A Segmentation Framework for Airlines' Passengers Based on their Value Using Data Mining

[1]Mohsen Sadegh Amal Nik and [2]Hamed Manshadian
[1]Department of Industrial Engineering, University of Tehran, Tehran, Iran
[2]Sharif University of Technology, Tehran, Iran

Abstract: Since, air transport organizations are associated with customers directly, implementation of an accurate and appropriate customer relationship management would seem essential for them. Segmenting customers based on their value is one of the issues which is challenging in this field nowadays and many customer-oriented organizations are interested in calculating customer value and segmenting customers based on their value so that they can have a true understanding of customers. In this study, a framework is proposed to determine customer value regarding airline transport organizations considering their lifetime and potential value. Moreover, some strategies are discussed for segmenting them using data mining. A criterion is suggested for obtaining customer value considering behavioral nature of airline passengers and their traveling motives which are named business and leisure incentive in literature as well as customers' previous purchases behavioral nature using LRFM Model. For weighing parameters of both models AHP technique is implemented. After that, k-mean algorithm is used to cluster customers based on their value. Eventually after coming up with a conceptual model for segmenting customers based on their lifetime and potential value, some customer attraction and retention strategies are proposed.

Key words: Airlines passengers, customer lifetime value, customer potential value, customer segmentation, data mining

INTRODUCTION

Nowadays organizations such as air transport organizations need correct understanding of customers and their needs in order to succeed in business. Customer relationship management is an efficient tool for maintaining and increasing the customers' satisfaction which leads to customers' loyalty. Air transport organizations should acquaint themselves with their customers and their expectations and know what kinds of service packages are in line with customer's demands and how they have to provide them. Some air transport organizations haven't still segmented their customers, so they are unable to identify those few customers who provide profit for the organization. Ergo, proposing a framework to recognize airline passengers looks necessary.

CUSTOMER RELATIONSHIP MANAGEMENT

There isn't any agreement on a unified definition for customer relationship management yet. Considering lack of a unified definition for customer relationship management can lead to unwanted and unpredictable outputs. Rababah proposed a unified definition for CRM considering the existing ones. He has investigated definitions from three different perspectives, namely philosophy, strategy and technology. Then, he has observed that roughly 50% of definitions are from strategic view, 22% from technologic view and 17% from philosophic view. Finally, according to overlapping and existing relationship among these three perspectives, he has proposed the following definition:

"CRM is the building of a customer-oriented culture by which a strategy is created for acquiring, enhancing the profitability of and retaining customers that is enabled by an IT application for achieving mutual benefits for both the organization and the customers".

Albadvi proposed a conceptual model for customer relationship management in airline companies. In his model, he has collected different aspects of customer relationship management that is operational CRM, analytical CRM and collaborative CRM with databases and customer relationship channels in technology process aspect and has combined it with human features of CRM process in airlines companies. Of course

Corresponding Author: Mohsen Sadegh Amal Nik, Department of Industrial Engineering, University of Tehran, Tehran, Iran

implementing the CRM requires some fields and before implementing the relevant systems to it, it should be studied to avoid incur losses. For example, Sarjoughian and Ansari proposed a method for feasibility of implementing CRM in an aircraft-manufacturing company.

CUSTOMER VALUE

There are different definitions and concepts for customer value. In recent years, the researchers have tried to develop the tools and processes for measuring and managing customer value (Day and Crask, 2000). There are three general views about customer value. The first one is organization value from customer's view (Pechlaner *et al.*, 2002). For example, Jensen and Hansen (2007) proposed a model to obtain restaurant value from its customers, standpoint. Boetsch *et al.* (2011) presented a framework for valuating air transport organizations from customer's view and studied the effects of customer value on adventing air transport organizations' brand. The second one which is under discussion in the present study is customer value from organization's view. That is how much valuable the customer is for the organization and it is known as customer lifetime value. Third view is a combination of both the first and the second views. In fact the value of relationship among the organization, customer and provider is assessed (Ravald and Gronroos, 1996).

Verhoef and Lemon (2013) proposed six key lessons for a successful customer value management. By presenting a conceptual model about effective factors on changing demand in airline industry, Bieger *et al.* (2007) concentrated on customer value concept and its oposition in proposed model. Pease divided the customer value management into three main parts: right customer, right relationship and right maintainance. From his vantage point, right relationship is to maximize lifetime value. Similar models were proposed to gain and maximize lifetime value. However, right customer aspect has been rather neglected. In this study, this aspect has been added to the model as a main aspect regarding customer potential value.

