

“© 2017 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.”

# A Customer Segmentation Framework for Targeted Marketing in Telecommunication

Anahita Namvar

School of Industrial Engineering  
Iran University of Science and  
Technology  
Tehran, Iran

[Anahita.namvar@gmail.com](mailto:Anahita.namvar@gmail.com)

Mehdi Ghazanfari

School of Industrial Engineering  
Iran University of Science and  
Technology  
Tehran, Iran

[Mehdi@iust.ac.ir](mailto:Mehdi@iust.ac.ir)

Mohsen Naderpour

School of Systems, Management  
and Leadership  
University of Technology, Sydney  
Sydney, Australia

[Mohsen.Naderpour@uts.edu.au](mailto:Mohsen.Naderpour@uts.edu.au)

**Abstract**— Telecommunication industry is highly competitive, and mass marketing is not applicable anymore. Moreover, Mobile customers have different behaviors that urge telecom industries to differentiate their strategies to meet customers' needs. At the same time, mobile operators have an enormous amount of customer records, and data-driven approaches can help them to draw insights from this huge amount of data. Therefore, a data-driven segmentation approach can support marketing strategies to tailor their marketing plans. In this research, we adopt behavior and beneficial segmentation in a two-dimensional framework to segment customers. The results indicate that our method has an outstanding performance for customer segmentation. Moreover, we have recommended some marketing strategies based on each segment's behavior with the aim of increasing in Average Revenue Per User (ARPU) and decreasing in marketing expenses.

**Keywords:** *Customer Segmentation; Clustering; Telecommunication; Mobile Broadband*

## I. INTRODUCTION

Recent developments in the 4G network have led subscribers to consume a huge amount of data for watching videos online, surfing the Internet, and communicating with their friends[1]. This high-speed Internet has increased the data revenue of mobile service providers and, broadband data services become more important than traditional voice and SMS services. Also, Mobile Network portability which enables subscribers to migrate from a mobile service provider to another one, increases the competition between mobile network operators. Consequently, these trends are reshaping the telecommunication industry and they had to pay special attention to broadband data services to maintain a competitive advantage over other competitors.

Competition plays a significant role in telecommunication market, and when the number of competitors increases, mass marketing does not work anymore, so mobile service providers should customize their products and services for their customers based on their expectations. Market segmentation enables organizations to separate a heterogeneous market into homogenous groups of customers with distinct behavior [2]. Therefore, they can differentiate

their marketing strategies based on different customer characteristics to meet customer needs [3]. Customer segmentation with applying the value of machine learning algorithms can better lead organizations to define suitable marketing strategies [1].

This paper aims to develop a customer segmentation framework for segmenting telecom Internet users for targeted marketing to define a proper marketing strategy for each group. This article has been divided into four parts. The first section gives a literature review on customer segmentation in the mobile broadband area. The second part presents the segmentation framework, and our experimental results described in third part. Finally, the last section discusses the conclusion and some future research recommendations.

## II. LITERATURE REVIEW

This section has attempted to provide a summary of the literature relating to the Customer Relationship Management (CRM), Customer Segmentation and the application of Data Mining for targeted marketing in telecommunication industry. The following parts move on to describe greater details.

### A. CRM and Data Mining

The number of mobile subscribers has increased remarkably in a way that in mega cities the contribution of mobile users is more than 100 percent. It means that some users have more than one mobile number. On the other hand, the cost of finding new customers for Internet Service Providers (ISPs) is much more than the cost of retaining current customers. Because of this, organizations shift their focus from customer acquisition to customer retention. According to Bose et al. [4], 1% increasing the retention rate can bring 5% progress in the firm's value. Consequently, most of the ISPs try to take advantage of their customer's data in order to become familiar with their needs for customized marketing.

Telecommunication industry has the huge amount of customer records and activities such as Call, Data, and SMS. This information is known as CDRs, Call Detail Records, that gathered and keep in the CRM databases routinely. AT&T, a big telecom operator in the USA, is a great illustration of generating and saving a plethora of CDRs, around 300 million

Call Detail Records per day [5]. On the one hand, CRM analysis plays a crucial role in the sale figure growth and marketing cost reduction [6]. So telecom companies have to use customers' data to gain an insight into their clients' needs for providing new products and services [7]. At the same time, the size of CRM databases is growing sharply, therefore the powerful methods and knowledge-based expert systems required to extract knowledge from these huge datasets [5, 8, 9]. Data mining algorithms are applied to discover hidden customer behavior and users' needs from this huge amount of CRM records [6]. CDRs that are available in telecom databases are raw data, and Data mining can change them to meaningful information to solve business challenges and issues [7]. Data mining is the process of extracting hidden patterns, rules and behavior from a large amount of data [8, 10].

