A Comparative Analysis of Recommender System For Social Networking Sites: A Case Study of Facebook, Twitter, Jeevansathi and Amazon in Punjab

[1]Ramanpreet Kaur, [2]Dr. Sawtantar Singh Khurmi
[1]Research Scholar at Desh Bhagat University, Mandi Gobindgarh, Punjab, India
[2]Head, Dept. of Computer Science, Bhai Maha Singh College of Engineering, Mukatsar, Punjab, India

Abstract
With the advent of computers and other means for mass digital storage, we started collecting and storing all sorts of data. Today, we have far more information than we can handle. Information retrieval is simply not enough anymore for decision-making. Confronted with huge collections of data, we have now created new needs to help people make better managerial choices. Recommender Systems are the tool and techniques which provide suggestions for the services and information of use. The systems are achieving widespread success in social networking sites nowadays. Recommender system applied in social networking sites are fully different from other kind of recommender systems. In Social network recommender system work on implicit and explicit data of the subscriber and provide the recommendation. Since the information recommendations are rational for human beings rather than products The current techniques that are mainly divided into five categories namely: Collaborative, Content based, Knowledge based, Demographic and Hybrid recommender techniques. This paper describes the anomalies of recommender technique by comparing the different tool of the recommender system.

Keywords
Recommender System, Social Network, KDD, Clustering, k-mean, Clope, SCClust, WSCClust and MOD.

I. Introduction

We are in an age often referred to as the information age. We have been collecting tremendous amounts of information as it leads to power and success. With the advent of computers and other means for mass digital storage, we started collecting and storing all sorts of data. Today, we have far more information than we can handle: from business transactions and scientific data, to satellite pictures, text reports and military intelligence. Information retrieval is simply not enough anymore for decision-making. Confronted with huge collections of data, we have now created new needs to help us make better managerial choices. These needs are automatic summarization of data, extraction of the "essence" of information stored, and the discovery of patterns in raw data. With the enormous amount of data stored in files, databases, and other repositories, it is increasingly important, to develop powerful means for analysis and interpretation of such data and for the extraction of interesting knowledge that could help in decision-making [1]. Data Mining, also popularly known as Knowledge Discovery in Databases (KDD), refers to the non trivial extraction of implicit, previously unknown and potentially useful information from data in databases. While data mining and KDD are frequently treated as synonyms, data mining is actually part of the knowledge discovery process [18]. Data mining originally was developed to act as an expert system to solve the problems. It is used in every field to solve the problems. Now data mining get introduced in social media. Social media is enjoying a great deal of success in these days. The user can visit social sites like Facebook, Twitter, MySpace, RSVP and eHarmony etc for different purposes. These social sites are designed for different purpose. Facebook, Twitter and MySpace like sites for communication and RSVP& eHarmony for dating purpose [4]. The social networks make the process of finding new friends on the network with the help of hobby, age and interest matching techniques. Social matching system can be defined as a recommender system that suggest people to people rather than items. Kaplan and Haenkein [15] define "Social Media as a group of internet–based applications that are built on the ideological and technological foundations of Web 2.0, and allow the creation and exchange of User Generated Content." Recommender system is a two way matching process. The task of recommender system is to help the user to find their prefect matches. The task of recommending people is more complicated as compared to recommending products. The Recommender System can be divided into five Categories: - Collaborative, Content-Based, Demographic, Knowledge-Based and Hybrid Filtering. Collaborative filtering is one of the most promising and widely used recommender techniques which recommend items to subscribers by determining the item that similar subscriber have positively rated or selected [5]. Collaborative Recommender Systems suffer from two main problems. The first problem is the lack of user feedback or user rating in some type of web sites. The second problem is referred to as a Cold start problem, which is identified when a new user joins [4]. Content-based Recommender System relies upon item features to produce the recommendations. Content-Based Recommender System use two approaches: the Classification approach, and second is Nearest-Neighbor approach. The Classification approach, items are classified, based on their characteristics. In contrast, the nearest neighbor approach uses the information of previously rated items for a specific user to recommender new items to him/her. The Content-Based Recommender System suffers from low accuracy problem [8].