CUSTOMER LIFETIME VALUE (CLV)

Customer lifetime value is a fundamental concept of CRM. According to a definition by Kotler *et al.* (2002), customers lifetime value is the art of marketing, attracting and keeping the profitable customer. They have defined a profitable customer as follow: a person, family or company whose obtained income during the time is more than the costs of company for attracting, selling and presenting the services to the customer. In literature, CLV is presented in other forms such as customer equity and customer profitability. Berger and Nasr have presented the following definition for CLV:

Customer lifetime value for a company is obtained net income of customer from some transactions which have been done during customer's lifetime minus the cost of attracting, selling and presenting the services to the customer considering the money time value.

Kotler (1974) defined customer lifetime value as a achievable net present value in customer lifetime. The companies may use CLV criteria in different methods. Verhoef and Lemon (2013) identified 5 different applications of CLV in companies:

- As a criterion for guiding toward culture which is more customer-oriented
- As a criterion for evaluating the marketing campaigns and investments
- As a based valuating tool for customer
- As a criterion for customer segmentation and resource allocation
- As an additional criterion in customer database

Rust *et al.* (2004) drew the investment on marketing in a conceptual model which led to an increased value from customer's view and as a result, attraction and retention rate enhanced and customer value amplified.

LENGTH-RECENCY-FREQUENCY-MONETARY (LRFM) ANALYSIS

LRFM analysis has been used for 50 years by direct marketers to target a subset of their customers in order to save mailing costs and increase profit (Kohavi and Parekh, 2004). However, in spite of oldness of this method it is still one of the most useful one in literature and according to the analysis by McCarty and Hastak (2007), it is the second common method in literature. Razmi and Ghanbari proposed a model for calculating customer lifetime value. After calculating the CLV they differentiated customers by their loyalty using mathematic calculation method and after that they estimated the probability of customer's future purchase by combining the parameters of RFM and ROI Models.

According to the view by Chang and Tsay (2004) and Reinartz and Kumar (2000), RFM can't differentiate the customers with long-term relationship and the customers with short-term relationship with organization. They proposed the idea of customer relationship length in their research and investigated its effect on customer loyalty and profitability. They expressed that increasing customer

relationship length will improve customer loyalty and then they defined "L" variable which indicates the time interval between the first and the last purchase of the customer. RFM Model defines the customers who create high financial value for company recently and in short-term have purchase frequency more than average purchase frequency as valuable customer while the factor of relationship length with company has been ignored. Therefore, the customer relationship length is added to RFM Model and this model will be promoted to LRFM. The results by Bhattacharya (1998) study show that the time length during which a person can be introduced as organization's customer is in positive correlation with times that customer tends to continue its relation with organization during it. Also, he has expressed that the customer relationships length with organization has positive correlation with the probability of his or her relationship stability in future. On the other hand, other basis has been found in resource advantage theory to follow customers. According to the Hunt and Morgan (1995) studies, the company resources are financial, physical, human-made organizational, informative and relational. Consistent with their studies it has been shown that constant relationship with customers can be utilized as a resource of organization which can be utilized as a competitive advantage. As said by Chang and Tsay (2004) with adding L index in RFM Model, the field for more exact analysis is provided. They suggest a matrix as the value matrix with aspects of purchase frequency and monetary value. They also have claimed that longer relationship of customer, higher loyalty and shorter recency show higher loyalty of customer. Two other indexes, customer relationship length and recent transaction time are defined as customer loyalty matrix (Fig. 1).

Sohrabi discussed a model for gaining customers long-term value in banking industry. For weighing the parameters of RFM Model, they used AHP technique. Jiale and Huiying discussed more different views of RFM. To do so, they considered M or monetary value as the customer current value and called R and F parameters as the customer potential value and calculated current value with mathematical method and the potential value using AHP. It is worth mentioning that researchers have also tried to present other techniques as a complement for RFM. For example, considering two variables of D (Discount and price) and R (Return times), Chiang (2011) presented RFMDR Model for online purchase industries. Moslehi presented a model for segmenting customers according to their lifetime value. They designed a general six-phase CRISP Method and clustered the customers after determining the measures of LRFM using AHP.
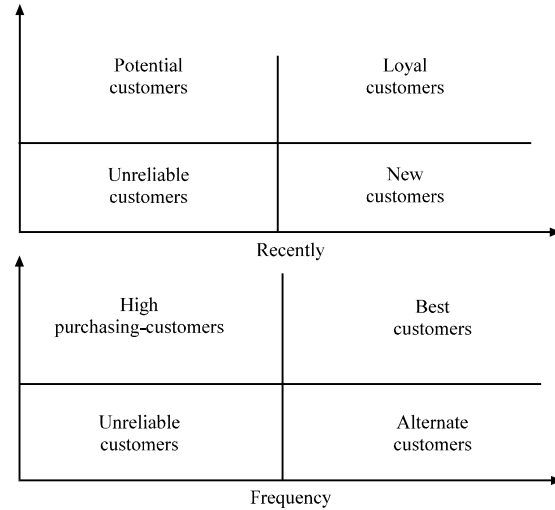

Fig. 1: Value and customer loyalty matrixes (Chang and Tsay, 2004)

Bradaran and Farrokhi added the variable of transaction days sequence to the RFM model and they converted it to RFMC Model and increased the model's accuracy by 5.5%.

