A RFMV Model and Customer Segmentation Based on Variety of Products

Saman Qadaki Moghaddam
Department of Electrical, Computer and IT Engineering, Qazvin Azad University, Qazvin, Iran
sam13671367@yahoo.com
Neda Abdolvand*
Department of Social Science and Economics, Alzahra University, Tehran, Iran
n.abdolvand@alzahra.ac.ir
Saeedeh Rajaee Harandi
Department of Social Science and Economics, Alzahra University, Tehran, Iran
saeedeh.rh@gmail.com

Received: 19/Dec/2016 Revised: 18/Aug/2017 Accepted: 25/Aug/2017

Abstract

Today, increased competition between organizations has led them to seek a better understanding of customer behavior through innovative ways of storing and analyzing their information. Moreover, the emergence of new computing technologies has brought about major changes in the ability of organizations to collect, store and analyze macro-data. Therefore, over thousands of data can be stored for each customer. Hence, customer satisfaction is one of the most important organizational goals. Since all customers do not represent the same profitability to an organization, understanding and identifying the valuable customers has become the most important organizational challenge. Thus, understanding customers' behavioral variables and categorizing customers based on these characteristics could provide better insight that will help business owners and industries to adopt appropriate marketing strategies such as up-selling and cross-selling. The use of these strategies is based on a fundamental variable, variety of products. Diversity in individual consumption may lead to increased demand for variety of products; therefore, variety of products can be used, along with other behavioral variables, to better understand and categorize customers' behavior. Given the importance of the variety of products as one of the main parameters of assessing customer behavior, studying this factor in the field of business-to-business (B2B) communication represents a vital new approach. Hence, this study aims to cluster customers based on a developed RFM model, namely RFMV, by adding a variable of variety of products (V). Therefore, CRISP-DM and K-means algorithm was used for clustering. The results of the study indicated that the variable V, variety of products, is effective in calculating customers' value. Moreover, the results indicated the better customers clustering and valuation by using the RFMV model. As a whole, the results of modeling indicate that the variety of products along with other behavioral variables provide more accurate clustering than RFM model.

Keywords: Clustering; Data Mining; Customer Relationship Management; Product Variety; RFM Model.

1. Introduction

Today, increased competition between organizations has led them to seek a better understanding of customer behavior and their partners through innovative ways of storing and analyzing the customer information [1]. One of the major challenges of customer relationship management (CRM) is to establish more profitable and long-term relationships with customers [1] [2]. The emergence of new computing technologies has brought about major changes in organizations' ability to collect, store and analyze large datasets. Therefore, over thousands of data can be stored for each customer, enabling the analysis of customer purchasing history [3]. To understand and identify valuable customers, they should be segmented based on their behavior [4] [5]. The most frequently used model in customer segmentation is the RFM model, which consists of three behavioral variables: R (Recency), F (Frequency) and M (Monetary) [4]. Buying goods or services is the organization's customer communication channel. It can be used, along with other behavioral variables, to better understand and categorize customers' behavior, and adopt appropriate strategies such as up-selling and cross-selling. Since the RFM model is incomplete, studies have tried to improve and develop the model by adding other variables. Although, purchasing goods and services is the main channel of communication in businesses, few studies have considered the importance of identifying the products purchased. Since the variety of products is one of the main parameters of assessing customer behavior, studying this factor in the field of business-to-business (B2B) communication represents a vital new approach. Accordingly, based on the sales strategies of up-selling and cross-selling as the important tools of increasing customer value to organizations, this study tries to improve the RFM model by adding the variable of product variety (V). This will help customer recognition and provide more accurate understanding of the customer in the field of B2B.

* Corresponding Author

For the purpose of this research, the literature is reviewed first. The research methodology then addresses the model development and measurement tools. After the statistical analysis is explained, the conclusions and recommendations of the study are discussed.