The Knowledge-based recommender systems are rarely used in social matching system due to their vast complexity. In this system implicit data are traced and used to improve the recommendations along with the explicit data [19]. Demographic recommender systems rely on users' characteristics to find the relationship between them and others, and to find what they like. One major weakness in demographic system is that the system use static user profile, which does not get updated when a user interests changes. So, the social matching system have some problem like ColdStart, Sparsity, Cyberstalking, Social profiling and third party disclosure, identify the membership between two persons. Some of these are reduced with the help of clustering techniques of data mining [4].

Jain et al. [12], Clustering is the unsupervised classification of patterns into groups. The clustering problem has been addressed in many contexts by researchers in many disciplines; this reflects its broad appeal and usefulness as one of the steps in exploratory data

analysis. However, clustering is a difficult problem combinatorial, and differences in assumptions and contexts in different communities have made the transfer of useful generic concepts and methodologies slow to occur. The k-means algorithm is effective in producing clusters for many practical applications. But the computational complexity of the original k-means algorithm is very high, especially for large data sets. Moreover, this algorithm results in different types of clusters depending on the random choice of initial centroids. K-means algorithm generates accurate and efficient results. But result is totally depended on the selection of initial centroids. Methods for refining the computation of initial centroids is worth investigating. As a consequence, using the standard "recommender system" suffer from some problems. The major problem is high complexity of the social matching system by using normal recommender system[4]. The "cold-start" problem describes situations in which a recommender is unable to make meaningful recommendations due to an lack of ratings initially. This problem can significantly degrade CF performance. It can occur in three scenarios: New User: - When a user first registers with a CF service, they have no ratings on record. New Item:- When a new item is added to a CF system, it has no ratings, so it will not be recommender. New Community: - The biggest cold-start problem is bootstrapping a new community [5]. The aim of this research is to identify the problems associated with recommendation system and provide suggestions to remove the same. Some of the problems have been identified by the different researchers, but no clear solution have been provided or identified. The main aim of this research is to identify all the problems and generate the solution to remove that problem.

