Study of Recommendation System for Web Portals

Lokesh Sharma
Jaypee University of Informational
Technology
Mathura, India

Ashok Kumar Agrawal
National Informatics Center, DeitY
Ministry of Communication &
Informational technology

ABSTRACT
Recommendation system (RS) is one of the most advanced approach which is popular commercially and in Research community. Many of the Web portals are using Recommender system to increase their customers and providing them better recommendation for purchasing of products. It learns from the customer's behavior of purchasing, rating and commenting, then deciding the score by help of Recommender system. In this paper, introducing about Recommendation system and its various types with their corresponding technologies that are currently used in E-commerce web portals. Later explaining some of the well known portals using Recommender system and comparison in techniques. Paper conclude with the applications of recommendation system and how they are increasing customer's to E-commerce.

General Terms
Collaborative Based Recommendation System, Content Based Recommendation system

Keywords
Group Recommendation System, Hierarchical Clustering, Video Recommendation System, Intelligent Recommendation System, Graph Recommendation system, Hotel Recommendation System, Cold start Context based Recommendation, Mobile Recommender System.

1. INTRODUCTION
Burke, 2002 defined Recommendation system as "any system that produces individualized recommendations as output or has the effect of guiding the user in a personalized way to interesting or useful objects in a large space of possible options" [1]. Recommendation system is a dynamic process of learning from Customers behavior of traditional information by which machine learns and do various texts mining analysis to recommend product's rating or score. It is found in many applications to help their customer for judging of products. This system produces a list on the basis of scores or ratings that is done by Recommendation system algorithms by which customer can decide appropriate items.

Recommender systems are now popular in both Research as well as commercially community. RS is now found in many modern applications that expose the user to a huge collection of items. These systems help users to decide an appropriate item, and ease the task of finding items in the collection. All these system are typically categorized as RS, even though they provide diverse services.

1.1 Types of Recommendation system
Collaborative based ;Recommender system accept collaborative filtering as one of the widely executed technique in web portals and in e commerce.[2,3,4,5]In collaborative system probability of accuracy will always increased whenever huge amount of data is there otherwise the system will not give the correct recommendation. There are so many parameters on which Collaborative filtering works such as purchasing habits, preferences, rating etc. Based on inter-user comparisons it generates Recommendations. Collaborative recommendation identifies users whose choices are similar to those of the available user and recommend items they have liked.

Content based; Recommender system that learns profile of the user's interest based on the feature present in objects the user has rated. Content-based recommendation systems used in various fields such as television programs, news articles, recommending web pages, restaurants and items for sale, apparently the details of various systems differ. Content-based recommendation systems share common means for describing the items that are recommended.

Knowledge based; Recommendation System which is completely reasoning oriented it have quantitative decision support tools[6]which allow customers to recommend better support for their judgments. Such recommendation systems are used to remove Statistical anomalies. Based on information, the system pursues a knowledge based approach to generating a Recommendation, by reasoning about what products meet the user requirements. It avoids some of the drawbacks: it does not have a ramp up problem since recommendation do not depend on rating, its judgment are independent of individual tastes.

2. APPLICATION OF RECOMMENDATION SYSTEM.
2.1 Facebook (Group Recommendation System)
Today Face book is well known for its methodology and its research in Recommender system. We all have used the applications of RS and by that it is very easy to search friends which having common interests. As there is huge diversity in cyberspace and it brings lot of uncertainty while joining groups, so FB propose a solution that is well known as Group Recommendation System . FB solve this problem by Hierarchical Clustering and Decision Tree techniques by which it select the most specified group according to the taste. Previously so many researchers had introspected over this topic and they conclude it as following.