SEGMENTATION

Undoubtedly, segmenting customers is considered as one of the most important issues in customer relationship management. Park and Baik (2006) called customer segmentation as the most important element in their proposed Client Request Filter (CRF) Model. As said by Kim *et al.* (2006), segmenting customers using their lifetime value is done with three general states:

- Segmentation using LTV measures only
- Segmentation using LTV elements
- Segmentation using LTV measures and other information
- In third state, LTV is usually placed in an axis of multi-dimensional segmentation space

Han *et al.* (2012) segmented the telecommunication industries' customers according to their value by decision tree model and have used AHP technique for weighing model's variables. Rajagopal divided the customers into four groups; the first group which includes the most valuable customers of organization who cause high income have low cost for it. The second group have high income and cost for organization. The third group includes the customers who have low income and cost for organization and finally, the fourth group or the least valuable customers for organization who have negative value for it that is low income and high cost.

DATA MINING

In a simple definition, data mining is the process of exploring useful information from massive and large data sources. Classifying, clustering, estimating, correlation, Regression, prediction, sequence analysis, deviation and indexing analysis can be named as data mining functions. As stated in literature, data mining is one of the most suitable tools for customer clustering. Huang *et al.* (2009) segmented customers based on their value using k-means, fuzzy c-means and bagged clustering algorithm they concluded that k-means technique is very convergent and it is very suitable for large data collection and it can be seen as a Fundamental Clustering Method. Liang (2010) used RFM technique for clustering customers and he applied k-means clustering and SOM technique for automobile industry and considered the k-means technique as the best one. Hu *et al.* (2014) clustered air transport customers into 8 groups according to RFM Model by k-mean clustering technique and afterward they have segmented customers into three final groups by advanced Apriori technique. By proposing a model for studying customers' behavior, Birant clustered customers of a sporting goods store regarding RFM Model using data mining. His model uses data mining technique for segmentation, prediction and recommendation related to CRM. Hashemi used data mining in order to identify different groups of banking customers' features. After determining the number of the optimized clusters, he used k-means technique for clustering customers using two step algorithm and afterward, he extracted the correlation rules of each cluster by using Apriori algorithm. Khajvand and Tarokh proposed a model for segmenting a private bank's customers using data mining and RFM Model and they classified the customers into 5 groups. They used k-mean method for clustering and have applied decision tree algorithm for classification. Kaffashpour divided the customers of a commercial company into 8 segments using RFM Model and k-means clustering method.

CONCEPTUAL MODEL

Considering discussed cases and urgent need of an organization for implementing a correct and comprehensive CRM System and explaining about customer value and general methods of customer segmentation, now we present a conceptual framework to determine airline's passenger value. In the following, essential cases will be proposed for presenting the framework and we will describe the model after explaining about parameters' weight calculation methods.

CALCULATING THE SI MODEL PARAMETERS FOR ESTIMATING THE CUSTOMER POTENTIAL VALUE

Regarding airline industries; there are two common reasons for traveling. Traveling due to business and traveling for leisure. Bieger *et al.* (2007) considered business and leisure as two main reasons of traveling and according to their researches in airline industry, 40% of air passengers travel due to business and 60% of them travel for leisure.

Considering their job, people may have numerous business trips. The probability of business travel usually is higher in jobs with high social status than the jobs with low social status. On the other hand in high-income classes, the probability of the demand for leisure travel is more than other income classes. Teichert *et al.* (2008) segmented the passengers regarding their travel reason and their travel class. They divided passengers into four general groups: the passengers who travel due to leisure and use economic class, the passengers who travel due to business but use economic class, the passengers who travel due to leisure but use business class and finally, the passengers who travel due to business and use business class. Looking at the personal aspect of customers for recognizing their potential value is a part of the suggested model of the present article as a result, customer value is acquired not only through financial and chronological view but also the personal aspect. In fact, the customer might have referred to the organization only once and in far past and perhaps he has not gained high profit for the organization financially. According to the LRFM Model such customer is not considered valuable customer and perhaps it is not essential to invest on them. However, if this customer has a high social status (e.g., specialist physician, lawyer and ambassador), the probability of his business travel will increase or if he has high income, the probability of his leisure traveling will increase. Therefore, regarding some factors this customer can be considered as a valuable customer. As we can see in different industries that in spite of existence of LRFM or RFM Models, the customer isn't considered as valuable customer but due to social position or high income, he/she is considered valuable and is so respected. The reason is considering the customer potential value. This occurs a lot in practice and the model that can consider this index is not observed in