II. Related Work

The main objective of literature review is to provide background knowledge of the related research areas. The networking concept gave birth to a business model for inter-connecting people for social causes. This idea was practically implemented as Social Networks and Social Matching System. An automated system was developed known as Recommender System for matching personal attributes of the subscribers. The Recommender System suffers from several problems like privacy problem, data accessibility problem, Cold Start, Sparsity, Cyberstalking, Social profiling and third party disclosures etc. Terveen and Mcdonald [22] raised some general questions on social network system. Their research specifies that data mining techniques can be used to discover the social relationship within the social network. It is possible to mine patterns of communications, find out the association rules from the users who get together, and divide users into clusters based on their similarities. Chen Lin [7] conducts a comprehensive analysis of an online dating network to solve social match making problem. The Researcher worked on five methods to improve matching process. The study shows that all these methods do not remove the cold start problem from the social matching system. The study suggested developing a better method to solve the cold start problem from social matching system. Greg Linden et al. [10] focus on the uses of recommendation algorithm for marketing in retail industry. The researchers developed a collaborative filtering algorithm which scales the massive data sets and produces high-quality recommendations in real time. Huang Z. et al. [11] worked on the sparsity problem by applying an associative retrieval framework and related spreading activation algorithm to explore transitive association amongst consumers through past transactions and feedbacks. They conducted an experimental study to evaluate the effectiveness of this approach using data set from an online bookstore and experimented with three spreading activation algorithms and compare these algorithms with collaborative filtering approaches that do not consider transitive associations. The study showed the over-activation effect of the spreading activation approach. The incorporating transitive associations with past transactional data that is not sparse may "dilute" the data used to infer preferences and lead to degradation in recommendation performance. The study suggested in its future work that a lots of work need to be done to improve the quality of spreading activation algorithm. Jilin Chen et al. [14] studied how effective different algorithms are in recommending people as potential friends, and what characteristics are there in term of recommending known versus unknown people. The researchers conducted this study to find out how could a recommender system effectively increase the number of friends and what would be the overall impact of such recommender system on the site. They developed a people recommender system on the concept of beehive. Two separate experiments were conducted, a personalized survey and a controlled filed study. It was found that the relationship-based algorithms are better in finding known contacts where as content similarity algorithms stronger at discovering new friends. The study suggested future research on developing better recommendation algorithm social networking sites. Another scope left is to investigate whether people recommendations can help bootstrap new comers, addressing adoption issues of social networking sites. Alsaleh et al. [4] concluded that the recommendations on the basis of implicit data in social matching system need to be explored. The study recommended that social matching system suffers from two important issues: computational complexity and the matching accuracy. To reduce these problems the research uses a new hybrid social matching system that uses implicit and explicit data to improve the recommendation process. Yang et al. [23] worked on the problem of categorical data clustering, especially for transactional data characterized by high dimensionality and large volume. To reduce this problem the researcher develop a novel algorithm CLOPE, which is fast, simple and based on the intuitive idea of increasing the height-to-width ratio of cluster histogram. Rui Xu [21] conducts a survey on clustering algorithms for data sets appearing in statistics, computer science and machine learning. The study illustrates that there is no clustering algorithm that can be universally used to solve all the problems. Alan Mislove et al. [2] reiterated that much work needs to be done on symmetric link and level of clustering. The study shows that the social networks are structurally different from other networks. After the analysis of online social networks, four popular online social networks have been suggested to gather data. Jain A.K [12] in his study on benchmark and clustering methods suggested that there is need to benchmark data to test and evaluate clustering methods. The study specify need to achieve a tighter integration between clustering algorithms. Researcher worked on k-means algorithm that is very simple and most popular partition algorithm. This algorithm requires three user-specified parameters : number of clusters, cluster initialization, and distance matric.

III. Proposed Work

In the world of Information Technology we have some sites which recommend person to person to establish the relation. The task of recommendation is performed with the help of Recommender System. The recommender system collects the data from the user and generates the results according to users' choice. Recommender system is a two way matching process. The task of recommender

system is to help the user to find their prefect matches. The task of recommending people is more complicated as compared to recommending products. The Recommender System can be divided into five Categories: - Collaborative, Content-Based, Demographic, Knowledge-Based and Hybrid Filtering [9].

Table 1: Comparison of Recommender Techniques

Techniques	Purpose	Upside	Downside
Collaborative	Recommendation based on User rating and User history	No need of manual maintenance of data.	Suffer from Sparsity cold start and scalability problem.
Content-Based	Recommendation based on item characteristics	Small groups are involve to achieve recommendation accuracy.	Error prone process because of manual data maintenance.
Knowledge-Based	Recommendation based on Implicit and Explicit data.	Use implicit data to predict items that the users may like, along with explicit data.	This technique is rarely used because of vast complexity. .
Demographic	Recommendation relies upon the users' characteristics.	Find the relationship between user and the item, they like	User profile don't updated when the users interest change.
Hybrid	Combination of two or more recommendation techniques		Suffer from Sparsity cold start and scalability problem and Less accuracy.