### B. Market Segmentation

The market segmentation definition was first introduced by Smith [11] in 1956. Customer segmentation is divided into four major categories: Lifestyle, Demographic, Behavioral, and Beneficial Segmentation [6]. Many researchers address variables related to these categories for market segmentation. For Lifestyle segmentation, the style of living is a major component, but finding a comprehensive pattern which is related to the lifestyle of a large number of customers, is difficult or maybe impossible. In demographic segmentation, geography has a pivotal role in this area; however, because of the market globalization and the progress in the field of information technology and e-Commerce the role of demographic features in customer segmentation is not applicable anymore. Behavioral segmentation is another type of segmentation which classifies customers based on their behavioral habits. Lack of behavioral data has existed as a general problem in this area but technological improvements in recent decades help industries to utilize a large amount of behavioral data to extract customer behavior patterns. Beneficial segmentation is another approach that segment customers according to the revenue that they generated for the organization. One important point is that behavioral and beneficial segmentations are not applicable exclusively [6], so historical habits of customers and their monetary value for the organization should be considered simultaneously. In order to achieve successful customer segmentation that can lead organizations to define suitable marketing strategies, machine learning algorithms are applied widely[1]. Segmenting customer employing data mining techniques is a challenging task which is confronted with two types of problems: 1) selecting or summarizing suitable features from a huge amount of raw transactional data. 2) Proposing an algorithm that has the best performance on huge dataset[12].

Various data mining techniques such as Decision Tree classification rules[1], Bayesian modeling[1], RFM analysis [13], K-Means clustering, K-medoids, Fuzzy Clustering [4], Hierarchical clustering, Self-Organization Map (SOM) [4, 14], Logistic regression, Support Vector Machine and metaheuristics methods have focused on segmenting customer in an efficient way[15]. For example, Dutta [16] grouped different data mining methods which adopted in market

segmentation field of study, such as RFM, K-means, hierarchical clustering. A great number of researchers considered RFM variables for customer segmentation [17-19], and loyalty based segmentation [14]. [20] proposed a support vector clustering algorithm in his study. In addition, a considerable amount of literature has been published on the application of data mining techniques for customer segmentation in telecommunication [3, 21-24].

### C. Clustering for Telecom

Clustering is one of the data mining techniques that has been widely used for customer segmentation, targeted marketing, and cross-selling. This method segments customers into similar groups and helps marketing strategist to target each group with tailored advertising and pricing campaigns to increase their satisfaction [4]. Han, et al. [25] defined clustering as a method that puts subjects with similar properties in one cluster, and it has been used widely by researchers for studying customer behavior. Chen, et al. [26] segmented mobile subscribers to find different payment behaviors of customer and focus on customers who had delayed payments. Wu and Chou [27] proposed a mixed class membership clustering approach to develop soft clustering method for segmenting online customers. Keralapura, et al. [28] offered a co-clustering method for customer behavior analyses by applying network data of a communication service provider in North America. Zhang, et al. [29] Applied hierarchical clustering and k-medoids in order to extract knowledge from sequential behavior patterns of mobile internet services.

Many studies suggested hybrid application of data mining methods to improve the segmentation result, Liao, et al. [30] segmented customer with Two-Step, and found some rule with Apriori algorithms in each cluster for product bundling. Furthermore, many investigations carried out on customer segmentation methods to notice changes in customer behavior over time. Bose and Chen [4] conducted FCM in different time intervals and presented changes in customer behavior by customer migration between clusters.

## III. SEGMENTATION MODEL

The framework which is presented in Fig. 1 segments mobile subscribers based on real customers historical data. The aim of this study is to determine marketing strategies to target mobile subscribers intelligently in a way that pushes the ARPU criteria to a higher level and achieves higher data revenue. Telecom operators must continually investigate their subscribers' behavior to meet their needs and provide new competitive advantages compared to their competitors. We will delve into each step of the proposed framework in the following.