Table 1: Classifying the jobs in terms of social status

Classes	Job
Excellent (5)	University professor, exarch, specialist physician, pilot, pharmacist, dentist, thepresident, university president, ambassador, governor, judge, industrialist, director of commercial company, electrical engineer, civil engineer, representative of parliament, minister, lawyer, priest or religious authority
Middle (4)	Fireman, elementary school tutor, general physician, chaplain, accountant, vet, goldsmith store, jewelry store, notary public, poet, printing and appearance laboratory, designer or graphic artist, photography, carpets wholesales, book store, computers goods store, confectionary, staff of medical laboratory, turning workshop, director, flower shop, furniture store, TV presenter, Hotel manager, sports coach, aviculture, car mechanics, musician, flight attendant, Car Expo, author, actor, painter artist
Good(3)	Hair salons, grocery, mason, real estate agency, postman, pizza store, typist, infusion therapy, servant of mosque and shrine, laundry, tailoring shop, newsstand, agency driver, taxi driver, narrator of religious tragedies, sandwich store, barbershop, auto dent removal, bookbinding, store dealer, cameraman, locksmiths, typical employee, security staff, rental of dishes, tables and chairs, plumber, secretary, bakery, carpentry, building painter
Low (2)	Junk shop, house servant, simple laborer, sweeper
Very low (1)	Brokerage, cigarette sale, bootblack, coupon retail, vendor

Table 2: Classifying the jobs in terms of social status

General classes	Detailed classes	Studied jobs
Very high	1	Ambassador, minister, executive manager of large company, University Professor, large industrialist, physician, representative of parliament, judge
	2	Exarch, pilot, senior officials in government ministries, lawyer, engineer
High	3	Professional footballer, cinema actor, jewelry seller, professional wrestler, musician, singer, architect, professional martial artist (karate, taekwondo and boxing players), the owner of the car firm
	4	Governmental office manager, oil company employee, small industrialist, market merchant, journalist, insurance company employee, bank employee
Middle to high	5	Constable, guard, police officer
	6	High school teacher, offices employee (Martyr Foundation, welfare organization, Red Crescent, Relief Committee, Documents registration), elementary school tutor
Middle to low	7	Craftsmen (carpenter, blacksmith, welder, electrician, turner), tradesman (grocery, apothecary butcher, dealer in junk shop), tailor
	8	Farmer, rancher, taxi driver, baker, plumber, barbershop, shoemaker
Low	9	Postman, artisans (broker, coffee house keeper, sweeper, vendor, bootblack), telephone operator, guardian, laborer

reviewed literature. Now this question is raised that how we can insert social and economic status in the framework?

There are different ways to classify jobs base on their social status. Bergman and Joye compared six of the most common social stratification methods in literature. Moghaddas compared the statuses of 97 different jobs in Shiraz, Iran with the statuses of those jobs in other societies. Chavashian and Sotoudeh Choubri clustered 95 jobs according to the position in Rasht, Iran, using fuzzy technique and they scored each job with very low, low, middle, good, very good scores (1-5) in terms of social status (Table 1).

Alizadeh and Rezaee scored 50 different jobs and then in terms of social status they divided them into 5 main groups and 9 sub-groups (Table 2).

Hauser and Warren (1997) classified a vast spectrum of different jobs in terms of social status based on gender using a new scoring method based on Nakao and Treas classifying technique. With reviewing related literature, it can be perceived that social position is assessed from the public people's view of that region and regarding their culture it may vary depending on their living area. For instance, Hauser and Warren (1997) considered the middle

score of social status, 53.19 for pilot and 52.32 for fireman which are in the same class Based on Chavoshian and Sotoudeh Choubri research they are not grouped in the same class. Due to this in implementing the main model which is proposed in the following it is worth mentioning that the scores should be inserted in the model with respect to the accomplished research in the related area. Consequently, considering the customer job and social status index, its measure which is shown with S is obtained. Like what was mentioned about differences in social status of persons in different societies, variances in economic status classifications are observable. Especially about this case, regarding the countries' currency and different value of goods in different countries, the diversity is more tangible. For instance, Table 3 shows 3 cases of household economic classes in the United States.