All the Online social networking sites that provide the user to user interaction use the recommender system. Facebook have experienced exponential growth in membership in recent years. This social networking site use demographic recommender system to provide the recommendation to their users [4]. These networks offer attractive means for interaction and communication, but also raise privacy and security concerns. Members use these sites for a number of purposes. The root motivation is communication and maintaining relationships. Popular activities include updating others on activities and whereabouts, sharing photos and archiving events, getting updates on activities by friends, displaying a large social network, presenting an idealized personal, sending messages privately, and posting public testimonials. These networks offer attractive means for interaction and communication, but also raise privacy and security concerns [6]. The misuse of personally identifiable information obtained online can raise many privacy concerns such as identity theft or even discrimination [17]. Twitter is a popular micro-blogging and social network service that allows people to share messages of 140 characters in length. As of September 2009, Twitter had over 50 million unique users [18]. While Twitter allows people to share information among friends or "followers", the default privacy setting on Twitter is that all messages are public, that is, anyone who signs up for Twitter may see them. In addition, all public tweets may be posted to a public timeline website which showcases the twenty most recent tweets. Profiles on Twitter are relatively short compared to Facebook; therefore the bulk of the information about a person is communicated through their Twitter messages or tweets. Incidental information such as when and where people may be can also have privacy implications. Time and location may constitute a second tier of personally identifiable information, which while seemingly mundane and minor can raise potential safety concerns when publically broadcasted and shared. Privacy concerns can arise due to the digital storage of personal information. Alterman [3] suggests that there are three kinds of concern associated with the storage of data. First is that someone "will legitimately gain access to information about you and utilize it to locate and harass or harm you in some manner" [3]. The second kind of privacy concern is that information you gave up for one reason or purpose will be accessed or used for purposes that you could not have thought about or did not approve. The third kind of concern that Alterman suggests is that your data will be illegitimately or even illegally accessed and used which will then put your at risk for embarrassment or worse. While the third is unlikely to occur through the public timeline of Twitter, the first two privacy concerns, which Alterman suggests arise from the storage of data, are quite relevant on Twitter in that followers on Twitter could potentially use tweets to harm or harass you [17]. Jeevansathi is one of the leading matrimonial websites in India, which was initiated in September 2004 keeping in view to serve the North and Western Indian communities. The website has more than 8.5 Lac registered members. Jeevansathi also has 14 match point across the India. It is a part of Bombay Stock Exchange and the growth of the company can be tracked in the annual report of Info Edge, shared on its website. Its searchable directory includes profiles of prospective brides and grooms. One can register and search by religion, community, caste and sub-caste, geographical area, occupation, and mother native language [14].

Amazon is an American Electronic Commerce company with headquarters in Seattle, Washington in 1994. It is developed in C++ and Java Languages. It provides online shopping, web hosting and content distribution services to their subscribers. Amazon use item –to- item recommender algorithm to provide the recommendation. One problem that's endemic to large-scale recommendation systems is scalability [10]. All the recommender tool suffer from data scalability and efficiency problems. Recommender System used clustering algorithm to improve the matching process. Clustering algorithms can be classified into four categories based on the employed techniques namely:- Partitional, hierarchical, density-based and grid based. Partitional clustering algorithms randomly choose k objects from the data set. Each object is considered as a single cluster. Then, the partitioned clustering algorithms iteratively reallocate other data points to their closest cluster unit a stopping criterion is met. This requires that users have at least some domain knowledge and another one is the difficulty of indentify clusters with large variations in size. Clustering people in social networks is the fusion of three main areas: large scale clustering, constrained clustering and semi supervised clustering. Each of these areas have been well researched and many algorithms have been presented for each area. To design an efficient algorithm for clustering people in social networks require a consideration of all three together. The major problem with the existing social matching system, and one that affect the accuracy of matching, is the compatibility of recommendation. Many social matching system consider the similarity between user x and user y, and then produce the recommendation without checking that y is also compatible with x. The new clustering algorithm Match opposite datasets (MOD) ensures the user preferences and their own information for each recommendation. The new clustering algorithm works on the opposite clusters and generates the results of highly ranked data. The basic idea behind this algorithm is similarity measurement.