### A. Data Preparation

Data pre-processing is a major task in data mining, and raw data should be cleaned and changed in a way that suitable for data mining tasks. Data preparation stage has a number of critical tasks: understanding source of error, data transformation, outlier detection, normalization, and feature selection.

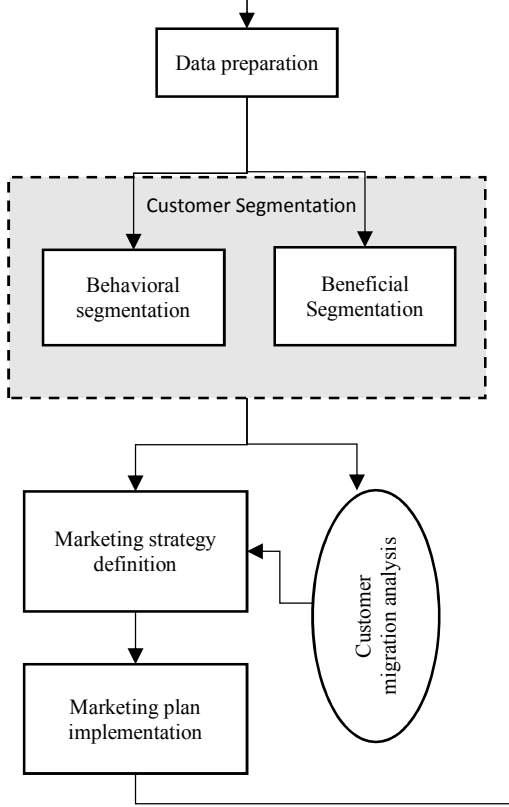


Figure 1. Customer segmentation framework

Telecom databases have a huge amount of raw data (CDRs) which generated every day, but they are not suitable for knowledge extraction. So, data transformation is essential to construct an appropriate data set [31]. One way is to derive statistical features such as mean, sum, min, max out of historical data (CDRs) which is used in this study. Another challenge in working with telecom data is that telecom database deals with many missing values and outliers. These outliers should be removed from our data set to have a reliable segmentation result. We applied SOM clustering (Kohonen) for removing outliers. Moreover, based on telecom marketing experts' advice records of subscribers who their total data usage per month is less than 15MB were removed from our dataset. We also used decision tree algorithm for predicting some of the missing values.

After removing outliers, we normalized our dataset to prevent the overall magnitude effect [12] and to provide appropriate cluster results. All selected features are normalized using Equation (1) suggested by [32].

$$X_n = 2 \left[ \frac{X - X_{Min}}{X_{Max} - X_{Min}} - 1 \right]. \quad (1)$$

where  $X_n$  represents the normalized value and  $X$  is original value.  $X_{Min}$  and  $X_{Max}$  refer to minimum and maximum value of the variables accordingly.

We utilized Pearson correlation analysis in order to choose appropriate features for clustering purposes. The Pearson coefficient equation is presented as follows:

$$\rho_{X,Y} = \frac{Cov(X,Y)}{\sigma_X \sigma_Y} \quad (2)$$

where  $Cov(X,Y)$  is the covariance between  $X$  and  $Y$  variables, and  $\sigma_X$  and  $\sigma_Y$  are the standard deviations of variables  $X$  and  $Y$  accordingly.

Variables that have a high positive correlation with revenue based features should be considered in clustering model. Table I illustrates features addressed in correlation analysis.

TABLE I. CORRELATION MATRIX

	Data Revenue	Age of Network	Purchased package count	Purchased Volume	Usage
Data Revenue	1	0.111	<b>0.541</b>	<b>0.615</b>	<b>0.492</b>
Age of Network	0.111	1	0.084	0.045	0.066
Purchased package count	0.541	0.084	1	0.075	0.095
Purchased Volume	0.615	0.045	0.075	1	<b>0.82</b>
Usage	0.492	0.066	0.095	<b>0.82</b>	1

The Pearson correlation coefficient that measures dependency between two variables ranges from -1 to 1. positive correlation coefficient between two variables means if one increases the other one also increases and vice versa. For clustering purpose, we choose features that their dependence on revenue based feature is more than 0.4. Furthermore, high correlation coefficient, 0.82, between usage and purchased volume presents high dependency of both variables, so considering one of them for clustering purpose would be enough, we choose usage based attributed for our model.