2. Literature Review

In recent decades, social and economic changes in the interaction between organizations and customers have made customer relationship management (CRM), one of the most important processes in the business environment (followed by business partner relationship management) [3]. CRM can be defined as the collecting, storing, and analyzing customer data in order to increase customers' loyalty and value, and to increase organizational benefits [4] [6] [7]. In order to increase customer satisfaction and prevent customers from leaving the organization, the organization should focus on segmentation and meeting customers' individual needs [8]. Generally, "customer segmentation" is the process of dividing customers into different groups based on geographical, demographic and ethnological information, in order to adopt strategies tailored to each group based on the consumption of goods and services and customers' purchase history [4] [8] [9]. The segmentation process continues to achieve a harmony in customer value; along with detecting and correct placement of clients in related groups, it is one of the most essential factors in CRM and business success [4] [9]. Since there isn't a preferred approach to customer segmentation, the best model is one that draws a proper insight of current and potential customers, and help organizations to achieve effective markets and appropriate customer feedback [10]. Different techniques and forms of data analysis can be used for customer segmentation; the most common are the use of data mining techniques and customer behavior variables. Data mining is the process of discovering and extracting hidden patterns from large amounts of data [2]. The RFM model, which tries to identify customers based on their behavioral characteristics, is one of the most widely used [4] [11] [12]. This behavioral model uses three criteria of customer transactions: recency, frequency and monetary. These categories are then interpreted by data mining tools and algorithms. Thus, the RFM model predicts the customer's next movements based on their behavior [13]. It is not only considered as a segmentation model, but also includes the concept of customer value. That is why many studies use this model for the discovery and analysis of customer value based on past purchase behavior [13]. On the basis of their purchasing behavior, the RFM reveals that valuable customers are those with the highest frequency and monetary value and the lowest recency. Despite its simplicity of understanding, interpretation, and implementation [14] [15] [16], the model has shortcomings, such as inattention to personal characteristics and demographic variables of customers [14] [16]. Thus, studies have been done with the aim of increasing the accuracy of RFM model output by adding variables which can be categorized in two groups of customer-oriented and product-oriented. In the first group, factors such as customer lifetime value and imposed costs (eg, [4] [14] [17] [18] [19] [20] [21]), are taken into consideration, and in the second group, product life cycle (eg, [16]) as well as the types of products are considered as the basis for the development of model [14] [16]. Cheng and Chen [13] based their study on RFM model, and, by adding customer credit to the model, segmented customers of an e-services company in Taiwan. After customer clustering using their proposed LEM2 algorithm, they extracted important rules for future marketing decisions and company's strategies. Moreover, Khvajvand et al [1] in their study in the cosmetics industry have classified and valuated customers based on the number of products purchased by customers along with other behavioral variables in their proposed model. Given the RFM variables. Chang and Tsai [22] generalized RFM model to the groups of products and services, and proposed a concept, named Group RFM or the GRFM, to discover better customer consumption behavior. They believed that the final value of customers purchase should be calculated, and behavioral variables should be classified based on purchases. Noori [11] added the variable of customer deposit to the RFM model to classify mobile banking users. He indicated that identifying customers by a behavioral scoring facilitates marketing strategy assignment.

Most studies used customer-business approach (B2C) to develop model, and a few studies have been conducted in the field of business-business (B2B). In the field of B2B, the firm will reap huge profits from the large volume of purchases. Therefore, understanding customers would be the success key in his area. Some researchers proposed reasons such as quality of communication, trust, participation, satisfaction, increased buying, and organizational changes as the effective factors of maintaining customers in B2B organizations [16]. Hosseini et al. [16] have added product life cycle / product duration to RFM model, and classified customers into 34 categories in a B2B company. Kandeil et al. [4] also used LRFM clustering techniques to segment customers of a B2B distributor. Moreover, Venkatesan and Kumar [17] have studied the customer lifetime value of a B2B Manufacturer, and indicated that selecting customers based on the CLV provides more profits than other metrics. RFM evolution is briefly shown in Table 1.