Input: Male and Female datasets
Output: List of ranked recommendation
Algorithm:
Begin:

1. Male User Cluster and their attributes (M1,M2,M3..Mx) into a number of groups;
2. Female User Cluster and their attributes(F1, F2,F3..Fn) into b number of groups;
3. For each cluster Ci do
   - Find the opposite gender from cluster Cj that has highest successful communication with Ci;
   - End.
4. For each user um belong to Cj do
   - Calculate the compatibility with all users in Cj where Cj is best matched cluster with Ci using User Sim $(U_m,U_f) = \sum^n_i=1 Sim(O_{mi}, P_{fi})$
   - End
5. End

This algorithm provides the recommendation on the basis of user preferences and their information attributes. This algorithm is implemented on online social matching system in which opposite gender users were recommended to each other. On the basis of their place (dataset1), hobbies (dataset2), qualification (dataset3) and age (dataset4) data sets are used to evaluate the new method. Each data set is grouped into male and female datasets. The empirical analysis demonstrates that the new system improves the accuracy of the recommendations, reduces the matching complexity, and overcomes the cold start problem. The new MOD algorithm helps to reduce the computational complexity of matching the users of social networks, because it emphasize on comparing every male with all female users. It limits the comparisons to the users within an assigned cluster.

V. Success Rate

The success rate, an external quality matrix, was used in this research to evaluate the social matching system based on the k-means, Clop, SCClust, WSCClust and MOD clustering algorithms. The success data rates of matching users in all four data sets based on these different clustering algorithms. Although the k-mean algorithm achieved the best results for both intra-similarity and inter-similarity, it gives the lowest success rate. Because it depends on all the clustering values. Further, the WSCClust perform about 20% better than Clope because it is used constraints and 14% better than SCClust because it used weighted constraints (Slah, 2012). The success rate increased in MOD algorithm because the constraints have less support and confidence as the constraints were ranked decreasing order.

Table 2: Success Rate

Clustering Method	Data Set1	Data Set2	Data Set3	Data Set4	Average
KM	0.38	0.41	0.32	0.38	0.3725
CLOPE	0.45	0.45	0.4	0.45	0.4375
SSClust	0.54	0.54	0.51	0.54	0.5325
WSCClust	0.62	0.6	0.59	0.62	0.6075
MOD	0.77	0.72	0.70	0.7	0.7275

Success Rate= Successful recommendations ∩ All recommendations All recommendation

This research is focused on finding the problems in recommender system. To achieve this objective four recommender tool are selected like: Facebook(Demographic Recommender Sytsem), Twitter(Content- Based Recommender system), Jeevansathi(Collaborative recommender system) and Amazon(item-based recommender system) and questionnaires are distributed in two district and collect responses from the internet subscribers. Research query specifies the Sex(Male, Female), age(16-24,25-30,31-35,35-45), place of residence (Urban, Rural, Semi-Urban), Education(Matric, SSC, Graduate, Postgraduate), Profession(Salaried, Business, Unemployed, Student), hobbies(Reading, Music & Movies, Shopping, Mobile-Computer Savvy) and Marital status of subscribers. The averages of internet subscribers on these parameters represented through these charts. When the subscribers visiting these social networking sites they receive unwanted notifications.The research shows that 66% subscribers receive the unwanted notification from Social Netwroking Sites, 20 subscibers sometime receive unwamted notification and sometime not. 13% respondents say they does not receive the unwanted notifications from SNS. There are 1 percent respodent never receive notifications. The study collect the data from the subscriber about the behaviour of users when they de-activate their account and how they feel when they are not subscribing the SNS. In Barnala district 53% respodents feeling bored when they are not subscribing the SNS and 22% respodents miss not being online badly and 25% users feel relaxed. In Sangrur district 50% respodents feeling bored when they are not subscribing the SNS and 25% respodents miss not being online badly and 25% users feel relaxed. 40% subscribers feeling very bad after deactivating their account and 25% users feeling bad. There are 178 respondents out of 400(90.22%) are not satisfied by the privacy setting provided by the social netwoking sites and 128 respondets are satisfied by the privacy settings of social netwoking sites.