### B. Customer Segmentation

In this study, we segment customers based on their beneficial and behavioral characteristics. Instead of considering all features in one stage and perform segmentation we address beneficial and behavioral segmentation separately, then placing both results in a two-dimensional matrix to construct the final segmentation.

#### 1) Behavioral segmentation

To begin this process, we apply usage-based features to segment customers' data with K-Means Clustering method. One of the big challenges of K-Means algorithm is that the number of clusters should be determined in advance before undertaking the segmentation. Following the VRC method, noted by [33] and performed well in many situations [34], we find the optimal number of  $K$  for K-Means clustering algorithm. For  $n$  objects and  $k$  segments The VRC equation is presented as follows:

$$VRC_K = \frac{(SS_B/(K-1))}{(SS_W/(N-K))} \quad (3)$$

where  $n$  is the total number of objects and  $K$  is the number of segments.  $SS_B$  and  $SS_W$  are the sum of the squares between the segments and the sum of the squares within the segments, respectively. For finding a value for  $k$  that minimizes the value of  $W_k$ , we could obtain appropriate number of  $K$  to feed into our model. The  $W_k$  equation is given as following.

$$W_K = (VRC_{K+1} - VRC_K) - (VRC_K - VRC_{K-1}) \quad (4)$$

The drawback of this criterion is that the minimum number of clusters can be determined as three. However, it does not affect our result while our input data is huge and we do expect to have customers in more than two clusters.

### 2) Beneficial segmentation

Our beneficial segmentation is based on customers' payment for data services over three months. We define a rule to segment customers according to data revenue characteristics.

### 3) Final segmentation

The matrix  $M$  with  $m$  rows as behavioral segments and  $n$  columns for beneficial groups presents the final segmentation result which is depicted in below.

$$M = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{mn} & \cdots & a_{mn} \end{bmatrix}$$

Where  $a_{ij}$  represents a group of customers that they belong to  $i^{\text{th}}$  behavioral segment and  $j^{\text{th}}$  beneficial segment. While  $i$  in  $\{1, 2, \dots, m\}$  and  $j$  in  $\{1, 2, \dots, n\}$ .

### C. Customer migration Analysis

Customers' behaviors change over time; mobile service providers should detect these fluctuations and respond effectively by defining appropriate marketing plans to shift customers into desirable segments. Our segmentation models could help marketers to track customer migration between clusters over multiple stages by repeating the model.

### D. Defining and implementing marketing strategies

The model output which is the segments of subscribers who have similar behavior can help marketing strategists to either develop new marketing plans or adjust their current strategies. Definitely, this data-driven customer segmentation approach can help them to meet customers' needs and increase the market share.

Managers should put emphasis on implementing defined strategies precisely, also monitoring the results dynamically. Otherwise, segmentation approach cannot lead to better business performance.

## IV. EXPERIMENTAL RESULTS

In this section, we assess the explained model to segments mobile broadband customers for targeted marketing with the aim of growth in ARPU.

We analyze three-month historical data of anonymized coded dataset, for protecting customer and company privacy. In preparation for analysis, the dataset is first cleaned as mentioned in the previous section; then features are constructed, normalized and selected for our segmentation model. As we mentioned before, adopting both behavioral and beneficial segmentation, we can have better segmentation results, therefore in this research, we consider both behavioral and beneficial segmentation.

### A. Behavioral segmentation

According to [35], K-Means clustering is sensitive to high variance dataset and tend to bias the clustering to larger values. Therefore, we divide our dataset into three major groups based on the knowledge of telecom experts to have better customer segmentation results. Fig. 2 demonstrates the groups that obtained from experts' advice, each group follows by K-Means clustering method.

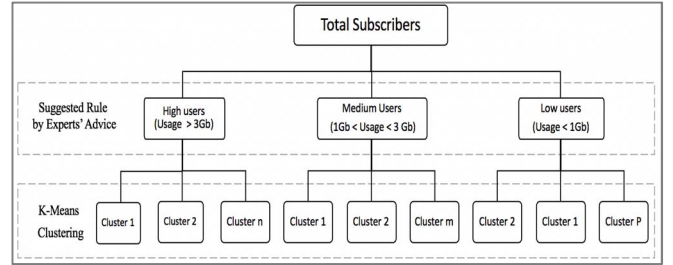


Figure 2. behavioral segmentation

The number of clusters in K-Means clustering algorithm must be defined before executing the algorithm. Therefore, in each group, according to VRC method the value of  $W_k$  calculated for different number of clusters in K-means algorithm to find minimum value for  $W_k$ . The final optimum number of  $K$  in each group are shown in table II.