Table 1. RFM Model Evolution

Added Variable	Resource	Industry
Customer Lifetime Value	(Ansari & Riasi, 2016)	Steel Company
	(Kandeil et al., 2014)	Distribution Company
	(Reinartz & Kumar, 2000)	Distribution Company
	(Alvandi et al., 2012)	Bank
	(Venkatesan & Kumar, 2004)	International Software and Hardware Producing Company
	(Wei et al., 2012)	Children's Dental Clinic
Customer Lifetime Value and Imposed Costs	(Soeini & Fathalizade, 2012)	Insurrance

Added Variable	Resource	Industry
Groups of Products and Services	(Chang and Tsai, 2011)	Educational Organization
Number of Products	(Khvajvand et al, 2011)	Cosmetics Industry
Customer Credit	(Cheng and Chen, 2009)	E-Services Company
Product Life Cycle / Product Duration	(Hosseini et al., 2010)	
Customer Deposit	(Noori, 2015)	Bank

According to studies carried out in this context, few studies have considered the importance of the product and their varieties, and reasonable valuation of customers based on product variety is not provided. In addition, most surveyed industries were in the field of B2C, and there has been insufficient attention to B2B domains. Therefore, with regard to the importance of the variety of products in leveraging strategies such as up-selling and cross-selling, increasing profitability of organizations, and customer maintenance, this study aims to add the variable of variety of products as a new behavioral variablein the RFM model, and to classify customers of the company in the field of B2B.

3. Methodology

Different algorithms and techniques are being used to classify, valuate, and model customer buying behaviors which are classified in two groups of clustering and association rule mining. Accordingly, clustering techniques and K-Means algorithms are used in this study. The CRISP-DM methodology, which is one of the greatest analytical methods for data mining projects, is used in this study. Moreover, K-Means algorithm and Silhouettes' measure of clustering quality is used to measure and to determine the number of clusters. This is done on 924 normalized records of customers by using SPSSModeler.14 software. Then, the ANOVA test is run on obtained clusters, and Duncan's post hoc test is used to determine distinct clusters using the SPSS.16 software. Finally, hierarchical analysis method is used for weighting the R, F, M, V variables, and to calculate the value of each customer.

3.1 Data Analysis

This study was conducted based on the CRISP-DM methodology which consists of six phases of business understanding; namely data understanding, preparation, modeling, evaluation and deployment [23] [24]:

Phase 1. Business Understanding

At this phase an overview of the type of business, Based on which the research is done is obtained, and the overall targeting is done based on the current strategies and business nature [24]. The objective of this study is to evaluate the variety of products purchased by customers along with other behavioral variables of the RFM model in the field of food and sanitary distribution company in the field of B2B to determine the similarity of customers based on their equity value, and to make company capable to identify high-value customers.

Phases 2 and 3. Data understanding and Preparing

The second phase involves collecting, describing and evaluating the data quality. In general, the aim of this phase is to select the appropriate data source in order to reach the goal [24] [25]. Thus, using the information available in our database, the data of over 1,000 customers, including customer code, name and total amount of purchase, are recorded in the database. In this stage, the output reports are received using Microsoft Excel 2010, and after the initial data sorting, the index values are extracted. Finally, the required variables, namely, the number of customers' purchases, customer's last purchase date, purchase amount, and the variety of purchased products in the last 6 months are extracted. Then data are prepared for modeling (third phase). Data preparation includes the process of excluding outliers and data normalization. In this study, 1112 records are collected; the number of data after removing duplicates data is reduced to 967 data records. Given that the values of 2, 2.5 and 3 are the common standard deviations values for detecting distortion, by considering these values and testing them on the data, the value 3 was used to determine the distortion data. Finally the number of data after removing duplicates and invalid data is reduced to 924 data records. The SPSSModeler.14 software is used for identifying outliers.

Then based on MIN-MAX method and using the formulas (1, 2, 3 and 4), indexes are normalized in the range of zero and one in order to prepare for modeling phase (next phase). In these formulas, MAXr, MAXf, MAXm, and MAXv represent the highest value, and MINr, MINf, MINm, and MINv represents the lowest value in the dataset, which are ultimately normalized to the final values of V', M', F', R'.

$$R' = \frac{(R - MIN_R)}{(MAX_R - MIN_R)} \qquad M' = \frac{(M - MIN_M)}{(MAX_M - MIN_M)}$$

$$F' = \frac{(F - MIN_F)}{(MAX_F - MIN_F)} \qquad V' = \frac{(V - MIN_V)}{(MAX_V - MIN_V)} \qquad (1, 2, 3 \& 4)$$