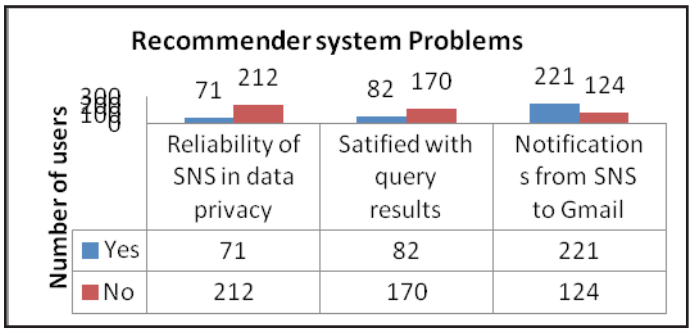

Fig. 1: Compare the result of Recommender System Problems

The 212 subscribers from 400 think that social newteoking sites are not reliable in keeping data accuracy.Besides it, 71 respodents think that social netwoking sites are reliable to provide tha data accuracy. However, 170 respodents are not satified with the querry results from social networking sites. Only 82 subscribers are statisfied with the querry result of social networking sites. There are 221 users who have seen the notification from SNS to their gmail account. The 124 users does not receive any notification from SNS to their gmail account. This impact on the database compatability.

V. Conclusion

SNS also makes the process of finding similar people a very

complicated and expensive task. Thus social networking sites need a recommender system that uses all type of available information regarding users' and their interest to provide the high accuracy recommendations. The purpose of recommender system is to provide the recommendation to their user on the basis of implicit and explicit data. The objective of recommender system is to provide the accurate results to their users in less time with high accuracy. Importantly, this research is focused on recommender system techniques and tools to identify the anomalies that arise at the time of recommendations. However, number of comparisons takes place to achieve the objective. The research is done to phases: one is empirical and another one is experimental. The results from both phases show that SNS users are not aware of third party involvement and tracking of data. These systems are based on actual user behavior i.e. objective reality. This is the biggest advantage - watching people in their natural environment and making design decisions directly on the results. The biggest issue facing recommender systems is that they need a lot of data to effectively make recommendations. Recommender system firstly needs item data (from a catalog or other form), then it must capture and analyze user data (behavioral events), and then the magic algorithm does its work. The more item and user data a recommender system has to work with, the stronger the chances of getting good recommendations. Recommender system is not capable to provide the recommendation when the user interest changes. Each person wants to stay in touch with their friends and other ones which are stranger for them. For this they join the SNS and do the chat with friends or other members of the SNS. They upload their picture on these SNS and time to time update the status, which effects the storage capacity at the server side and decrease the data accessibility rate. Some users create their ID on the SNS to take the benefits of SNS. However having the member of SNS user create the id but never use it. This activity arise data storage anomalies. The recommender system should find out those users who are not using their accounts for more than 6 months or one year and should deactivate their accounts and the user must follow new creation of account. To increase the storage there should be a specified limit on the number of time the user updates their status in a week. These results show that interaction of trust and privacy concerns in social networking is not yet removed. The result of the study encourages further research on data storage and privacy concerns.

Refrences

[1] Akanmu Semiu Ayobami and Shamsuddeen Rabi'u, "Knowledge Discovery in Database: A knowledge Management Strategic Approach", proceeding in Knowledge Management International Conference (KMICe) 2012, Johor Bahru, Malaysia, 4 – 6 July 2012.

[2] Alan Mislove, M. Macron, K.P.Gummadi, "Measurement and Analysis of Online Social Networks", IMC'07, ACM 978-1-59593,-908-1/07/0010, 2007, California, USA.

[3] Alterman, A. (2003). "A piece of yourself": Ethical issues in biometric identification. Ethics and Information Technology, 5(3), 139-150.

[4] Alsaleh, Slah, Nayak, Richi, Xu,Yue , and Chen, Lin, "Improving Matching Process in Social Network Using Implicit and Explicit User Information". In: Proceeding of the Asia-Pacific Web Conference lecture notes in Computer Science, 2011.

[5] Ben J. Schafer, Dan Frankowski, Jon Herlocker, and ShiladSen, "Collaborative Filtering Recommender System", in Proceeding of The Adaptive Web, LNCS 4321, pp. 291-324, 2007.

