Group Recommender Model for Boosting and Optimizing Customer Purchases

Saravanan M and Prasad G
Ericsson R & D, Chennai
India
m.saravanan@ericsson.com
prasad.garigipati@ericsson.com

Jagadeesan.M, Revathy Raman, Rekha.S
Department Of Information Technology
Meenakshi Sundararajan Engineering College
Chennai, India
mjagadeesan90@gmail.com, sheetalrevathy@gmail.com
rekha.sampathkumar@gmail.com

Abstract - **Group recommender systems generate a set of recommendations that will satisfy a group of customers with potentially competing purchase interests. This paper proposes a research and operational model which effectively enhances *Group Recommender Model* to boost the customer purchases. For this purpose, it uses the communication and collaboration of two major sources namely *Mobile Money Operator* and Outlet. MMO proactively monitors the spending pattern of the customers who make purchases using their mobile money. Outlet performs customer segmentation based on *RFM (Recency, Frequency and Monetary)* score after which a *Recursive Cluster Elimination* is performed that eliminates customers within the targeted segment. *Recursive Frequent Itemset Mining* and *Recursive Market Basket Analysis* are performed for the rest of customers in the targeted segment. From the obtained results, the product preferences of the remaining customers in the segment are identified based on which offers are formulated and recommended for the entire segment. It is then communicated to the MMO that intimates these offers to the potential customers among the segment. This model results in boosting customer purchases, expanding customer base and effects in the profitability of the combined source.**

Keywords: Customers; Mobile Money Operator (MMO); Outlet; Offers; Recursive Frequent Itemset Mining (RFIM); Recursive Clustering Elimination (RCE); Recursive Market Basket Analysis (RMBA)

I. INTRODUCTION

The customer is the end goal of businesses since it is the customer who makes purchase and creates demand that helps in the growth of the business. So, customer satisfaction has been regarded as the crucial activity in the field of Marketing. Different customer groups have different consumption behavior. As the toll of high spenders and frequent shoppers has grown like topsy, the liabilities to meet the interests of these customers have increased rapidly. To explore the purchase interests of the customer, a Recommender System [1] that attempts to analyze a customer's purchase history is employed so that offers for those specific products can be recommended to customers. When this recommender system is applied for a group of customers with similar purchase interest then this becomes a Group Recommender System. The important sources that connect customers with the system are described below.

Mobile Money is a *mobile financial service* provided by the MMO that enables their subscribers to convert their liquid cash into e-money in order to pay bills and accounts, top up mobile credit and receive money across the country's telecom networks [2]. The MMO stores all the transaction detail records of customers making purchases through mobile money over a period of time. The outlet is considered a *place of business for retailing goods.* These outlets store a wide range of datasets that contain the customer's transaction and purchase details over a period of time. These datasets are used to evaluate and explore the offer recommendations to boost the customer purchases. In the existing system, customers need to accustom to offers available for products that may not interest them but compromise on their choices just for the sake of saving their shopping expense.

Our Group Recommender System evolves from a Personalized Recommender System that targets the customer group rather than an individual. This paper builds a system that is not only able to extract consumer behavioral features from their consumptive information, but also to analyze, operate and evaluate these extracted features with enhanced techniques such as Recursive Cluster Elimination, Recursive Frequent Itemset Mining and Recursive Market Basket Analysis. In this kind of an approach, the scope of product purchase will be limited since the items focused for offers and their MBA rely on these items.

Generally, offers vary for each individual. When these offers are intimated individually, the cost for formulating offers and the time to perform the process increases making this approach a failure. Our contributions to the proposed work which enhances the traditional system are as follows:

- We define a group recommender model as shown in Fig 1 for customers in the outlet that recommends offers to the group on the basis of the