Phase 4. Modeling

In this phase, the proposed model is explained and developed. According to the purpose specified in the first phase, the CRISP- DM methodology as well as K-means algorithm that is one of the most famous and widely used clustering algorithm are used to develop the model. The K-Means algorithm is based on partition clustering that the number of clusters should be preset. Then, based on the number of initial clusters, data are placed in different clusters. One of the fundamental issues in this algorithm is to find the optimal number of clusters. In this study, the silhouette measure that uses the combination of two criteria of solidarity and density is used to determine the optimal number of clusters ranging from 3 to 10. In this measure, the average distance of samples in a cluster is compared with the average distance of samples in other clusters. The results are indocated in table 2. According to the results, the best value for silhouette criteria in K-means algorithm is 3 clusters with the value of 0.540. Then, by increasing the number of clusters to a value of 10, standard numeric value of silhouette is declining. Whatever the output of this standard is closer to one, the quality of clusters resolution is better. Therefore, the optimal number of clusters will be 3. After that, the Mean value of behavioral variables in each cluster is evaluated. This helps to ensure that before proceeding to the next test, customers are grouped in significant clusters.

Table 2. Values of Silhouette Criteria in K-Mean Algorithm

	Cluster's Number	Value
1	**3**	**0.540**
2	4	0.491
3	5	0.492
4	6	0.461
5	7	0.468
6	8	0.476
7	9	0.445
8	10	0.448

- **Valuation and Determination of the RFMV Model Parameters Weight Using a Hierarchical Analysis:**

The weight of each parameter should be determined in order to rate and cluster customers. In this study, hierarchical analysis method, the most efficient decision support technique, is used for weighing the indexes of R, F, M and V. Then, respondents were asked to do paired comparison, and give the value of 1 to 9 to each index. The company's managing director and sales manager's ideas are used to determine the weight of parameters through the Expert choice software. According to formula (5), considering the total value of 1 and Inconsistency Index of 0.1, the total weight of 0.054, 0.075, 0.636, and 0.236 are obtained for the indexes of R, F, M and V respectively. Since the Inconsistency of variables is less than 0.1, the results are reliable.

$$WR + WF + WM + WV = 1 \quad (5)$$

Then, the indexes' value of each customer is calculated based on the formula (6):

$$CLVi = WR'_{*}R' + WF'_{*}F' + WM'_{*}M' + W\,v'_{*}V' \quad (6)$$

After determining the indexes' weight and value of all customers, customers' value is determined according to the Mean value of each cluster. Then, each index Mean is determined from the data extracted in the first phase (Table 3).

Table 3. Customers' Average Value

Cluster	Average Value	Total Average Value	Cluster Value Than the Average
1	0.33040	0.16718	Upper
2	0.07996		Lower
3	0.09155		Lower

Then, the total value of each cluster is divided by the number of members of that cluster, in order to determine the mean value of each cluster. The value obtained for each cluster is shown in Table 4.

Table 4. Clusters' Indexs' Mean

Cluster	$\overline{R''}$ Average	$\overline{F''}$ Average	$\overline{M''}$ Average	$\overline{V''}$ Average	Average Value
Cluster 1	0.0043	0.0384	0.1763	0.1113	0.3304
Cluster 2	0.0332	0.0027	0.0282	0.0154	0.0796
Cluster 3	0.0084	0.0110	0.0442	0.0276	0.0915
Total	0.0154	0.0174	0.0829	0.0606	0.1672

With regard to the main objective of the study, and given the value of customers based on behavioral variables, customers are categorized into three groups: high-value, middle-value and low-value. Due to the fact that each index can be dual-mode (above average and below average) the factors related to the variety of products are examined. The symbols ↑ and ↓ are used to compare the average of each cluster with total means. If the index is higher than the total average, the symbol ↑ is used and ↓ is used otherwise. The lower R (Recency) is better which means that customer purchased the products in closer intervals, while the value upper than mean is better for other variables. Given the customer value based on the behavioral variables, three categories of customers are considered in this study including, high-value customers that has better situation in terms of the average value of the indexes in the data, and indexes' mean in the cluster. This group has lower recency, and in terms of the other variables is better than other clusters. Moreover, the variety of products in the cluster is significantly higher and is distinct from other clusters.