Hoff et al. [7] done mathematically find relationship between different groups by the help of probability approach. Backstrom et al.[8] had find how to grow the network by separate adding of users and finally made them in a group with having common parameters. Lada A Adamic et al.[9] developed small world effect, clustering and the strength of weak ties. Their work was over online portals such as Club Nexus, at Stranford University. They draw Correlations between individual parameters and with others attributes. Due to the extensive efforts they have find maximum comparisons between the attributes to draw new trends and make interesting friends in Group Recommendation System. Gary William Flake et al.[10] observed Recommendation system by Graph theory that is suited for all types of graphs and link structures.
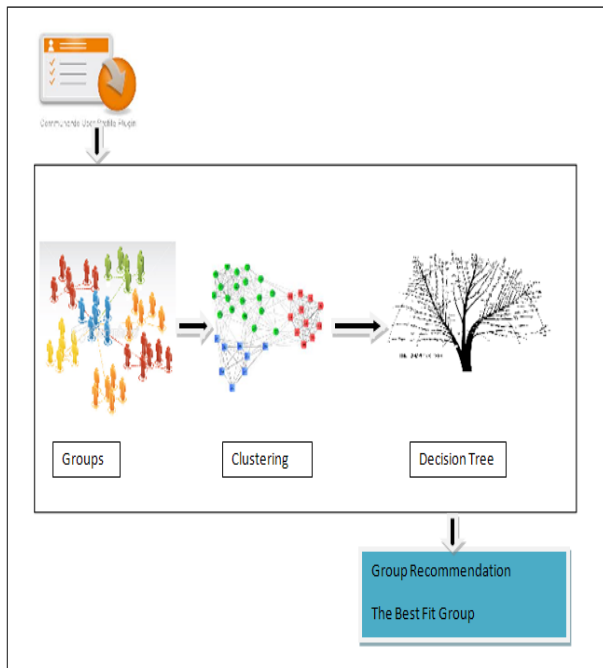
Fig1 Group recommendation system

Enkh Amgalan et al.[11] categorized into three parts 1)Similarity inference 2)Clustering Coefficient 3)Decision tree.

In similarity inference, similarity is drawn between different members it is done by Euclidian distance formulae. Similarity in nodes calculated by Hierarchical clustering analysis in multidimensional space. Calculated normalize data compute distance matrix and then generate hierarchical cluster tree from unweighted pair group method using Arithmetic Averages.

Fig2 Screenshot from face book

Clustering coefficient (C) is innermost part of the group and it should be reduced from the calculated value to reduce noise impact. Decision Tree finds the most suitable groups for a user based on binary recursive partitioning. It also split the data by some splitting algorithm to attain highest homogeneity on each side.

2.2 YouTube (Video Recommendation System)

YouTube is well known site for online video community, it is world's most populous site and it also recommends personalized sets of videos. Search and goal oriented browsing [12] is the biggest challenge for YouTube and it is excellently driven by them. YouTube delivers personalized group of videos according to the user's recent click which shows its interest.

Renjie Zhou*y* et al.[13] perform measurement study on recommendation video system by the video view sources and they have found that there is strong affinity between view count of a video and average view count of its top referrer videos. It conclude that YouTube rather than showing the most populous videos it shows according to user's interest which is the only reason for YouTube Diversification.

James Davidson et al.[12] has divided YouTube Recommendation system into three parts 1)Data Collection 2)Recommendation generation 3) Recommendation Serving.

In Data Collection, whenever user upload any video it also input details regarding to the video such as title, description etc. Video viewer users do the following activities such as rating, comments, likes etc. which is having different impact while the video weight is calculated. In recommendation generation by the techniques co-visitation counts, video visits are counting within 24 hours. By all these sequences of videos top N candidate videos are ranked and then only they will be displayed at user profile. Again while ranking so many parameters are checked about videos such as video quality, specify and diversification which tell about users choice of videos, and then only they will be served to the users.
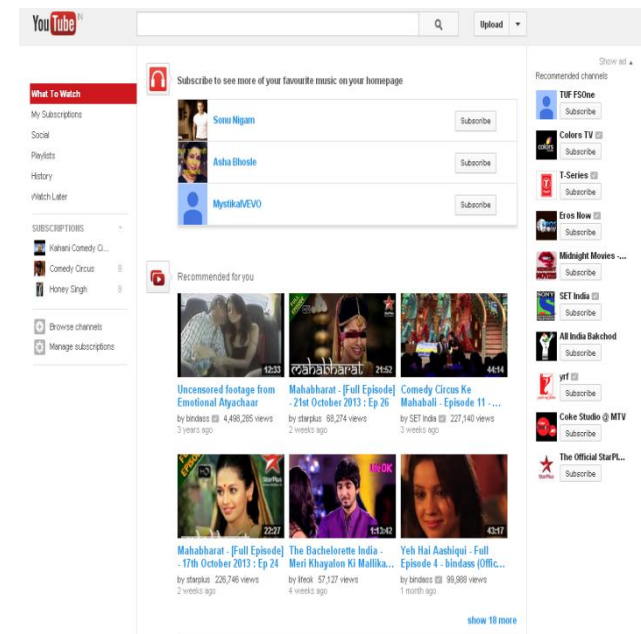