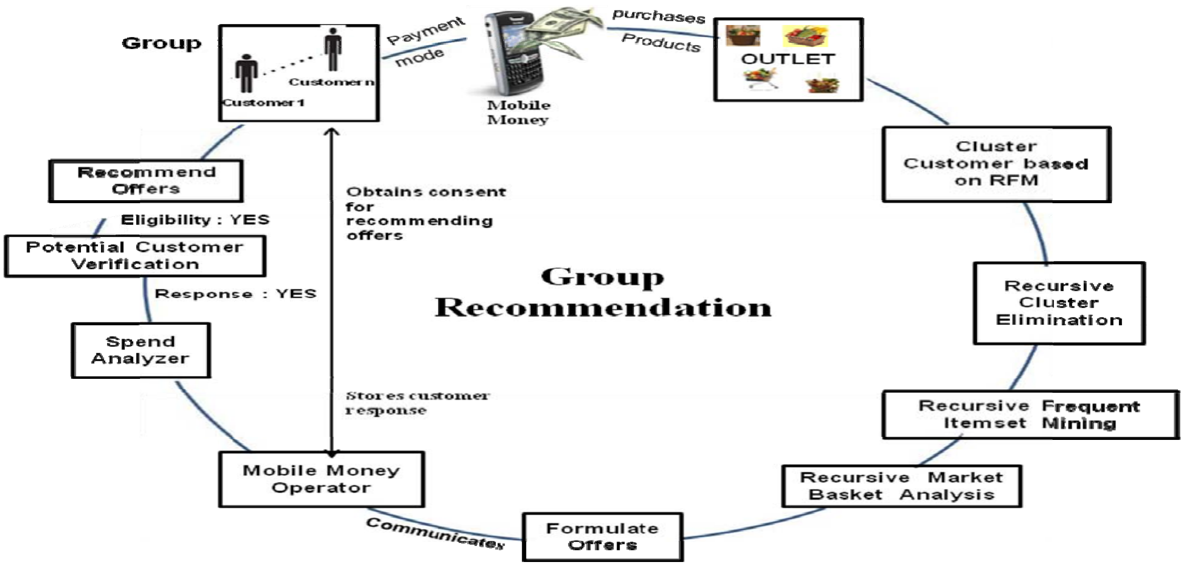

Fig 1: Group Recommender Model for Boosting and Optimizing Customer Purchases

purchase interests of the most potential customers in that group.

- We establish a link between two different sources namely MMO and Outlet that connects customers with the enticing offers for their product preferences.

- Frequent Itemset Mining and Association Rule Mining for customers within the targeted cluster is performed using a platform called HPCC (High Performance Computing Cluster) [3] which serves as a massive parallel-processing computing platform.

- We perform K-Means Clustering using Weka [4] and Market Basket Analysis using HPCC recursively to come out with an enhanced and efficient result that recommends offers not only for their product preferences but also for the items correlated with those products. This enhances the purchase behavior that was prevalent in Individual Recommender System and produces better results.

This paper is organized as follows: the following section presents the overview of the existing work in this system. Later, we describe about the varied techniques adopted in this paper. Subsequently, we elucidate the experimental analysis of the dataset used in our model. Finally, we conclude with the implications of our proposed research model.

II. RELATED WORK

The major part of the research on group recommendation investigated the core algorithms used for generating the group recommendations. Different strategies are available and two main approaches have been proposed [5]. The first consists of creating a joint user profile for all the users in the group and then performing a recommendation for this artificial user represented by the group profile [6] [7]. This would provide a recommendation for the group that is based on all user profiles and in some way represent group interests mediated in the group profile. The second approach aggregates the recommendations for each individual member into a single group recommendation list [8].

One of the most commonly-used and successfully-deployed recommendation approaches is collaborative filtering. In the field of collaborative filtering, two types of methods are widely studied: neighborhood-based approaches and model-based approaches. Neighborhood-based methods mainly focus on finding the similar users [9] or items [10] for recommendations. Model-based methods provide item recommendations by first developing a model of user ratings [11].

In our work we are concerned with recommendations to the entire group based on items purchased by potential members of the group. This does not incorporate the above mentioned approaches to promote recommendations. Here, the product preferences of the most potential customers are recommended for the entire group. The most potential customers (most loyal customers) have a good record of purchase history. When their items are recommended, there will be offers for myriad products and the scope of purchase can be improved significantly.