It should be regarded that the number of classes may vary depending on the society. Ergo, regarding the society, the number related to the customer economic class should be considered in the model. In case of having 5 classes, number 1 is related to the lowest class and number 5 is related to the highest economic class. Consequently, the related measure which is shown with I in the following is obtained.

Table 3: Comparing 3 cases of household economic classes in United States

Gilbert		William and Hickey		Beeghley	
Class	General features	Class	General features	Class	General features
The capitalist class (1%)	High rank executive managers, high rank politicians	High class (1%)	High rank executive managers, famous persons with income >500000 dollars in a year	Very rich (0.9%)	Millionaires with income >350000 dollars in a year like famous persons and politicians and high rank managers
High middle class (15%)	High education usually with university degree, often recipients of salary, professionals and middle managers with high work independency	High middle class (15%)	High education (often university education), experts and managers with income >100000 dollars in a year	Rich (5%)	Families with asset value >1 million dollars that most of it is from the price of their house, usually with university education
Low middle (30%)	Semiprofessionals and crafts men with average standard of life, often with university degree and so-called white collar worker	Low middle class (32%)	Semi-professionals and craftsmen, some with independent jobs, the family income between 35000-75000 dollars in a year, usually with university degree	Middle class (46%)	Workers with university education with significant income class more than average, men's income is >57000 dollars in a year and women's income is >40000 dollars in a year
Worker class (30%)	Often blue collar workers with typical and routine work, the life standard depends on the numbers of persons who work in family but usually have enough income, often with high school education (Diploma)	Worker class (32%)	Pink collar and blue collar workers with low job security, yearly income of family is usually between 16000-30000 dollars in a year, high school education (Diploma)	Worker class (40-50%)	Blue collar workers with very typical job and low job security, the men's income is averagely 40000 dollars in year and women's income is 26000 dollars in a year
Poor employed (13%)	Inferior employees and some blue collar workers, lack of income security and with high poverty risk, some with high school education (Diploma)	Lowest class (14-20%)	They are engaged in activities with very low incomes, some of them with high school education (Diploma)	Poor class (12%)	The persons who are accounted below the poverty line and aren't very effective on work and employment of society, family income is about 18000 dollars in year, some of them with high school education (Diploma)
Poor class (12%)	With very limited work activity or without role, depending on the government transmission, some with high school education (Diploma)				

CALCULATION OF LRFM PARAMETERS

As it was mentioned, RFM Method is for scoring the customers according to their purchase behavior. RFM technique is an analytic technique which has been used for years. In recent years, more complicated methods and models have been developed but RFM is still used due to its feasibility and efficiency and as said by McCarty and Hastak (2007), it is still the basis of scoring customers. Regarding to the studied industry that is airline industry; the time unit of transaction is one day. The measure of delay is equal to the number of days which has lasted since the last customer's purchase. The number of transactions of each customer is equal to the number of days that the customer has purchased from organization. The average financial transaction of customer is equal to the measure of profit that the customer has gained for the organization. Customer relationship length is equal to the numbers of days which have been elapsed since the first purchase till the last one.

NORMALIZING PARAMETERS

To begin with parameters' normalization should be calculated. For this purpose, the following formulas should be used. For normalizing L, F, M, S and I variables, (Eq. 1) is used:

$$L_i^N = \frac{L_i - L^s}{L^l - L^s}, \quad F_i^N = \frac{F_i - F^s}{F^l - F^s}, \quad M_i^N = \frac{M_i - M^s}{M^l - M^s},$$
$$S_i^N = \frac{S_i - S^s}{S^l - S^s}, \quad I_i^N = \frac{I_i - I^s}{I^l - I^s} \tag{1}$$

Due to negative effect of R-index on calculation of customer lifetime value (Eq. 2) is used for normalization:

$$R_i^N = \frac{R^s - R_i}{R^l - R^s} \tag{2}$$

That we have in mentioned formulas:

- x^l: maximum value of Xs
- x^s: minimum value of Xs
- x_i^N: normalized value of x_i
- i: customer index

Normalized measures of L, F, M, S and I variables are calculated for each customer by using the mentioned equations.