Middle-value customers whose customers have an average level in terms of indexes, and have relatively low distance with the mean values obtained for indexes. In terms of the recency and variety of the products these customers are in the middle. In this study, nearly 51% of our customers are in this level. With regard to the number of members in the cluster, the movement of customers to other clusters is examined by dividing the cluster into sub-clusters. Firstly, for further investigation of the third cluster, this cluster is hierarchically divided based on K-mean algorithm.

According to the Means of R, F, M, and V variables, the Mean-values will be considered as a basis for of sub-clusters comparison. Based on the results, the highest silhouette is 0.454, which is related to the clustering with 5 clusters. Therefore, the data of third cluster is divided into 5 clusters based on K-Mean algorithm (Table. 5).

Table 5. Sub-Clusters Information

Variables (Higher than Mean-Value)	Total Percent	Number of Items	Sub-Clusters
FM	2.6%	12	1-3
RFMV	21.5%	100	2-3
MV	12.7%	59	3-3
R	37.1%	173	4-3
-	26.2%	122	5-3

The final group includes low-value customers. This group doesn't have a desirable condition based on the indexes' values relative to the total value. The customers of this group have higher Recency, and lower Frequency and Monetary compared to the others. Moreover, the variety of purchased products is lower than, others.

Phase 5. Evaluation

After clustering and analysing the results, the differentiation of clusters created by K-mean algorithm and their resolution are evaluated through ANOVA test. As indicated in Table 6, significance levels (sig) of all variables (R (F=1.580, Sig= 0.000), F (F=790.748, Sig= 0.000), M (F=342.468, Sig= 0.000), and V (F=649.677, Sig= 0.000)) are lower than 0.05. Therefore, the homogeneity of populations' mean is rejected showing that clusters have different mean.

Table 6. ANOVA Test Results

Variable	Source of Change	SOS	df	Mean Square	F	Sig
R	Inter group	45.21	2	22.608	1.580	0.00
	Intra group	13.17	921	0.014		
	Total	58.39	923	-		
F	Inter group	24.23	2	13.117	790.75	0.00
	Intra group	15.27	921	0.017		
	Total	16.33	923	-		
M	Inter group	6.96	2	3.482	342.47	0.00
	Intra group	9.36	921	0.010		
	Total	16.33	923	-		
V	Inter group	20.82	2	10.410	649.67	0.00
	Intra group	14.75	921	0.016		
	Total	35.57	923	-		

After determining the incompatibility between variables, Post-Hoc Duncan test is used to ensure that clusters are distinct. In this test, if the variables do not have a significant distinction in the cluster, the similar values are placed in a sub-group. In other words, if we consider three clusters, the output of Duncan test should have three columns for each variable. The results of each variable are indicated in table 7. According to the results, the average values in all three clusters are quite distinct.

Table 7. Distinguishing Variable between Clusters

Cluster	variable	Number of Items	Clusters' Mean 1	2	3
1	R	181	0.803		
	F				0.512
	M				0.277
	V				0.471
3	R	466		0.160	
	F			0.147	
	M			0.069	
	V			0.117	
2	R	277			0.616
	F		0.038		
	M		0.044		
	V		0.065		

Phase 6. Deployment

After reviewing the results and valuation, the model is assessed. If the results are consistent with the primary targets of business, and seem to satisfy business needs, they will be used in a real environment. A new phase will be defined otherwisw, and the process will be repeated.

- **RFM and RFMV Comparison**

In this part, the results of the RFM and RFMV are compared with each other. Based on comparative analysis test, weights of R, F and M are equal to 0.088, 0.139 and 0.733 respectively in RFM model, while they are equal to 0.054, 0.075, 0.636 and 0.236 for R, F, M and V respectively in RFMV model. The average value obtained for the clusters is indicated in Table 8.