III. BACKGROUND

A. Group Recommender Model

Recommendation is a particular form of information filtering, which exploits past behaviors and user similarities to generate a list of information items that is personally tailored to an end-user's preferences [12]. Recommender systems have traditionally recommended items to individual users, but there has recently been a proliferation of recommenders that address their recommendations to groups of users. Group recommenders are built to help groups of people decide a common activity or item. The recommendation for the group will in general be based on information about preferences of individual group members [13]. Therefore, some type of aggregation method is required, by which information about individual preferences is combined in such a way that the system can assess the suitability of particular items for a group as a whole [5].

B. Recursive Cluster Elimination

Customer groups are formed according to customer behavior. RFM is behavioral model which considers past customer purchases [14] to prospect the future customer behaviors. The parameters Recency (R) defines the latest purchase made by the customer, Frequency (F) defines the number of purchases made within a certain period and Monetary (M) defines the money spent during a certain period. This paper involves the "User Definable Band" method to calculate RFM score that adds up the values of R, F and M values. Using this method, we can define specific bands for R, F, M that can be set based on certain criteria. Here the bands of R, F, M are measured in a scale of 1 to 5.This may vary from user to user. For example, *if (last visit>=0 day and last visit<51 days)*, then *R can be scaled as 5*, *if(number of transactions>=15)*, then *F can be scaled as 5*, *if(Transaction Amount>=15000),* then *M can be scaled as 5*.Similarly, the values 4 to 1for the parameters can be calculated. Now, a Partitioning Clustering technique called *K-Means Clustering* [15] is adopted to segment the customers in the outlet. The results of the clustering are obtained using a WEKA [4]. The K-Means clustering is a centroid based approach that accurately performs the clustering. Consider Fig 2 in which the *(1) customer purchases* in the outlet. *(2) Using the stored purchase details, RFM score is calculated. (3) Customers (here n=15) in the outlet are clustered based on RFM value into three segments namely "High, Medium and Low"*. This is considered as the first cluster step. Here, the centroids are chosen randomly based on RFM scores. It is then recalculated to get a modified centroid using which the scores nearest to the modified centroid is clustered in one group until no score switches clusters. After the first cluster step, *(4) a particular group for which recommendation has to be made is targeted* and *(5) Recursive Cluster Elimination* [16] *is performed* as shown in Section IV. In this case, the high cluster is being targeted. The RCE output holds only three most preferred customers as depicted in Fig 2. *Here, by using RCE we are only eliminating the customers with low R, F scores within the high cluster rather than eliminating the whole of the low cluster. (6) The output of RCE is used as the input of RFIM* – C4, C5, C6 customers are eliminated and C1, C2, C3 are most active within the targeted cluster.

C. Recursive Frequent Itemset Mining

The Frequent Itemset Mining is used to find frequent itemset present in a database. Apriori Algorithm [17] is designed to operate on databases containing transactions (For example, collections of items bought by customers). The input for this algorithm is derived from results of the FIM. Here, products which occur frequently are identified

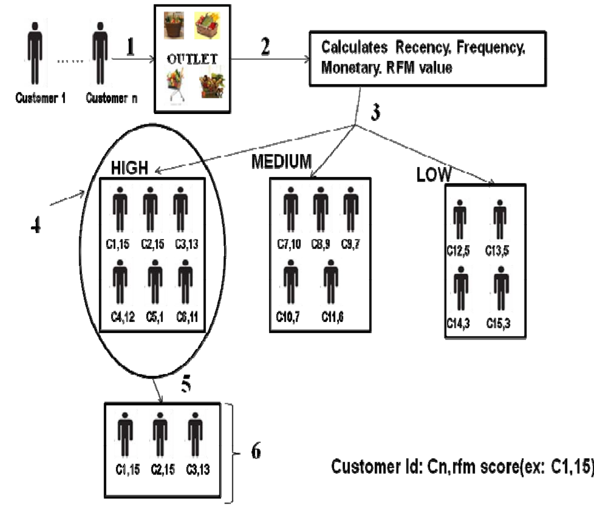

Fig 2: Recursive Cluster Elimination

and various scans are performed for it. This scan is nothing but the iterative search for products that occur frequently. This can be determined with a certain support count. In traditional FIM, the support count of an itemset X is defined as the number of transactions that contain X [18]. Products whose support count is below a certain threshold are pruned. Here, the frequent itemset scan is performed for products purchased by customers in the preferred group and this is done recursively to find the most preferred products by them. Frequent item set mining leads to the discovery of associations and correlations among items in large transactional or relational datasets present in the outlet.