Fig 3 Screenshot of Youtube Video recommendation system with having diversification

2.3 Cognizant (Intelligent Recommendation System)

Intelligent Recommendation system[14] increased awareness among scientists exploring areas, technologies on which colleagues are working and experts views. Apache Mahout is

a Conventional open source framework whose primary goal is to scale Machine learning algorithms. Apache is the superior way to include as it implements Collaborative filtering, Clustering and Classification Machine learning Techniques.

Collaborative filtering generate recommendation in four ways, user based: users with similar characteristics, item based: Similar items, Slope one : Based on previous user rated items, Model based : Compare profile with active users.IRS adopted user based Collaborative filtering technique since the client recommend scientist based on multiple preferences.
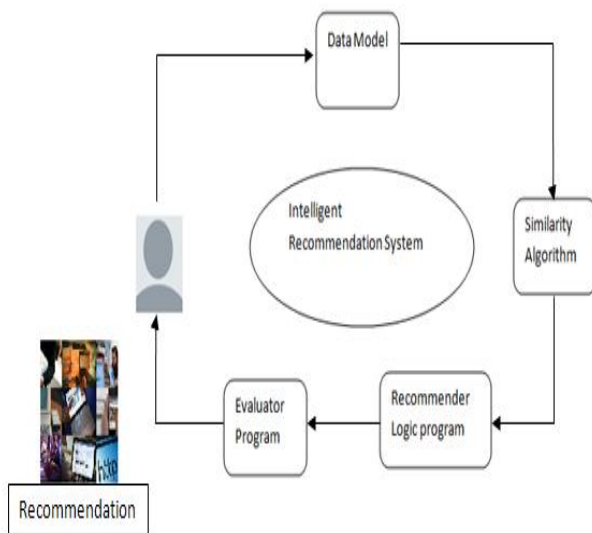

Fig 4 Flow of Intelligent Recommendation system

TanimotoCoefficientSimilarity takes all types of input data. It provides similarity on the basis of rating each items. Demographic properties plays a vital role in establishing similarity among users and it is best strategy for clustering of users.

Classification support two types of classifier approaches: Naïve Bayes Classifier and Complementary Naïve Bayes Classifier. In first stage it analyze the data and searches common features and in second predict the category of a new item.

2.4 Chinese Digital libraries (Graph Recommendation system)

Graph based recommendation system is demographic based approach combination of Content based and Collaborative recommendation[15,16,17].Such Recommendation system have application in today's online digital library stores even they work well for multilingual environment.[18,19]Amazon.com is one of the best example of collaborative filtering and manual recommendation. In Recommendation system three types of information are usually available factual, demographic and transactional information.

A two layer graph approach is proposed by Zan huang et al. for graph based recommendation system[20] exclusively in the context of digital libraries. In this approach there are two stages of computation first stage, draw vectors between customers and book information and calculating similarity by the help of similarity function[21]. At the second stage, a model is proposed in which two layers are there upper layer represent Books and down layer shows customers. In Book