Table 8. Clusters' Average Value OF RFM & RFMV

Variable	Model	Cluster 1	Cluster 2	Cluster 3
Average value of R (Recency)	RFM	0.06	0.08	0.17
	RFMV	0.004	0.032	0.009
Average value of F (Frequency)	RFM	0.03	0.53	0.13
	RFMV	0.038	0.003	0.011
Average value of M (Montery)	RFM	0.04	0.26	0.07
	RFMV	0.176	0.028	0.044
Average value of V (Variaty)	RFM	-	-	-
	RFMV	0.111	0.015	0.028
Cluster's Average Value	RFM	0.09	0.28	0.09
	RFMV	0.330	0.080	0.092

Based on the results, in valuation using the RFM model the value of clusters is not distinct. Moreover, the values of cluster 1 and 2 are equal, therefore they are not distinct. However, the clustering based on a RFMV model by adding a variety of products (V), three clusters have different values. Thus, valuation would be better by using variable V. On the other hand, the role of the variable V in clusters' detachment and valuation is well confirmed. The results indicate that compared with clustering with RFM model, using the variety of products along with other behavioral variables provides more accurate clustering for companies.

4. Discussion

With the aim of finding groups with more product diversity along with other behavioral variables, the clusters are discussed in this section:

Cluster 1 (↑↑↑↑) high-value customers are the most valuable clusters among our customers and have the best condition in terms of the variety of the products. Moreover, regarding other factors are at a good level. This cluster has the potential to buy more and apply incentive programs, and is an appropriate group to impose the strategies of up-selling and cross-selling. The organization can make loyal customers by identifying and targeting these customers. Moreover, company can increase the probability of sales success and profitability by offering more products to this group.

Cluster 2 (↓↓↓↓) low-value customers are not in a desirable situation in terms of the obtained indexes. The notable point is that more than half of the firm's clients are in this cluster. This group has the least indexes' value. Low variety of the products of this cluster indicates that the customers of this group have low interest in diversified purchases. Therefore, offering additional products to these customers has not significant impact on the company's sales. Moreover, low recency of the cluster indicates its low frequency. This means that customers of this group for did not interact with the company for a long time. The remoteness of the firm and the high recently can be one of the reasons for turning customer away from the firm.

Cluster 3 (↓↓↓↑) middle-value customers have an average value in terms of indexes. The recency of this group is less than average, which means the continuation of the customer relationship with the firm. Thus, with increasing volume and variety of products of this group, the monetary will increase. This group has a potentiality to become high-value customers through incentive programs and increasing the variety of products. The low average of this cluster in terms of monetary, frequency, and variety of products could be too risky. Low recency indicates that customers haven't interacted with the company for a long time. This will lead to the customer's defection. Since the cost of attracting new customers is high, company should provide appropriate strategies to keep its recency in a suitable level and to maintain customers, and then to increase the customers' value.

The sub-cluster of 3-2 (↑↑↑↑) has a higher average among other sub clusters of third cluster. Due to the high variety of products, this group has the potential to migrate to a higher level in the first main cluster. The group consists of 21.5% of total customers of third cluster.

Sub-cluster 3-5 (↓↓↓↓) is the most critical sub-cluster among others. The Average values of all the variables in this group are lower than total average values of the third cluster.

This group is on the verge of moving to the second main cluster, and then leaving the firm, and hence their disconnection. This group includes 26.2 % of total customers of third cluster that is a significant digit compared with other clusters.

Sub-clusters 3-1 (↓↑↑↓) and sub-cluster 3-3 (↓↓↑↑) are complementary. The organization could increase the variety of products and monetary factors of these groups to move into the sub-cluster 2-3, and then move to the first cluster of general classification by increasing the potential value of both groups using up-selling and cross-selling strategies. These two groups include about 15 percent of the third cluster's customers.

Sub-cluster 3-4 (↑↓↓↓) has kept its relationship with the firm, but their low volume of transactions is noticeable. The high recency of this group compared with the total mean is the only positive point of the group.