TABLE II. VRC RESULTS

Groups based on Usage amount	Cluster Numbers (K)	VRC	$W_k$
High Users	3	10.82	4.67
Medium Users	3	17.61	-4.8
Low Users	5	13.32	-1.39

### B. Beneficial segmentation

According to telecom experts' advice, the following rule defined to segment customers in this stage. Regards to data revenue attribute, sorting subscribers in descending order, then choosing the top 20 percent of subscribers as "high-value customers." The second segment is the next top 30 percent of subscribers that we name them "Medium value customers,"

Finally, remaining subscribers, the bottom 50 percent, would be considered as “low-value customers”. Fig. 3 represents the distribution of business data revenue among three groups of subscribers.

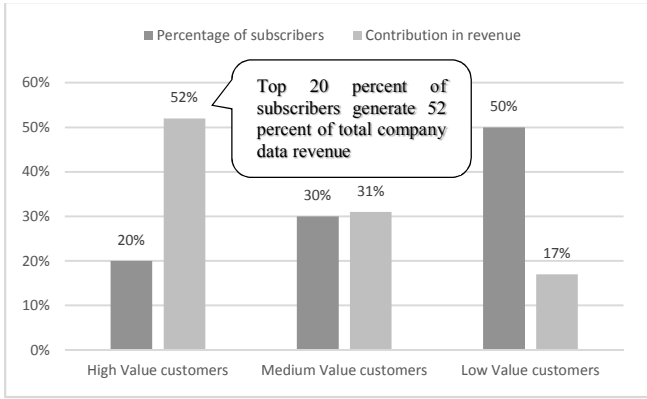


Figure 3. Beneficial segmentation

### C. Final Segmentation

The combination of behavioral and beneficial segmentation results provides a matrix with 11 rows and 3 columns. Rows and columns indicate behavioral and beneficial segmentation respectively. Final segmentation result is shown in Table III.

TABLE III. BEHAVIORAL-BENEFICIAL MATRIX

		Cluster No.	Beneficial Segmentation		
			HIGH VALUE CUSTOMERS	MEDIUM VALUE CUSTOMERS	LOW VALUE CUSTOMERS
Behavior segmentation	High users	1	6 %	1 %	0.1 %
		2	0.2 %	0.01 %	-
		3	0.6 %	0.005 %	-
	MEDIUM USER	1	3 %	5 %	0.1 %
		2	3 %	3 %	1 %
		3	6 %	7 %	2 %
	Low users	1	-	2 %	26 %
		2	0.67 %	10 %	7 %
		3	0.14%	-	13%
		4	-	2%	0.01%
		5	-	-	0.16%

### D. Marketing approaches

Customer segmentation is a success factor for understanding different customers’ behavior and determining profitable customers. After dividing customers into a set of groups, the organization should provide differentiated

strategies based on customers’ segments. Table IV shows some marketing strategies that target different groups of customers to increase revenue, improve retention rate and decrease marketing expenses.

TABLE IV. MARKETING APPROACHES

Marketing Strategies	Segment Code*
Retention strategies (loyalty plans)	(H1,H)-(H2,H)-(L1,M)-(L2,M)-(L4,M)
Promoting long term data packages	(H1,M)-(H1,L)-(H2,M)-(H3,H)-(M1,L)-(M2,L)-(M3,M)-(M3,L)-
Promoting short term data packages	(H3,M)-(M1,M)-
High value customers that need careful marketing approaches (wrong strategy can lead to drop in ARPU)	(M1,H)-(M2,H)-(M3,H)
Promotional plans (discounted offers)	(M2,M)-(L1,L)-(L3,L)
More investigation in segment behavior is required	(L2,H)-(L3,H)
Allocating marketing budget wisely	(L2,L)-(L4,L)-(L5,L)

\* Segment Code represents related behavioral beneficial segment of customers e.g. H1,H according to Table III indicates the segment of High users, Cluster 1, High-value customers.