layer there are nodes which represent books and shows interconnectivity with other books based on similarity function same with customer layer. There are one more links between Books and customers which shows the transactional history. This model adapts content based Collaborative and hybrid approach.
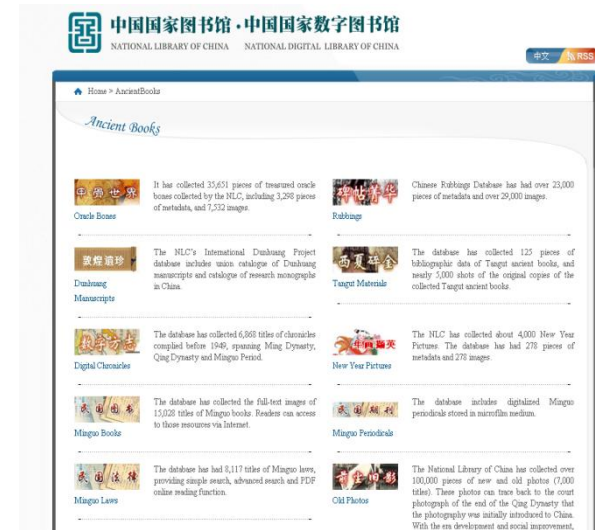

Fig 5 Chinese Digital Library

2.5 TripAdvisor.com and Venre.com (Hotel Recommendation System)

Cold start Context based Recommendation[22] is unique type of Recommendation system in which the system needs to recommend a product to a new user that has no past information and very few or no ratings[23,24,25] that is why it is called to Cold Start. This system extract information from reviews and a very novel type of text mining to read emotions of users.
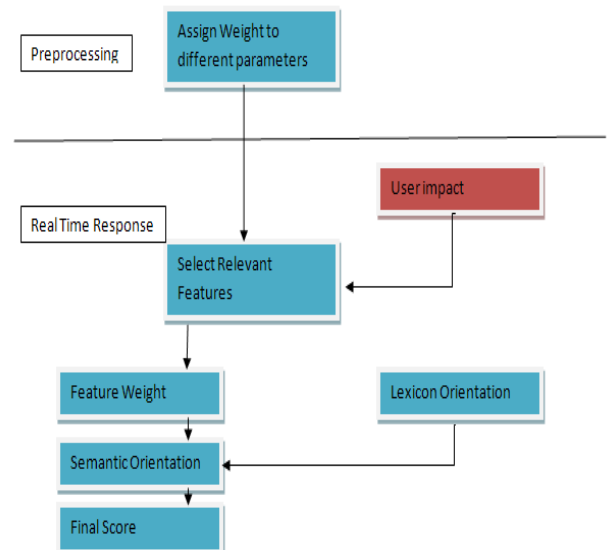

Fig 6 Flow of Hotel recommendation system

Today internet has overtaken word of mouth as primary medium for choosing destination[26]:63% of consumers plan travel by searching the Internet. Context based Hotel Recommendation System will minimize users effort and likely to recommend product of users interest. It generalize three types of Context information first is purpose second

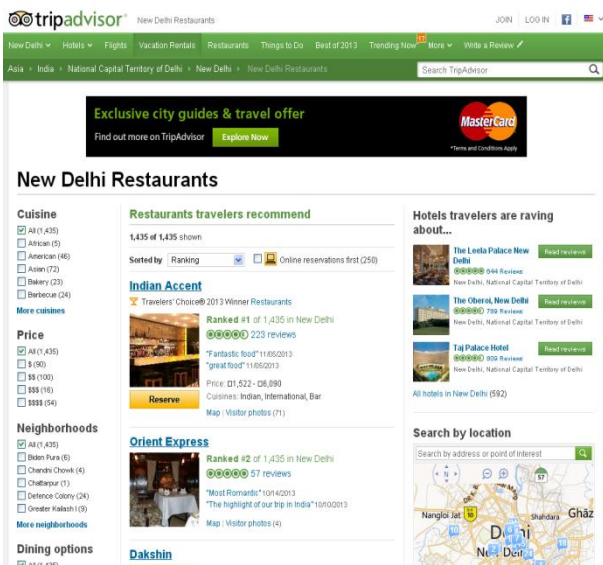nationality and third is preferences for various standards of hotels.

Fig 7 Screenshot of TripAdvisor.com

Researchers obtain data from Hotel Recommending sites such as reviews, rating and general information. In pre-processing stage all the texts are mined into groups according to their traits and in next stage Real time response feature weight is allotted to texts, then on the basis of semantic orientation give final scores to Hotels[27] by which customers can choose according to their demands.

