2011 2nd International Conference on Networking and Information Technology
IPCSIT vol.17 (2011) © (2011) IACSIT Press, Singapore

Product Recommendation Based on Customer Lifetime Value

—An Electronic Retailing Case Study

Zahra Tabaei and Mohammad Fathian

Department of Industrial Engineering, Iran University of Science and Technology, Tehran, Iran
Zahra_tabaei@ind.iust.ac.ir, fathian@iust.ac.ir

Abstract—Electronic commerce contains a lot of applications. Product recommendation is one of useful application of electronic commerce. Recommending right products to right customers enhances the customer's utility and firm profitability. Different customer types have different interests, so firms should firstly segment customers in to groups and recommend right product to them. The main purpose of this paper is to clustering customers based on Customer lifetime Value and then recommending product to different groups of customers with association rule mining technique. We use one electronic retailing in our study.

Keywords-component; product recommendation; Customer Lifetime Value; electronic commerce;data mining; customer loyalty

1. Introduction

With growing electronic commerce in all over the world, considering to customers and their needs and desired is so important. Customers attract to electronic retailing with high compliance with their desired [1]. So considering the voice of customers and providing right product for them is one of key business successful. In electronic commerce because of not face to face relation with customers this subject is more important.

Different groups of customers prefer some special products. Customer's type recognition is one of main aim of each business and firms know who want what? [2] So firms should try finding customer groups according to customer behavior.

Customer loyalty is a suitable feature for segmenting customers. Customers past purchasing behavior show their loyalty [3]. The customer is loyal, if he/she purchases more at his/her lifetime, buys products recently and spending more money during the lifetime. But if a customer doesn't purchase recently, total number of his/her purchasing is low and spent money is low, he/she is disloyal customer.

RFM model is one of Customer Lifetime Value models that have been used for customer loyalty in a lot of studies is useful model ([3], [5], [6], [7], [8], [9]). This model is easy. The parameters need for constructing model is available. RFM is behavioral model and consider all past customer purchases to prospect the future customer behaviors.

With using CLV models for customer loyalty and data mining techniques for customer segmentation according CLV, we can recommend right products to right customers and provide individual marketing decision for each customer. So customers receive products according their requirements, will been satisfied and purchase more at time and finally they will be loyal customers. On the other hand, firms spend low lost for customer retention and reach profitability.

The organization of this study is as follows. In Section II, we review the related works in customer loyalty and recommendation systems. In section III, we propose methodology for product recommendation with findings in our case study. Finally, a discussion on the study results is described in section IV.

2. Literature Reviewe

2.1 Customer Lifetime Value

Customer Lifetime Value models are used widely to identify customer loyalty and determinate marketing strategies to different groups of customers [10]. There are several models for CLV. Some of them are based on past behavioral models and some of are based on future customer revenue.

One of useful behavioral model for customer loyalty is RFM model. RFM was constructed based on 3 variables [11]:

- Recency: refers the duration time between last customer purchasing and present time.
- Frequency: refers the total number of customer purchasing during life time.
- Monetary: refers the average money spending during past customer purchases.

In weighted RFM, each normalized variables multiply to their estimated weights which defined by experts. Experts make decision about each weight according to their experiences [12].

Past customer value (PCV) is another CLV model based on the past monetary value of customers. PCV extrapolate the past monetary spent by customers in past purchases into present time. This model doesn't consider the future expected customer value [13].

Some LTV models, consider future monetary value of customers prospect the future monetary value of customers. These models extrapolate the future prospected monetary will spend by customer into present time [14].

Some studies such as ([14], [16]) have prospected future customer acquisition profits and active time for each customer, then they have sued them to estimate the LTV.

Retention rate is an important factor in LTV estimation. If the firm identifies the right retention rate, it can estimate the valid LTV. Some studies such as ([17], [18]) have constructed the LTV based on retention rate.

Buyer and seller's commitment is an effective factor that has been considered in [19]. In fact customer commitment builds the loyalty and seller's commitment is led to customer trust and as result LTV models can been constructed based on these commitments.

According to literature review about customer value we can understand although LTV model is the best one, but prospect the acquired future revenue from customers and their lifetime are so difficult. But RFM model is a behavioral flexible model and can adapt itself to every business.