5. Conclusion

Social and economic changes between organizations and customers, have turned the transactional marketing to relational marketing. Although purchasing goods and services shape the nature of communication in organizations and businesses, few studies have considered the importance of this item. So, the nature of products is the missing link of researches in the field of customer behavior. In this study, the variety of products along with behavioral variables of RFM model was considered to identify the high-value customers. This would help organizations to find the best customers to implement the strategies of up-selling and cross-selling. Moreover, they can increase the customer buying domain through offering diverse products. The subsidiary purpose of this study is to compare the function of K-Mean algorithm on the data. In this study, the accurate identification of customers' groups as well as behavioral and demographic data was obtained using the RFM model besides adding the variable of variety of products to this model. Therefore, organizations can adopt the appropriate strategies in each category, and can increase their Probability of success. As the results indicated, the recency and frequency values in cluster 1 and 3 are pretty close together, but the significant differences in the values of variables F and R make a huge difference in CLV of these clusters. According to the results, variable V has the significant role in the distinction and valuation of clusters. On the other hand, the valuation using RFMV model is more distinctive than RFM model. Moreover, in clustering using the RFM model, despite the difference in variables' Mean, the average value of clusters 2 and 3 are equal, while are distinct in the RFMV model.

The results indicate that the variable V, variety of products, is effective in calculating customers' value. Moreover, from a practical perspective, it could have advantages such as accurate recognition of different groups, identifying subgroups to assess the movement of customers between groups, increased effectiveness of applied strategies, and increased profitability of organizations. As a whole, the results of modeling indicate that the variety of products along with other behavioral variables provide more accurate clustering than RFM model. It also could provide a more accurate understanding of customers' groups. This variable is important in providing campaigns and strategies that are directly related to the variety of products. Clustering based on RFMV model can specify the most distinctive and meaningful groups to apply up-selling and cross-selling strategies. In other words, the variety of products associated with each group can assist organizations to identify the target groups.

According to the results, the surveyed company that is B2B, can group customers in appropriate clusters by differentiating their characteristics.

Previous studies on the clustering customers using RFM model have only added one variable to the model. For more general information, to achieve more practical results, it is suggested that future studies add two or more variables to RFM model so to identify the variable with greater impact on understanding the customer. Furthermore, future studies can use methods such as Markov chain to study the movement of customers between groups. This will help organizations to analyze and evaluate the increase and decrease of customers in groups by combining and comparing the results. Moreover, it is suggested that future research expand the scope of study to various industries, and compare their results.

References

[1] 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, 2011, pp. 57-63.

[2] P T. Kord Asiabi,, and R. Tavoli, "A Review of Different Data Mining Techniques in Customer Segmentation." Journal of Advances in Computer Research. Vol. 6, No.3, 2015, pp. 51-63.

[3] V. L. Miguéis, D. Van den Poel, A. S. Camanho, and J. F. e Cunha, “Modeling partial customer churn: On the value of first product-category purchase sequences”. Expert systems with applications, Vol. 39, No.12, 2012, pp.11250-11256.

[4] D. A. Kandeil, A. A. Saad, and S. M. Youssef, “A Two-Phase Clustering Analysis for B2B Customer Segmentation”. In: Intelligent Networking and Collaborative Systems (INCoS), 2014 International Conference on. IEEE, 2014. pp. 221-228.

[5] C. Liu, “Customer Segmentation and Evaluation Based on RFM, Cross-Selling and Customer Loyalty”. In: 2011 International Conference on Management and Service Science. 2011.

[6] E. S. Margianti, R. Refianti, A. B. Mutiara, and K. Nuzulina, “Affinity Propagation and Rfm-Model for CRM's Data Analysis". Journal of Theoretical and Applied Information Technology, Vol. 84, No. 2, 2016, p. 272.

[7] H. Hwang, T. Jung, and E. Suh, “An LTV model and customer segmentation based on customer value: a case study on the wireless telecommunication industry”. Expert systems with applications, Vol. 26, No. 2, 2004, pp. 181-188.

[8] C. F. Tsai, Y. H. Hu, and Y. H. Lu, “Customer segmentation issues and strategies for an automobile dealership with two clustering techniques”. Expert Systems, Vol. 32, No.1, 2015, pp. 65-76.

[9] W. Wang, and S. Fan, “Application of Data Mining Technique in Customer Segmentation of Shipping Enterprises”. In 2010 2nd International Workshop on Database Technology and Applications, on. IEEE. 2010, November. pp. 1-4.