2.6   Mobile Recommender System

Today's Mobile phones usage are not limited up to traditional manufactures designs, they have unique property of location awareness and ubiquity which make cell phones to be a in hand computers with exposed to global world. Mobile Recommendation system is knowledge based structure having capability to predicted Case Based Reasoning(CBR)[28] which effect today's generation a lot. It is characterized by three fundamental dimensions: user mobility, device portability and wireless connectivity[29].As wireless communication is getting pace a new type of business is emerged in market called to be M-commerce or M-business[30].To cultivate business by this technology it need devices which works over sensor and radio frequency efficiently such as mobile phones, personal digital assistant and there variants.

While developing recommender system for mobile system for mobile phones there were challenges of small screen devices, user performances, limited inputs, access results and have shorter browser sessions which make recommendation systems to be problematic on mobile devices. To overcome with these problems the immediate need is to work over base mobile designs and sooner QWERTY keyboards and starfield displays were part of cell phones.

Mobile recommender system works efficiently for tourist guide finding, relevant city attraction like museums, art galleries etc and also guide how to reach your destination point by transportation services. It navigate the user from their current geographical location using GPS to their destination having optimal route computation in road networks using shortest route between two points.

Recently researchers Tumas and Ricci, 2009 developed a personalized solution uses knowledge based approach. Burke, 2007 defined a new standard that can guide citizens and city guests to their destinations by having possible transport means such as bus , train ,taxi etc. PECITAS(Personalized City Transport Advisory System) generate multiple route and additional route constraints between source and destination[31,32].

2.7   Recommendation based on user clustering

Recommendation system provides suggestions for online stores such as books, magazines, movies, hotels and other items. In this items are not intended for personal usage but for bulk of users having common interests. It comes into the category of Collaborative filtering approach which offer users to the similar liked items. [33]

Algorithm clustered the users with similar preferences and grouped them in a cluster. In this set of items and users are interacting with some recommendation system and expresses their preferences for a rated item denoted by *preference* and the items which are unrated denoted by *relevance.*

Personnel Recommendation, Different Researchers estimate relevance by different ways. It is following Collaborative filtering approach and produce relevance score for unrated items. By the help of *similarity* function it calculate the similar behavior and *support* function calculate the percentage of friends that have expressed the preference to the same item. The *personal value* is defined by the addition of product of worth, relevance and worth.

Group Recommendation, In this the preferences of single users predict the preferences for the group as a whole. There are three different strategies regarding the aggregation method (i)*least misery design*, the minimum value score for the group.(ii)*fair design,* having the average value score of group.(iii)*most optimistic design*, it capture the most satisfied members of the group.

3. COMPARISON IN VARIOUS RECOMMENDATION SYSTEMS

Refer to Table 1

4. CONCLUSION

Recommender System is a most advanced mass customization for web portals that increase E-commerce. It is increasingly adapted in future, as globalization increases value to the customers for their Business. Recommending previously was a very complex way but as internet is making whole world small such techniques are highly effective for users in judging products according to their tastes. In this survey paper, authors provide an intensive study of different well known web portals that are using Recommendation system and technologies associated with them to enhance their clientele. It also discussed the classification of different Recommendation system and their application in web portals.