CALCULATION OF PARAMETERS' WEIGHT

Parameter's weight are not necessarily identical. Some parameters may have more or less superiority over others. There are some methods for identifying the parameter's weight, like Entropy, Eigenvector, SMART, Paired Comparisons and Hierarchy Analysis Process Method. With regard to literature, the most common method for calculating the weight of the LRFM and RFM Model variables is AHP Method. AHP Method is a powerful technique and a flexible and multi criteria tool in order to make decisions in complicated situations. It considers two qualitative and quantitative concepts. This

Table 4: Paired comparisons matrix of LRFM variables

Variables	Frequency (F)	Monetary (M)	Recency (R)	Length (L)
Length (L)	1/5	1/3	3	1
Recency (R)	1/7	1/5	1	1/3
Monetary (M)	1/3	1	5	3
Frequency (F)	1	3	7	5

Table 5: Paired comparisons matrix of binary variables of S and I

Classes	Economic class (I)	Social class (S)
Social class (S)	1/3	1
economic class (I)	1	3

Table 6: Relative weight of variables

Variables	Income class (I)	Social class (S)	Frequency (F)	Monetary (M)	Recency (R)	Length (L)
Relative weight	0.75	0.25	0.5579	0.2634	0.0569	0.1218

process is one of the most practical techniques which was introduced by Thomas L. Saaty in 1970s. AHP technique links the personal judgments and values to each other by a logical method. This process forms the structure of hierarchy using imagination, knowledge and the experience of experts and causes desired judgments using experience and logic. For this purpose, the comparison of indexes with each other is usually used by numbers 1-9 where number 1 shows the equal importance and number 9 shows absolute superiority.

In this stage, some interviews were conducted with related experts and they announced their opinions about the weight of L, F, M, S and I variables (Table 4 and 5).

The consistency rate (R) of LRFM matrix is 0.065437 and as it is <0.1 it shows the consistency of variables. Consequently, the coefficients can be used to gain the customer value and use it in customer segmentation process. Table 6 shows relative weight of model's variables which have been gained by AHP Method.

WEIGHED MODEL APPLICATION FOR CUSTOMER SEGMENTATION

Due to variability of parameters' importance in different industries their weight may vary. Nowadays, segmentation techniques according to weighing is considered as one of the best branches of customer segmentation. Different methods for calculation of RFM and LRFM have been proposed. Alencar *et al.* (2006) combined RFM with Genetic algorithm and applied a new method for RFM calculation. Kim *et al.* (2006) combined RFM with decision tree and introduced a new segmentation method. Hsieh (2004) used Bayesian Weighing Method and neural networks to gain better results of RFM Model. In this study, AHP technique was used for weighing the RFM Model parameters whose

Table 7: Final weights which are used in model

WI_i	WS_i	WM_i	WF_i	WR_i	WI_i
$W_I \times I_i^N$	$W_S \times S_i^N$	$W_M \times M_i^N$	$W_F \times F_i^N$	$W_R \times R_i^N$	$W_L \times L_i^N$
$0.075 I_i^N$	$0.025 S_i^N$	$0.05579 M_i^N$	$0.2634 F_i^N$	$0.0569 R_i^N$	$0.1218 L_i^N$

amount were achieved formerly. Now we should multiply the calculated weights by their related variables to obtain weighed LRFM. The related measures are shown in Table 7.

CALCULATING CUSTOMER POTENTIAL VALUE

In previous parts the SI Model parameters' weight was gained. Equation 3 shows the way to achieve each customer potential value:

$$PV_i = WS_i + WI_i = 0.25\, S_i^N + 0.75 I_I^N \qquad (3)$$

CALCULATING CUSTOMER LIFETIME VALUE

The weighed measure of each of LRFM Model variables was calculated by the proposed method. To gain lifetime value of each customer it is enough to sum up their gained measures (Eq. 4):

$$LTV_i = WL_i + WR_i + WF_i + WM_i$$
$$= 0.1218\, L_i^N + 0.0569\, R_i^N + 0.2634\, F_i^N + 0.5579 \qquad (4)$$

CLUSTERING CUSTOMERS BASED ON THEIR VALUE

The purpose of clustering is to divide the existing data into some groups so that each cluster has the maximum possible difference with each other. Customers in each cluster are very similar. The duty of clustering is to divide different groups into a number of more similar groups or clusters. In other words, clustering divides an unordered society into a collection of ordered. One of the most important clustering algorithms is k-means. It is a Cluster Recognition Method which is very common in practice. The main stages of k-mean include the following cases:

- A: selecting k points as the clusters' central points
- B: creating a new part by determining each sample with centrality of the closest class
- C: calculating the centers of new class as main classes
- D: repeating B and C stages to gain a desired measure of criterion performance
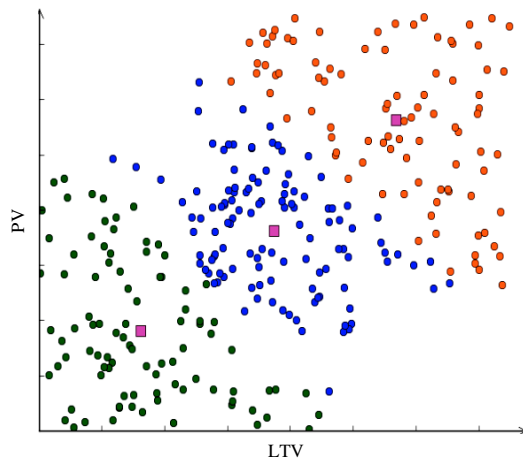
Fig. 2: Potential value and customer lifetime value in two-dimensional space