[10] K. N. Lemon, and T. Mark, “Customer lifetime value as the basis of customer segmentation: Issues and challenges”. Journal of Relationship Marketing, Vol. 5, No. 2-3, 2006, pp.55-69.

[11] B. Noori, “An Analysis of Mobile Banking User Behavior Using Customer Segmentation”. International Journal of Global Business, Vol. 8, No.2, 2015, p. 55.

[12] Y. L. Chen, M. H. Kuo, S. Y. Wu, and K. Tang, “Discovering recency, frequency, and monetary (RFM) sequential patterns from customers’ purchasing data”. Electronic Commerce Research and Applications, Vol. 8, No. 5, 2009, pp. 241-251.

[13] 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, No. 3, 2009, pp.4176-4184.

[14] J. T. Wei, S. Y. Lin, C. C. Weng, and H. H. Wu, “A case study of applying LRFM model in market segmentation of a children’s dental clinic”.Expert Systems with Applications, Vol. 39, No. 5, 2012, pp. 5529-5533.

[15] A. Asllani, and D. Halstead, “A Multi-Objective Optimization Approach Using the RFM Model in Direct Marketing”. Academy of Marketing Studies Journal, Vol. 19, No. 3, 2015, p. 49.

[16] S. M. S. 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, No. 7, 2010, pp. 5259-5264.

[17] R. Venkatesan, and V. Kumar, “A customer lifetime value framework for customer selection and resource allocation strategy”. Journal of marketing, Vol.68, No. 4, 2004, pp.106-125.

[18] M. Alvandi, S. Fazli, and F. S. Abdoli, “K-Mean clustering method for analysis customer lifetime value with LRFM relationship model in banking services”. International Research Journal of Applied and Basic Sciences, Vol. 3, No. 11, 2012, pp. 2294-2302.

[19] W. J. Reinartz, and V. Kumar, “On the profitability of long-life customers in a noncontractual setting: An empirical investigation and implications for marketing”. Journal of marketing, Vol. 64, No. 4, 2000, pp. 17-35.

[20] R. A. Soeini, and E. Fathalizade, “Customer Segmentation based on Modified RFM Model in the Insurance Industry”, Proceedings of 4th International Conference on Machine Learning and Computing IPCSIT. Singapore: IACSIT Press. 2012, Vol. 25, pp. 104-108.

[21] A. Ansari, and A. Riasi, “Customer Clustering Using a Combination of Fuzzy C-Means and Genetic Algorithms”. International Journal of Business and Management, Vol. 11, No. 7, 2016, p.59.

[22] H. C. Chang & H. P. Tsai, “Group RFM analysis as a novel framework to discover better customer consumption behavior”. Expert Systems with Applications, Vol. 38, No. 12, 2011, pp. 14499-14513.

[23] D. T. Larose, Data mining methods & models. John Wiley & Sons, 2006.

[24] S. Moro, R. Laureano, & P. Cortez, “Using data mining for bank direct marketing: An application of the crisp-dm methodology”. In Proceedings of European Simulation and Modelling Conference-ESM, 2011, (pp. 117-121). Eurosis.

[25] P. N. Tan, Introduction to data mining. Pearson Education India, 2006.

Saman Qadaki Moghaddam holds his master’s degree in Information Technology Engineering from Islamic Azad University of Qazvin, Qazvin, Iran. He is interested in research in the field of information systems & technology, including innovation in ICT, and Electronic Commerce.

Neda Abdolvand is an assistant professor in Alzahra University and was a postdoctoral research fellow at Tarbiat Modares University. She holds her PhD and MS in information technology from Tarbiat Modares University and a post graduate certificate in information systems from Melbourne University. She is interested in research in the field of information systems & technology, including innovation in ICT, business intelligence, Electronic Commerce, and big data analytics. She has been a member of the Association for Information Systems since 2009.

Saeedeh Rajaee Harandi holds her master’s degree in Information Technology Management from Alzahra University, Tehran, Iran. She is interested in research in the field of information systems & technology, including innovation in ICT, electronic commerce, and business intelligence.