D. Recursive Market Basket Analysis

Association Rule Mining represents an unsupervised learning method that attempts to capture associations between groups of items. MBA has also been referred to in the literature as Association Rule Mining or Affinity analysis[19]. Apriori is a seminal algorithm for mining frequent itemsets and identifying associations for varied customer purchases. In this model, the customers are clustered in the outlet using RFM score as discussed in sub-section B of section III. The output of frequent itemset mining is taken up as input for the MBA. Then, the Market Basket Analysis [20] made for an individual is merged to perform a recursive MBA which represents the group recommendation using which offers can be correlated and intimated to the customers. Here, the offers are formulated on the basis of the results obtained from the RMBA which are then intimated to the customer's mobile through MMO.

IV. PROPOSED SYSTEM

This section consists of establishing a relationship between Outlet, MMO and Customer to design the group recommender model. Sub-section A deals with the architecture of the proposed system and Sub-section B deals with the algorithm of proposed system by incorporating the K-Means, RFIM, RMBA techniques in our system.

A. Basic Architecture

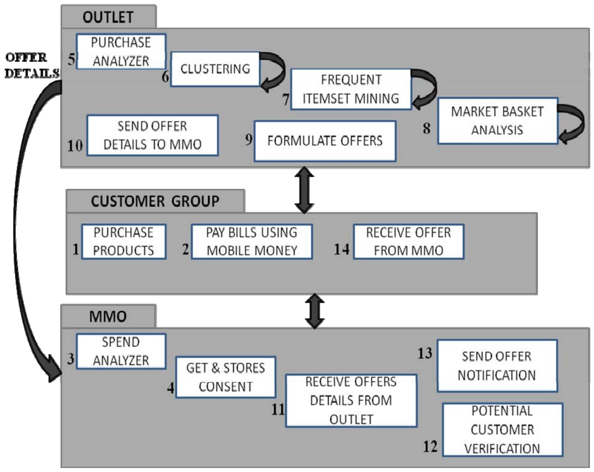

Fig 3: Overall Architecture of Proposed System

The above architecture consists of varied components represented in small rectangular boxes that are employed to build our group recommender system. Customers are the initiators of the model. They *make purchase* ***(1)*** in some outlet and make *payment using their mobile money* ***(2)***. The transaction details are logged in the MMO and purchase details are recorded in the outlet for future references. With those details as the base, the MMO performs the *Spend Analyzer* ***(3)*** to identify the domain in which a subscriber is interested in and it *obtains the customer's willingness for the service and stores it for future reference* ***(4)***. The outlet also performs a *purchase analyzer* ***(5)*** that examines the customer's purchase details and determines their distinct purchase interests. It *clusters customer* ***(6)*** and performs a *frequent itemset mining recursively* ***(7)*** to identify the product preferences of the customer. It also performs *market basket analysis recursively* ***(8)*** to find the product affinity and to *formulate offers* ***(9)*** for the group. Finally, the *outlet communicates the offer details of the group to the MMO* ***(10)***. The MMO performs *potential customer verification* **(*12)*** and *notifies offers to customer group* (***13)***.

Algorithm of Proposed System

The algorithm for Recursive cluster elimination*(1)* has been designed to identify the most potential customers whose purchase records are used as a tool to recommend offers for the entire group. In this, the cluster step includes the customer segmentation in the outlet using K-Means clustering based on RFM scores. The results are run on Weka in which the customers are segmented into three clusters. We cluster them into three categories as small number of segments can target a large number of customers. This saves cost, time and also ensures a greater probability of customers availing the offers. The next step targets a particular group from the three groups (high, low, medium). Using an incrementer i, the R, F, M values of each customer in that group targeted is checked for elimination. The threshold values for R and F are compared with the R and F values of each customer in that group.