The clustering method after determining the potential value and lifetime value of each customer is in this manner that after determining each value that was already mentioned, we put each of them in an axis of two-dimensional space and then cluster them based on their values (Fig. 2).

PROPOSED FRAMEWORK

Conceptual framework of the implementation process of customer relationship management strategy according to the customer value is shown in Fig. 3.

Suggested CRM strategies for customer segments acquiring customer value and clustering them isn't profitable for the organization as long as it isn't use practically. Therefore, after recognizing each customer's cluster, essential decisions and suitable behavior with them should be made. Less valuable customers for the organization can be more costly than profitable. As a result, the organization can't adopt unique behavioral policy for all customers.

Considering the strategies like discount and reward can be one of the suitable strategies for loyal and valuable customers. The organization can suggest free travels for its more valuable customers or it can consider gift cards for them.

Paying attention strategies like congratulating customer's birthday via phone call, SMS, e-mail and so on can provoke a good feeling in their memory. Presenting special pre-flight in-flight and post-flight service can be another strategy for valuable customers. For example, pre-flight services can include free in-town transportation of the passenger to the airport. In-flight services can
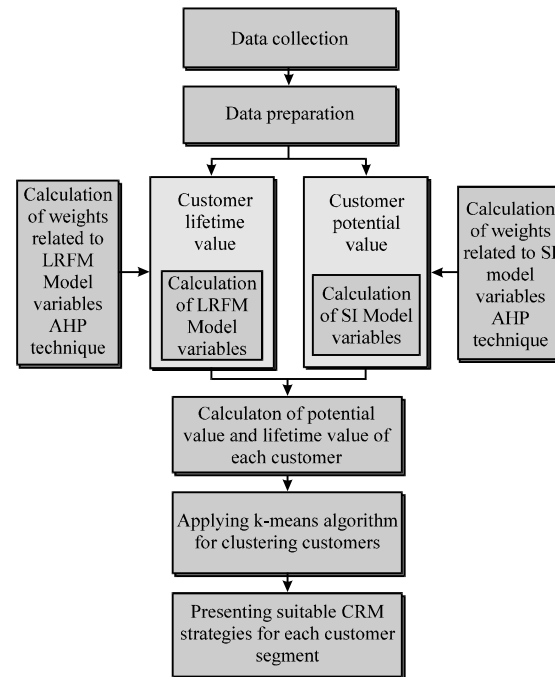

Fig. 3: General framework for clustering the airline passengers

include presenting the special and free reception services. Post-flight services can include free in-town transportation and hotel reservation. Anyhow, regarding the special policies of organization and considering the financial resources and existing facilities, special decisions can be made for attracting and keeping valuable customers.

CONCLUSION

In this study, a model was proposed for segmenting airline passengers based on their value. Considering related literature, different models have been suggested for segmenting customers in different industries each of the models have had their own weaknesses and strengths. Most of the suggested models focus on customer lifetime value, considering financial transactions, relationship length, recency and frequency of customer purchases. However, considering the nature of the industry studied in this study (i.e., air transport organizations), two factors of social status and income class were added to the model as the customer potential value. Finally considering the calculated values for each customer, all of them can be put in different clusters. This was modeled by means of data mining tool using k-mean Clustering algorithm. Eventually by implementing right and suitable customer management policies, we can be more successful in achieving customers' attraction and retention.

The present study provides new direction for further exploration in the future. For weighing the parameters, AHP technique and the experts, opinion were used. This technique is considered as one of the most applicable techniques in this domain. However, new models can be compared with the proposed model by designing a similar model and calculating the parameters' weight with other techniques like paired comparison analysis or neural networks technique. Furthermore, in this research k-mean method was used. Other methods like fuzzy C-mean and CRISP can be used for clustering and the findings can be compared with each other. Besides, the proposed model in this research can be scrutinized in similar industries such as hotel management and tourism industry whose customers have either a business or leisure incentive. However, the weight of parameters may vary. So, to obtaining the right weight, we should consult experts of related industries.