Table

Methodologies used for Recommendation system	Techniques that are implemented	Challenges	Goals	Accuracy	Application
1.Face book Similarity Inference		Difficulty for users to select a right group to join	User friendliness and flexible in cyberspace	It overcome the challenge by provide suitable Group recommendation having diversified nature.	Group Recommendation system
	Hierarchical clustering analysis and Euclidean distance to infer similarity				
Clustering Coefficient	Differences in group are noticed by Noise Coefficient.				
Decision Tree	Based on Binary Recursive Partitioning Based on best split of data with maximum homogeneity				
2.Youtube Data Collection	Content Data, User Activity data Association rule mining or co-visitation counts	Videos having very poor metadata. User interactions are relatively short and noisy. YouTube have a short life cycle.	Direct navigation. Search and goal oriented Browse.	Diversification is achieved	Video Recommendation System.
Recommendation generation	Seed Set: Union of liked videos. Candidate set by taking limited transitive closure				
Recommendation Serving	Video quality, User specificity and diversification				
3. Cognizant Collaborative	Generate Recommendation based on user , item, slope, model Collaborative filtering	Online businesses are very hard to increase their customer bases. Increase information awareness. Bolstering sales	Recommendation system to be applied in Scientific research Domain. Providing Client with flexibility of using pre-existed data.	Provide flexibility to client in pre-existing information and enable researchers to use information more efficiently	Intelligent Recommendation system
Clustering	GenericUserSimilarity, TanimotoCoefficientSimilarity, PearsonCorrelationSimilarity, SpearsonCorrelationSimilarity, EuclideanDistanceSimilarity				
Classification	Naïve Bayes Classifier and Complementary Naïve Bayes Classifier				
4. Digital Libraries Chinese online stores Inclusion of Content based approach with Collaborative filtering approach.	Correlation between book to book, user to user and book to user. First stage, vectors from factual customers and book information.. Second stage, Interconnectivity between Customers and books.	Transactions to be implemented by links in vector diagram Implementation of graph. Weighting scheme to be implemented.	User to user, books to books and user to books correlation. Content based, Collaboration and hybrid approach are applied.	Based on usage history books are recommended in multilingual digital libraries.	Graph based recommendation system for multilingual digital libraries.
5. TripAdvisor.com and Venere.com Preprocessing	Intent: business trip, single traveler, family, group, couple. Nationality Standards: location, service, food, room, price value, quality.	Review text as it is the main data. Weighted algorithm for text mining. Vocabulary for hotel aspects	Context based search Recommendation. Vector aspect model with metadata. Utilize Demographic information	On Average 20% more satisfied with hotel recommendation	Cold start Context based Hotel Recommendation system.
Real time Response	Context based search to overcome Cold Start problem Devise Weighted text mining algorithm that leverages common traits found per context group.				

6. REFERENCES
[1] Interactive Critiquing for Catalog Navigation in E-Commerce R. Burke. *Artificial Intelligence Review 18(3--4):245--267 (2002)*

[2] Burke, R.D.; Hammond, K.J.; Yound, B.C., "The FindMe approach to assisted browsing," *IEEE Expert*, vol.12, no.4, pp.32,40, Jul/Aug 1997 [3]Goldberg, D., Nichols, D., Oki, B. M., Terry, D. 1992.Using collaborative filtering to weave an

information tapestry. Communicationosf the A CM,3 5(12) 61-70.Maes, P., Guttman, R. H., and Moukas, A. G. 1999. Agents that buy and sell. Communication of the ACM4, 2(3), 81-91.

[4]Resnick, P., Iacovou, N., Suchak, M., Bergstrom, P., and Riedl, J. GroupLensa: n open architecture for collaborative filtering of netnews. In CSCW9 4: Proceedings of the conference on Computer supported cooperative work, 175-186. New York: ACM Press.

[5] Resnick, P., Varian, H.R.: Recommender Systems. Communications of the ACM 40(3), 56–58 (1997)

[6] Bhargava, H.K.; Sridhar, S.; Herrick, C., "Beyond spreadsheets: tools for building decision support systems," *Computer* , vol.32, no.3, pp.31,39, Mar 1999

[7] Hoff, P.D., Raftery, A.E., Handcock, M.S.: Latent space approaches to social network analysis. Journal of the American Statistical Association, vol. 97(460), pp.1090-1098 (2002)

[8] Backstrom, L., Huttenlocher, D.P., Kleinberg, J.M., Lan, X.: Group formation in large social networks:Membership, growth, and evolution. In KDD, pp. 44-54 (2006)

[9] *L. A. Adamic*, *O. Buyukkokten*, and *E. Adar*. A *social network caught* in the *web*. *First Monday*, *8*(6), June *2003*.

[10] Flake, G.W., Lawrence, S., Giles, C.L., Coetzee, F.: Self-organization and identification of web communities. IEEE Computer, vol. 35(3), pp. 66-71 (2002)

[11] Book Section D 2008 978-3-540-88874-1 B On the Move to Meaningful Internet Systems: OTM 2008 Workshops 5333 Lecture Notes in Computer Science E Meersman, Robert E Tari, Zahir E Herrero, Pilar R 10.1007/978-3-540-88875-8_41 T Group Recommendation System for Facebook

[12] Davidson, James, et al. "The YouTube video recommendation system." *Proceedings of the fourth ACM conference on Recommender systems*. ACM, 2010.