1. Algorithm for Recursive Clustering

C_i=customer in the outlet
n=initial number of customers in targeted cluster.
t1=Threshold value for Recency(R)
t2=Threshold value for Frequency(f)
i = incrementer
1. Cluster the customers C into k clusters $c_1, c_2 \ldots c_k$ using K-means based on RFM scores (RFM score=R+F+M values)(Cluster Step).
2. Target a cluster from c1 to ck
3. Set threshold t_1 and t_2
for(i=1;i<=n;i++)
{
if($R_i > t_1$)and($F_i > t_2$) do
4. Retain the customer (C_i) in the target cluster
else
5. Remove the customer(C_i)having low R_i,F_i values within the targeted cluster. (RCE step) }

Here, the threshold values of R and F are considered because Recency and Frequency is very important for Loyalty based Marketing even though their monetary value might change from time to time. In this case, the threshold value for R(t_1) and F(t_2) may be set as 3 for a scale ranging from 1 to 5. This varies according to the user as already mentioned in section II. The customers are retained only if their R and F values are greater than the threshold values set. The other customers are eliminated within that group. Now, the purchase interests of these customers are analyzed and those items and their correlated items are recommended to the entire group.

2. Algorithm for Recursive Association

i - Number of customers within the group obtained from the result of RCE (1, 2, 3...N)
C_i - Targeted customer
n -Total number of products bought by Customer C_i
M_{ij} - Count of product j bought by Customer C_i in targeted cluster
I_i - Frequent itemset for customer C_i
AI_i - Associated itemset for customer C_i
t_3 - Threshold support count for recursive frequent itemset mining
t_4 - Threshold support count for recursive association

1. Target a customer C_i from the result of RCE.
2. Identify the frequent itemsets I_i of C_i.
3. Also identify associated items AI_i of C_i.

for(j=1;j<n;j++){
if($M_{ij} > t_3$)
Add to frequent itemsets I_i of Customer C_i
Identify corresponding associated itemsets AI_i=I_1 I_2, I_1, I_3, I_2 I_3,
I_1 I_n.....>t_4;
I=I_i;
AI=AI_i;
else
Discard the product M_{ij} }

Frequent Itemset Mining (2) is applied to the purchase details of the customers retained in the targeted group obtained using RCE. Initially, the frequent itemset mining and the MBA for the individuals in the group is carried out. The customers from the output of Recursive Cluster Elimination are targeted and their corresponding frequent itemsets are considered. From these itemsets purchased by the customers the possible and the frequent combination of itemsets are identified by using the machine learning algorithm tools. From the available combination of itemsets the frequent itemset bought by the customers in the targeted cluster is identified and recursive market basket analysis is made for them. The resulting combination of itemsets is recommended for the entire group which covers almost all the customers present in the targeted cluster. Here the thresholds t_3 represent the summation of support count of maximum products present to the total number of products of each customer and t_4 represent the average of total number of combination varieties of itemsets present. If the support count of every product is greater than the threshold value (t_3) add the item to the frequent itemsets and identify their corresponding itemsets. The itemsets and the associations which are ranked below this threshold are eliminated. If the maximum combination of itemsets is greater than the threshold value (t_4) that item is considered to be resultant itemset that covers the maximum number of customers present in the group.

V. EXPERIMENTS AND DISCUSSIONS

A. Data Set

The data for the empirical study involves the customers' purchase history stored in the outlet and their corresponding transaction detail records stored in the MMO for a period of one year. This study involves a in a simulated dataset containing the records of the customers single outlet and single MMO derived from a real dataset consisting of thousands of customers, hundreds of outlets, specific MMOs and lakhs of transactions. Due to the space constraints, we briefly outline the experimental results of our proposed system using the above result set.

B. Comparison with other models

Table 1 describes specific experimental comparison between our proposed system and other models.

