users Interest", International Journal of Advanced Intelligence Paradigms, Inderscience, Accepted.

- Papneja, S., Sharma, K., Khilwani, N.," Context Aware Personalized Content Recommendation Using Ontology Based Spreading Activation", *International Journal of Information Technology*, Accepted.

- Papneja, S., Sharma, K., Khilwani, N., "*Movie Recommendation to Friends using Whale Optimization Algorithm",Bentham Science,* Accepted.

INTERNATIONAL CONFERENCES

- Papneja, S., Sharma, K., Khilwani, N.," Content Recommendation Based on Topic Modeling", International Conference on Computational Methods and Data Engineering, Accepted.