2.2 Recommendation Systems and Association Rule

According to reference [6] "Recommender systems are technologies that assist businesses to implement one-to-one marketing strategies". Recommender systems review the history of past customers purchasing and identify the customer desired products. Using recommendation systems are led to both customer and firm profitability [20].

For product recommendation first stage is to gathering information about past customer's purchases. This information contains all data about any combination of products have been purchased by each customer groups. With this information recommendation system extracts the useful rules and can propose the right products to customers [21].

One of useful data mining technique have been used in recommendation systems is association rule mining. This technique extracts useful patterns according to customer transaction records ([22], [23]).

After rules extraction, recommendation system works according to following method:

If x_u is the set of all products have been purchased by customers at past, for each $x \rightarrow y$ rules in RS_{i}, if $x \subseteq x_u$, then all $y \subseteq x_u$ products were proposed to candidate customer. RS_{i} is the set of rules related to customer group j [24].

Tow rule evaluation metrics contain 'support 'and 'confidence'. For every $x \rightarrow y$ rule, support refers to fraction of transaction that contain both *x* and *y*, confidence measures how often items (products) in *y* appear in transactions that contain x [25].

3. Methodology With Findings

In our survey we have used weighted RFM model for customer loyalty calculation. In the second stage customers have been segmented according weighted RFM and have been assigned ranking score to each groups. Finally we use association rule mining to extract useful rules for product recommendation. The brief stage of our methodology has been shown in fig. 1.

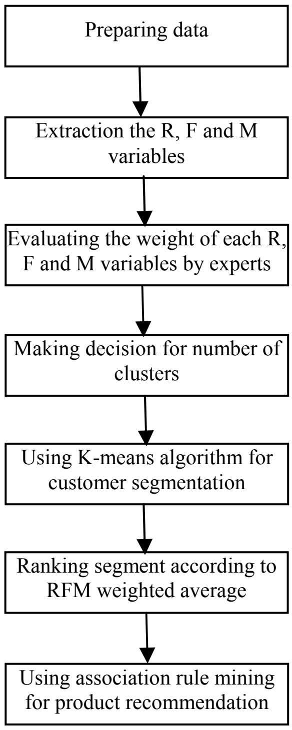

Figure 1. Research design

3.1 Data preparing

We use a data set from one of electronic retailing, in our study. This electronic retailing sells different types of CD (software, educational, film, game and music) to customers. Data set contains 65535 transaction records with 22086 customers. Each record contains customer ID, last purchase time, price of purchased product and product ID. The R, F and M value for each customer were extracted. In our survey the customers who purchase recently has higher R value.

3.2 Clustering Customer with similar FRM value

After evaluating the weight of each variable with AHP analysis, K-means algorithm has been used in order to customer segmentation. We set the number of clusters at 8, because each variable R, F and M is higher (↑) or lower (↓) than its average (2*2*2=8). Table I shows the result of customer segmentation according to normalized weighted R, F and M value.

As table I shows cluster 2 and 5 have few customers, they may be outliers. So with putting aside them only 4 cluster patterns ($R\downarrow F\downarrow M\downarrow$,$R\uparrow F\uparrow M\uparrow$,$R\uparrow F\uparrow M\downarrow$, and $R\downarrow F\downarrow M\downarrow$) have been identified. so it seems that we should decrease the number of clusters. We set number of clusters at 4 and again perform K-means algorithm. The results with 4 clusters have been shown in table II. Results show that with 4 clusters, 3 pattern types have been identified and cluster 4 with 9 members is outlier. So we repeat last stage again with 3 clusters. There is no outlier customers. The results with 3 clusters have been shown in table III. Results show that with 3 clusters, 3 pattern types have been identified and there is no outlier customers. So the results are valid.