TABLE 1 EXPERIMENTAL COMPARISON

CASES	I	II	III
MODEL	Individual Recommender	Group Recommender	Group Recommender **(Proposed System)**
TARGETS	Individual	Group	Group
CLUSTERING (WEKA)	0.02 sec	0.02 sec	0.02 sec
FIM & MBA (HPCC)	10 sec	11.36 min	42 sec
RECURSIVE CLUSTERING (WEKA)	NO	NO	0 sec
RFIM & RMBA (HPCC)	NO	25 sec	16 sec
Summation	10.02 sec	12 min 03 sec	58.02 sec
Total Time Taken in Outlet	Outlet=n customers Say n=100, n*10.02 sec =17min 10 sec	Clusters =k Say k=3 k*12 min 03 sec =36min 15 sec	Clusters =k Say k=3 k* 58.02 sec = 3 min 20 sec

Inference:

The above result set is obtained by carrying out experimentation in an outlet consisting of n customers (say n=100). Here, the customers are clustered into 3 segments say High, Medium, Low by performing clustering using K-means.

Case I: A particular customer is targeted from any of the segment and techniques such as FIM, MBA are carried out. The total time taken to formulate and send custom built offers to all customers in the outlet is 17 min 10 sec.

Case II: Here, a whole segment is targeted and the techniques such as FIM, MBA are carried out recursively. By performing this, we can send custom built offers to the customers in a particular segment targeted and total time taken to formulate and send offers to all the customers in the outlet is 36 min 15 sec.

Case III: Here, a whole segment is targeted and the techniques such as FIM, MBA are performed for the targeted cluster recursively. Next, the most potential customers within the targeted segment are identified by performing clustering recursively and similarly FIM and MBA are executed recursively. By adopting this kind of approach, we can send custom built offers to all the customers in a segment based on the purchase interests of the most potential customers. The total time taken to formulate and send offers to all the customers in the outlet is 3 min 20 sec.

So by comparing the above cases, Case III performs clustering, FIM and MBA recursively and is considered the best since it is less time consuming and custom built offers are sent to many customers unlike the other cases.

The most potential customers' purchase influences the purchase of the whole segment. When this model was carried out without RCE and with RFIM and RMBA, the same output was obtained. This kind of an approach time involved lot of complexities and time delay. The ratio of taken for the approach with (No RCE+ RFIM+ RMBA) to the approach with (RCE+RFIM+RMBA) was evaluated to 12:1. Since, the approach with (RCE+ RFIM+ RMBA) produced similar results in much lesser time; this is regarded as the accurate method of recommending offers. This magnifies the purchase power of the customers and they can choose from a wide variety of products tagged with offers. The offers for the above mentioned products and the other products associated with it are recommended to the group. As shown in Fig 4, the final results of RFIM and RMBA for the products purchased by the most potential customers are P, O, E, D, B. The offers for these products are recommended to the whole targeted segment. By adopting this model, the execution time can be exponentially reduced and this makes our group recommender model unique compared to other models.

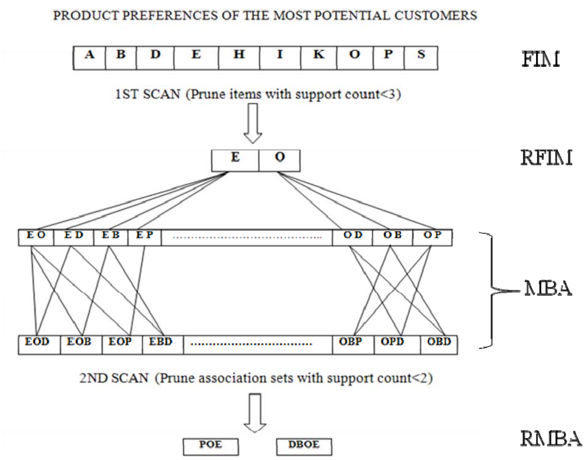

Fig 4: RFIM and RMBA for most potential customers

VI. CONCLUSION

Group Recommender Systems allow businesses to leverage their customer history to notify custom built offers to their customers. In this system, the customer's spending pattern and buying behavior is monitored and examined based on which relevant offers are notified to the customer's mobile. Here techniques such as Recursive Cluster elimination, Recursive Frequent Itemset Mining and recursive market basket analysis provide an enhanced and optimized result using which offers can be correlated and recommended to the customers via SMS. This can be achieved by the communication and collaboration of the two sources namely MMO and outlet. This kind of Product-associated recommendation allows businesses to respond to each customer's current interests and allow the natural associations among different products to guide customers to the right purchase. When such innovative, enticing ads for products and its corresponding associations are promoted from secured and legal organizations such as MMO, customers spare a thought to subscribe to this service which depends on the willingness of him/her. Loyalty based marketing approach is implemented where the customer base is expanded in the MMO and outlet. Rewarding the customers with valuable recommendations based on their input results in boosting their purchase which serves as a boon to the outlet in terms of revenue.