[6] Catherine Dwyer, Starr Roxanne Hilte, Katia Passerini, "Trust and Privacy Concern within Social Networking Sites: A comparison of Facebook and Myspace," proceeding of the Thirteenth Americas conference on information system, Keystone, Colorado August 09-12 2007.

[7] Chen Lin , "A Social Matching System: Using Implicit and Explicit Information for Personalized Recommendation in Online Dating Services", 2013.

[8] Christian Desosriers and George Karypis, "A comprehesive survey of neighbourhood-based recommendation methods", pp 107-144, DOI-10.1007/978-0-387-85820-3-4, ISBN-978-0-381-85819-7, Springer US, 2011.

[9] Dietmar Jannach, Markus Zanker, Alexander Felfering, Gerhard Friedrch, 352pages, Cambridge University Press, 30 September 2010, ISBN-10: 0521493366, ISBN-13: 0521493369.

[10] Greg Linden, Brent Smith, and Jeremy York, "Amazon. Com Recommendation: item-item Collaborative Filtering," Industry Report, IEEE Internet Computing, January-Feburary,2003.

[11] http://computer.org/internet.

[12] Huang. Z, H. Chain, and Daniel Zeng, "Applying Associative Retrieval Techniques to Alleviate the Sparsity Problem in Collaborative Filtering ", in ACM Transactions on Information System, Vol. 22, No.1, 2004, pp116-142.

[13] Jain A.K, "Data clustering: 50 years beyond k-means. Pattern Recognition Letters", 31(8):651-666, 2010.

[14] Jain A.K, M.N.Murty, and P.J.Flynn, "Data Clustering: A Review", ACM Computing Surveys, Vol.31, No.3, pp 264-323, USA, 1999.

[15] Jiban K. Pal, "Social Networks Enabling Matrimonial Information Services in India," International Journal of Library and Information Science Vol. 2(4), pp. 54-64, May 2010 Available online http://www.academicjournals.org/ijlis

[16] Jilin Chen, Werner Geyer, Casey Dugan, Michael Muller, and Ldo Guy, "Make New Friends, but Keep the Old"-Recommending People on Social Networking Sites. ACM 978-1-60558-246-7/08/04, 2009.

[17] Kaplan A.M, and M. Haenkein, "Users of the wold, unite! The Challenges and opportunities of Social Media, Business Horizon, 53(1),59-68, 2010.

[18] Lenhart, A., & Madden, M. Teens, "privacy & online social networks. Washington DC:Pew Internet and American Life Project," pp 10-29, 2007.

[19] Moore, R. J. (2009, October 5). Twitter data analysis: An investor's perspective. Retrieved November 1, 2009, from http://www.techcrunch.com/2009/10/05/twitter-data-analysisan-investors-perspective/.

[20] Neelamadhab Padhy, Dr. Pragnyaban Mishra, and Rasmita Panigrahi, "The Survey of Data Mining Applications and Feature Scope", in International Journal of Computer Science, Engineering and Information Technology (IJCSEIT), Vol.2, No. 3, 2012.

[21] Robin Burke, Alexander Felfernig, and Mehmet H. Göker, "Recommender System: An Overview", Association for the Advancement of Artificial Intelligence (www.aaai.org). 2011.

[22] RuiXu, "Survey of Clustering Algorithm", in IEEE Transactions on neural networks, Vol. 16, No.3, May-2005, Pages 645-

678, ISSN-1045-9227, DOI-10.1109/TNN.2005.845141, INSPEC Accession Number- 8422347.

[23] Terveen, L. And McDonald, D. W. " Social Matching : A Framework and Research Agenda " , in ACM Transactions on Computer-Human interaction, Vol. 12, No. 3, 2005, pp. 401-434.

[24] Yiling Yang, Xudong Guan, Jinyuan You, "CLOPE: A Fast and Effective Clustering Algorithm for Transactional Data", in SIGKDD, 2002, Edmonton, Alberta, Canada, ACM 1-58113-567-X/02/0007.