TABLE I. 8 CLUSTER RANKING BY WEIGHTED SUM OF NORMALIZED RFM VALUE

cluster	No. of customers	Weighted Recency	Weighted Frequency	Weighted Monetary	Pattern Type	WRFM	Rank
1	13374	0.06	0.000	0.000	$R\downarrow F\downarrow M\downarrow$	0.060	8
2	3	0.728	0.164	0.058	$R\uparrow F\uparrow M\uparrow$	0.950	1

3	2127	0.586	0.004	0.001	$R\uparrow F\uparrow M\downarrow$	0.591	4
4	1	0.134	0.048	0.038	$R\downarrow F\uparrow M\uparrow$	0.220	7
5	1979	0.254	0.002	0.001	$R\uparrow F\downarrow M\downarrow$	0.257	6
6	71	0.708	0.037	0.012	$R\uparrow F\uparrow M\uparrow$	0.757	2
7	2462	0.692	0.006	0.002	$R\uparrow F\uparrow M\uparrow$	0.700	3
8	2069	0.438	0.003	0.001	$R\uparrow F\uparrow M\downarrow$	0.442	5
Overall Average		0.236	0.0021	0.0011			

TABLE II. 4 CLUSTER RANKING BY WEIGHTED SUM OF NORMALIZED RFM VALUE

cluster	No. of customers	Weighted Recency	Weighted Frequency	Weighted Monetary	Type	WRFM	Rank
1	13937	0.065	0.000	0.000	$R\downarrow F\downarrow M\downarrow$	0.065	4
2	4856	0.638	0.005	0.002	$R\uparrow F\uparrow M\uparrow$	0.645	2
3	3284	0.336	0.002	0.001	$R\uparrow F\downarrow M\downarrow$	0.369	3
4	9	0.713	0.116	0.034	$R\uparrow F\uparrow M\uparrow$	0.863	1
Overall Average		0.236	0.0021	0.0011			

TABLE III. 3 CLUSTER RANKING BY WEIGHTED SUM OF NORMALIZED RFM VALUE

cluster	No. of customers	Weighted Recency	Weighted Frequency	Weighted Monetary	Type	WRFM	Rank
1	13937	0.065	0.000	0.000	$R\downarrow F\downarrow M\downarrow$	0.065	3
2	4863	0.648	0.006	0.002	$R\uparrow F\uparrow M\uparrow$	0.645	1
3	3286	0.338	0.002	0.001	$R\uparrow F\downarrow M\downarrow$	0.369	2
Overall Average		0.236	0.0021	0.0011			

As have been shown at table III, 3 pattern types have identified. First one is $R\uparrow F\uparrow M\uparrow$ with the highest CLV ranking. This segment contains loyal customers, because they have purchase recently, the total number of their purchasing is high and they have spent much money. These customers are valuable for firm. Second valuable segment with $R\uparrow F\downarrow M\downarrow$ pattern contains new customers, because they have purchased recently but their frequency and monetary values are low. Finally the least valuable customers are belonged to cluster 1 with $R\downarrow F\downarrow M\downarrow$ pattern type. They are disloyal customers and all variables are low. The most customers have been belonged to this segment.

3.3 Product Recommendation

The association rule mining is used for extracting useful rules in each separate segment of customers and then proposing desired product to customers. The extracted rules from loyal, disloyal and new customer segments have been shown in tables IV, V and VI.

TABLE IV. EXTRACTED RULES FOR LOYAL CUSTOMERS

Rule	Instances	Support (%)	Confidence (%)
$M\rightarrow F$	3388	69.661	85.394
$G,M,E\rightarrow F$	1545	31.778	85.198
$G,M\rightarrow F$	2052	42.199	85.046
$G,M,S\rightarrow F$	1527	31.408	84.882
$G,M,S,E\rightarrow F$	1277	26.269	84.585
$S\rightarrow E$	3395	69.805	84.069
$F\rightarrow M$	3447	70.874	83.933
$E\rightarrow S$	3404	69.99	83.847
$G,M,S\rightarrow E$	1527	31.408	83.639