REFERENCES

Alencar, A.J., E.M. Ribeiro, A.L. Ferreira, E.A. Schmitz, P. Lima and F.S.P. Manso, 2006. Optimized RFV analysis. Marketing Intell. Plann., 24: 106-118.

Berger, P.D. and N.I. Nasr, 1998. Customer lifetime value: Marketing models and applications. J. Interact. Marketing, 12: 17-30.

Bhattacharya, C.B., 1998. When customers are members: Customer retention in paid membership contexts. J. Acad. Market. Sci., 26: 31-44.

Bieger, T., A. Wittmer and C. Laesser, 2007. What is driving the continued growth in demand for air travel? Customer value of air transport. J. Air Transp. Manage., 13: 31-36.

Boetsch, T., T. Bieger and A. Wittmer, 2011. A customer-value framework for analyzing airline services. Transp. J., 50: 251-270.

Chang, H. and S. Tsay, 2004. Integrating of SOM and K-mean in data mining clustering: An empirical study of CRM and profitability evaluation. J. Inf. Manage., 11: 161-203.

Chiang, W.Y., 2011. To mine association rules of customer values via a data mining procedure with improved model: An empirical case study. Expert Syst. Appl., 38: 1716-1722.

Day, E. and M.R. Crask, 2000. Value assessment: The antecedent of customer satisfaction. J. Consum. Satisfaction Dissatisfaction Complaining Behav., 13: 52-60.

Han, S.H., S.X. Lu and S.C. Leung, 2012. Segmentation of telecom customers based on customer value by decision tree model. Expert Syst. Appl., 39: 3964-3973.

Hauser, R.M. and J.R. Warren, 1997. Socioeconomic indexes for occupations: A review, update and critique. Sociological Methodology, 27: 177-298.

Hsieh, N.C., 2004. An integrated data mining and behavioral scoring model for analyzing bank customers. Expert Syst. Appl., 27: 623-633.

Hu, X., H. Zhang, X. Wu, J. Chen and Y. Xiao et al., 2014. A new customer segmentation framework based on biclustering analysis. J. Software, 9: 1359-1366.

Huang, S.C., E.C. Chang and H.H. Wu, 2009. A case study of applying data mining techniques in an outfitter's customer value analysis. Expert Syst. Appl., 36: 5909-5915.

Hunt, S.D. and R.M. Morgan, 1995. The comparative advantage theory of competition. J. Market., 59: 1-15.

Jensen, O. and K.V. Hansen, 2007. Consumer values among restaurant customers. Int. J. Hospitality Manag., 26: 603-622.

Kim, S.Y., T.S. Jung, E.H. Suh and H.S. Hwang, 2006. Customer segmentation and strategy development based on customer lifetime value: A case study. Expert Syst. Appl., 31: 101-107.

Kohavi, R. and R. Parekh, 2004. Visualizing RFM segmentation. Proceedings of the 4th SIAM International Conference on Data Mining, April 22-24, 2004, Florida, USA., pp: 391-399.

Kotler, P., 1974. Marketing during periods of shortage. J. Marketing, 38: 20-29.

Kotler, P., G. Armstrong, J. Saunders and V. Wang, 2002. Principles of Marketing. European Edn, Prentice-Hall, London..

Liang, Y.H., 2010. Integration of data mining technologies to analyze customer value for the automotive maintenance industry. Exp. Syst. Appl., 37: 7489-7496.

McCarty, J.A. and M. Hastak, 2007. Segmentation approaches in data-mining: A comparison of RFM, CHAID and logistic regression. J. Bus. Res., 60: 656-662.

Park, H.S. and D.K. Baik, 2006. A study for control of client value using cluster analysis. J. Network Comput. Appl., 29: 262-276.

Pechlaner, H., E. Smeral and K. Matzier, 2002. Customer value management as a determinant of the competitive position of tourism destinations. Tourism Rev., 57: 15-22.

Ravald, A. and C. Gronroos, 1996. The value concept and relationship marketing. Eur. J. Marketing, 30: 19-30.

Reinartz, W.J. and V. Kumar, 2000. On the profitability of long-life customers in a non-contractual setting: An empirical investigation and implications for marketing. J. Marketing, 64: 17-35.

Rust, R.T., K.N. Lemon and V.A. Zeithaml, 2004. Return on marketing: Using customer equity to focus marketing strategy. J. Marketing, 68: 109-127.

Teichert, T., E. Shehu and I. von Wartburg, 2008. Customer segmentation revisited: The case of the airline industry. Transp. Res. Part A Policy Pract., 42: 227-242.

Verhoef, P.C. and K.N. Lemon, 2013. Successful customer value management: Key lessons and emerging trends. Eur. Manage. J., 31: 1-15.