REFERENCES

[1] Prem Melville and Vikas Sindhwani, "Recommender Systems", Encyclopedia of Machine Learning 2010: 829-838

[2] Alberto Jimenez, Prasanna Vanguri, "Cash Replacement through Mobile Money in Emerging Markets: The FISA Approach", July 2010

[3] http:// www.hpccsystems.com

[4] Ian H. Witten; Eibe Frank, Mark A. Hall (2011). "Data Mining: Practical machine learning tools and techniques, 3rd Edition". Morgan Kaufmann, San Francisco.

[5] A. Jameson and B. Smyth, "Recommendation to Groups", In P. Brusilovsky, A. Kobsa, and W. Nejdl, editors, The Adaptive Web, volume 4321 of Lecture Notes in Computer Science, pages 596–627. Springer, 2007.

[6] J. F. McCarthy and T. D. Anagnost. Musicfx, "An arbiter of group preferences for computer supported collaborative workouts", In CSCW '98: Proceedings of the 1998 ACM conference on Computer supported cooperative work, pages 363–372, New York, NY, USA, 1998. ACM.

[7] Z. Yu, X. Zhou, Y. Hao, and J. Gu, "TV program recommendation for multiple viewers based on user profile merging" User Modeling and User-Adapted Interaction, 16(1):63–82, 2006.

[8] J. Masthoff, "Group modeling: Selecting a sequence of television items to suit a group of viewers. Use Modeling and User-Adapted Interaction" , 14(1):37–85, 2004.

[9] G. Linden, B. Smith, and J. York, "Recommendations: Item-to-item collaborative filtering", IEEE Internet Computing, pages 76–80, Jan/Feb 2003.

[10] J. S. Breese, D. Heckerman, and C. Kadie., "Empirical analysis of predictive algorithms for collaborative filtering", In Proc. of UAI '98, 1998.

[11] Badrul Sarwar, George Karypis, Joseph Konstan, and John Riedl, "Item-Based Collaborative Filtering Recommendation Algorithms", Department of Computer Science and Engineering,, University of Minnesota, Minneapolis.

[12] http://recsys.acm.org/2011

[13] Joel Sobel, "Information Aggregation and Group Decisions", University of California, San Diego, January 17, 2006.

[14] Chia-Cheng Shen, Huan-Ming Chuang, "A Study on the Applications of Data Mining Techniques to Enhance Customer Lifetime Value", Wseas Transactions on Information Science and Applications, Issue 2, Volume 6, February 2009.

[15] Ching-Hsue Cheng, You-Shyang Chen, "Classifying the segmentation of customer value via RFM model and RS theory", Department of Information Management, National Yunlin University of Science and Technology.

[16] Lin-Kai Luo, Deng-Feng Huang, Ling-Jun Ye, Qi-Feng Zhou, Gui-Fang Shao, and Hong Peng, "Improving the Computational Efficiency of Recursive Cluster Elimination for Gene Selection", Ieee/Acm Transactions On Computational Biology And Bioinformatics, Vol. 8, No. 1, January/February 2011.

[17] Pramod Prasad, Dr. Latesh Malik, "Using Association Rule Mining for Extracting Product Sales Patterns in Retail Store Transactions", International Journal on Computer Science and Engineering (IJCSE),Vol. 3, May 2011.

[18] Chun-Kit Chui, Ben Kao, and Edward Hung, "Mining Frequent Itemsets from Uncertain Data", PAKDD 2007: 47-58.

[19] Pang-Ning Tan, Michael Steinbach, Vipin Kumar, "Introduction to Data Mining"

[20] Troy Raeder, Nitesh V. Chawla, "Market Basket Analysis with Networks", Interdisciplinary Center for Network Science and Applications, University of Notre Dame, USA.