$G,F,S \rightarrow E$	1581	32.518	83.565
$G,M,S,F \rightarrow E$	1296	26.66	83.346
$G,E \rightarrow S$	2090	42.98	83.166
$G,E,F \rightarrow S$	1590	32.703	83.092
$S,G \rightarrow E$	2099	43.165	82.81
$G,E,F \rightarrow M$	1590	32.703	82.778
$G,M,E \rightarrow S$	1545	31.778	82.665
$G,F \rightarrow M$	2125	43.7	82.126
$G,M,E,F \rightarrow S$	1316	27.071	82.08
$S,G,F \rightarrow M$	1581	32.518	81/985
$G,S,E,F \rightarrow M$	1321	27.174	81.77
$E,M \rightarrow F$	2264	46.557	81.589
$S,M \rightarrow F$	2228	45.817	81.202
$S,M \rightarrow E$	2228	45.817	80.709
$S,F \rightarrow E$	2268	46.639	80.608
$E,F \rightarrow M$	2295	47.194	80.488

Note: G=Game, E=Educational, F=Film, S=software, M=music

TABLE V. EXTRACTED RULES FOR DISLOYAL CUSTOMERS

Rule	Instances	Support (%)	Confidence (%)
$G,E \rightarrow S$	163	1.170	74.847
$G,S \rightarrow E$	175	1.256	89.714
$G,M \rightarrow F$	165	1.184	68.485
$G,F \rightarrow M$	169	1.213	66.864
$S,M \rightarrow E$	214	1.535	66.355
$E,M \rightarrow S$	222	1.593	63.964
$S,F \rightarrow E$	194	1.392	62.371
$F,E \rightarrow M$	201	1.442	60.697
$F,E \rightarrow S$	201	1.442	60.199

Note: G=Game, E=Educational, F=Film, S=software, M=music

TABLE VI. EXTRACTED RULES FOR NEW CUSTOMERS

Rule	Instances	Support (%)	Confidence (%)
$E \rightarrow S$	1837	55.907	84.586
$S \rightarrow E$	1854	56.425	83.81
$F \rightarrow M$	1793	54.568	83.594
$M \rightarrow F$	1793	54.568	83.595

Note: G=Game, E=Educational, F=Film, S=software, M=music

As have been shown, the rules have been extracted from disloyal segment have low support (1%) and the maximum confidence is 74%. Reference [6] has proposed that for disloyal customers preference-based CF Collaborative Filtering (CF) method is useful and can improve the recommendation quality. But for loyal and new customers the minimum support and confidence are high $F \rightarrow M$ so the results are reliable.

4. Conclusion

In this survey product recommendation was studied. At first customers were segmented with CLV model and three types of customers were identified: loyal, disloyal and new customers. Then association rule mining was used for extracting rules from each segment. Results show that the rule support and confidence for loyal and new customers is relatively high, so this method is useful and reliable for these groups. But because of low support for disloyal customers this method is not so useful and CF method is applicable for disloyal segment, as proposed at [6].

5. References

[1] L.-S. Chen, F.-H. Hsu, M.-C. Chen and Y.-C. Hsu, “Developing recommender systems with the consideration of product profitability for sellers” Information Sciences, Vol. 178, No. 4 , February 2008, pp. 1032-1048 , doi:10.1016/j.ins.2007.09.027.

[2] D.-R. Liu and C.-H. Liou, “Mobile commerce product recommendations based on hybrid multiple channels”, Electronic Commerce Research and Applications, Vol. 10. No.1, January-February 2011, pp. 94-104, doi:10.1016/j.elerap.2010.08.004.

[3] H. –C Chang H. -P and Tsai (2011) ‘Group RFM analysis as a novel framework to discover better customer consumption behavior’, Expert Systems with Applications,Vol. 38, No. 12, November-December 2011, doi:10.1016/j.eswa.2011.05.034.

[4] M. Khajvand, K. Zolfaghar, S. Ashoori, and S. Alizadeh, “Estimating customer lifetime value based on RFM analysis of customer purchase behavior: Case study,” Procedia Computer Science,Vol. 3,pp.57-63, February 2011.

[5] M. Khajvand, and M.J. Tarokh, “Estimating customer future value of different customer segments based on adapted RFM model in retail banking context,” Procedia Computer Science,Vol. 3,pp.1327-1332, February 2011

[6] D.-R. Liu, and Y.-Y Shih, (2005) “Hybrid approaches to product recommendation based on customer lifetime value and purchase preferences,” Journal of Systems and Software, Vol. 77, pp. 181–191, August 2005

[7] C.-H. Cheng , and Y.-S. Chen, “Classifying the segmentation of customer value via RFM model and RS theory,” Expert Systems with Applications, Vol. 36, pp.4176-4184, April 2009.