[13] Zhou, Renjie, Samamon Khemmarat, and Lixin Gao. "The impact of YouTube recommendation system on video views." *Proceedings of the 10th ACM SIGCOMM conference on Internet measurement*. ACM, 2010.

[14] Anup Prakash Warade, Vignesh Murali Natarajan, Siddharth Sharad Chandak: cognizant 20-20 insights

[15Ahmad Wasfi, Ahmad M. "Collecting user access patterns for building user profiles and collaborative filtering." *Proceedings of the 4th international conference on Intelligent user interfaces*. ACM, 1998.

[16] Balabanović, Marko, and Yoav Shoham. "Fab: content-based, collaborative recommendation." *Communications of the ACM* 40.3 (1997): 66-72.

[17] Claypool, Mark, et al. "Combining content-based and collaborative filters in an online newspaper." *Proceedings of ACM SIGIR workshop on recommender systems*. Vol. 60. 1999.

[18] Resnick, Paul, and Hal R. Varian. "Recommender systems." *Communications of the ACM* 40.3 (1997): 56-58.

[19] Herlocker, Jonathan L., et al. "Evaluating collaborative filtering recommender systems." *ACM Transactions on Information Systems (TOIS)* 22.1 (2004): 5-53.

[20] Huang, Zan, et al. "A graph-based recommender system for digital library." *Proceedings of the 2nd ACM/IEEE-CS joint conference on Digital libraries*. ACM, 2002.

[21] Condliff, Michelle Keim, et al. "Bayesian mixed-effects models for recommender systems." *Proc. ACM SIGIR*. Vol. 99. 1999.

[22] Levi, Asher, et al. "Finding a needle in a haystack of reviews: cold start context-based hotel recommender system." *Proceedings of the sixth ACM conference on Recommender systems*. ACM, 2012.

[23] Adomavicius, Gediminas, and Alexander Tuzhilin. "Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions." *Knowledge and Data Engineering, IEEE Transactions on* 17.6 (2005): 734-749.

[24] Rashid, Al Mamunur, et al. "Getting to know you: learning new user preferences in recommender systems." *Proceedings of the 7th international conference on Intelligent user interfaces*. ACM, 2002.

[25] Schein, Andrew I., et al. "Methods and metrics for cold-start recommendations." *Proceedings of the 25th annual international ACM SIGIR conference on Research and development in information retrieval*. ACM, 2002.

[26] N. M. T. Watch, "Online Travel Market," April 2011. Available: http://www.newmediatrendwatch.com/worldoverview/91-online-travel-market

[27] Levi, Asher, et al. "Finding a needle in a haystack of reviews: cold start context-based hotel recommender system." *Proceedings of the sixth ACM conference on Recommender systems*. ACM, 2012.

[28] Bridge, Derek, et al. "Case-based recommender systems." *The Knowledge Engineering Review* 20.03 (2005): 315-320.

[29] Steiniger, Stefan, Moritz Neun, and Alistair Edwardes. "Foundations of location based services." *Lecture Notes on LBS* 1 (2006).

[30] Turban E, Lee JK, King D, McKay J, Marshall P (2008) Electronic Commerce – A Managerial Perspective. Prentice Hall.

[31] [Tumas and Ricci, 2009] *Personalized Mobile City Transport Advisory System*. In *Information* and *Communication Technologies* in *Tourism 2009*, *Pages*: *173--184*, *Springer*

[32] [Burke, 2007] *R. Burke*. The *adaptive web*. chapter *Hybrid web recommender systems*, *pages 377–408*. *Springer*-Verlag, *Berlin*, *Heidelberg*, *2007*.

[33] Ntoutsi, Irene, et al. "grecs: A group recommendation system based on user clustering." *Database Systems for Advanced Applications*. Springer Berlin Heidelberg, 2012.