[8] C.-H. Cheng and H.-P. Tsai “Group RFM analysis as a novel framework to discover better customer consumption behavior,” Expert Systems with Applications, in press.

[9] S.M. Seyed Hosseini, A. Maleki, and M.R. Gholamian, “Cluster analysis using data mining approach to develop CRM methodology to assess the customer loyalty,” Expert Systems with Applications, Vol. 37, pp.5259-5264, July 2010.

[10] J. Miglautsch “Thoughts on RFM Scoring”, *Journal of Database Marketing, Vol. 8, No.*1, 2000, pp. 67-72.

[11] B. Stone, Successful direct marketing methods, Lincolnwood, IL: NTC Business Books, 1995, pp. 37-57.

[12] V. Kumar, and W.J. Reinartz, Customer Relationship Management: A Data based Approach, First ed., Wiley & Sons, 2006, pp.108-128.

[13] V. Kumar and R. Venkatesan, “Who are the multichannel shoppers and how do they perform?: Correlates of multichannel shopping behavior”, Journal of Interactive Marketing, Vol. 19, No. 2, March 2005, pp. 44-6.

[14] C. S. Yamamoto Sublaban and F. Aranha , “Estimating cellphone providers' customer equity”, Journal of Business Research, Vol. 62, No. 9, September 2009, 891-898, doi:10.1016/j.jbusres.2008.10.006.

[15] A.P. Sant’Anna, and R.O. Ribeiro, “Statistical modeling and probabilistic composition in the prediction of the customer lifetime,” Benchmarking: An International Journal, Vol. 16, pp.335 – 350, 2009.

[16] N. Glady, B. Baesens, and C. Croux, “A modified Pareto/NBD approach for predicting customer lifetime value,” Expert Systems with Applications, 36:2, 2062-2071, March 2009.

[17] E. Malthouse, and F. Mulhern, “Understanding and using customer loyalty and customer value,”,Journal of Relationship Marketing, Vol. 6, pp.59 - 86,September 2008.

[18] P. Hidalgo, E. Manzur, S. Olavarrieta, and P. Farias, “Customer retention and price matching: The AFPs case,” Journal of Business Research, Vol. 6, pp.691-696, June 2008

[19] G.E. Fruchter and S.P. Sigué, S. P. “Social relationship and transactional marketing policies—Maximizing customer lifetime value,” Journal of Optimization Theory and Applications, Vol. 142, pp.469–492, 2009.

[20] Y. H. Cho and J. K. Kim, “Application of Web usage mining and product taxonomy to collaborative recommendations in e-commerce”, Expert Systems with Applications, Vol. 26, No. 2, February 2004, pp. 233-246, doi:10.1016/S0957-4174(03)00138-6.

[21] H. Yun, D. Ha, B. Hwang and K. Ho Ryu, “ Mining association rules on significant rare data using relative support”. The Journal of Systems and Software,Vol. 67, No. 3, September 2003, pp. 181–191,doi:10.1016/S0164-1212(02)00128-0.

[22] S. W. Changchien and T.-C Lu, “Mining association rules procedure to support on-line recommendation by customers and products fragmentation”, Expert Systems with Applications, Vol. 20, No. 4, , May 2001, pp. 325-335, doi:10.1016/S0957-4174(01)00017-3.

[23] D.-Ren Liu, C.-H,Lai and W.-J. Lee, “A hybrid of sequential rules and collaborative filtering for product recommendation”, Information Sciences, Vol. 179, No. 29, September 2009, pp. 3505-3519, doi:10.1016/j.ins.2009.06.004

[24] B. Sarwar, G. Karypis, J. JKonstan and J. Riedl, “ Analysis of recommendation algorithms for e-commerce”. In: Proceedings of the ACM Conference (Electronic Commerce), pp. 158–167, 2000, doi:10.1145/352871.352887.

[25] J. Han and M. Kamber, Data Mining: Concepts and Techniques, 3rd ed ,Morgan Kaufmann, 2000, pp. 186-220, Jiawei Han and Micheline Kamber.