### E. Model Evaluation

To investigate the validity of the proposed model, we compare the result of two approaches, the approach of segmenting customers in one phase compare to what we did in our research “addressing both features separately and then merge the results in two-dimensional matrix “. The evaluation result is depicted in Table V.

TABLE V. EVALUATION RESULTS

CRITERIA	ONE PHASE CLUSTERING	RESEARCH FRAMEWORK
NUMBER OF CLUSTERS	7	26
COMPREHENSIBLE RESULTS	2	4
WITHIN CLUSTER SIMILARITY	320260	304595
BETWEEN CLUSTER SIMILARITY	0.32	0.34
DIFFERENTIABLE CLUSTERS	2	3

### V. CONCLUSION

The mobile telecommunication marketplace is highly competitive, and mass marketing is not applicable the same as

before. Therefore, telecom industries should define different marketing strategies which targeted different segments of customers to improve their business performance. Mobile operators have huge amount of CDRs which can help them to extract customer's habits and behaviors. Segmentation approaches can support marketing managers to define better targeted marketing strategies. This research has shown that how customer segmentation based on behavioral and beneficial features as a two-dimensional approach could have a better outcome compare to considering all features in one phase. Moreover, we have developed various marketing strategies based on the customers' behavior to increase data revenue. Future work includes studying the performance of clustering with applying other noise cancelling methods, adopting categorical features in K-Means or other methods, and implementing this framework in other domains and industries.

#### REFERENCES

- [1] C. Dullaghan and E. Rozaki, "Integration of Machine Learning Techniques to Evaluate Dynamic Customer Segmentation Analysis for Mobile Customers," arXiv preprint arXiv:1702.02215, 2017.
- [2] G. Armstrong and P. Kotler, *Marketing: an introduction*. Prentice Hall, 2005.
- [3] J. Yang, C. Liu, M. Teng, M. Liao, and H. Xiong, "Buyer targeting optimization: A unified customer segmentation perspective," in *Big Data (Big Data)*, 2016 IEEE International Conference on, 2016, pp. 1262-1271: IEEE.
- [4] I. Bose and X. Chen, "Detecting the migration of mobile service customers using fuzzy clustering," *Information & Management*, vol. 52, no. 2, pp. 227-238, 2015.
- [5] G. Weiss, "Data mining in telecommunications," *Data Mining and Knowledge Discovery Handbook*, pp. 1189-1201, 2005.
- [6] Z. Wang, L. Tu, Z. Guo, L. T. Yang, and B. Huang, "Analysis of user behaviors by mining large network data sets," *Future Generation Computer Systems*, vol. 37, pp. 429-437, 2014.
- [7] L. Luan and H. Shu, "Integration of data mining techniques to evaluate promotion for mobile customers' data traffic in data plan," in *Service Systems and Service Management (ICSSSM)*, 2016 13th International Conference on, 2016, pp. 1-6: IEEE.
- [8] J. D'Haen, D. Van den Poel, and D. Thorleuchter, "Predicting customer profitability during acquisition: Finding the optimal combination of data source and data mining technique," *Expert systems with applications*, vol. 40, no. 6, pp. 2007-2012, 2013.
- [9] M. Siami, M. R. Gholamian, and J. Basiri, "An application of locally linear model tree algorithm with combination of feature selection in credit scoring," *International Journal of Systems Science*, vol. 45, no. 10, pp. 2213-2222, 2014.
- [10] M. Siami, M. R. Gholamian, J. Basiri, and M. Fathian, "An Application of Locally Linear Model Tree Algorithm for Predictive Accuracy of Credit Scoring," in *International Conference on Model and Data Engineering*, 2011, pp. 133-142: Springer.
- [11] W. R. Smith, "Product differentiation and market segmentation as alternative marketing strategies," *Journal of marketing*, vol. 21, no. 1, pp. 3-8, 1956.
- [12] P. W. Murray, B. Agard, and M. A. Barajas, "Market segmentation through data mining: A method to extract behaviors from a noisy data set," *Computers & Industrial Engineering*, vol. 109, pp. 233-252, 2017.
- [13] L.-C. Cheng and L.-M. Sun, "Exploring consumer adoption of new services by analyzing the behavior of 3G subscribers: An empirical case study," *Electronic Commerce Research and Applications*, vol. 11, no. 2, pp. 89-100, 2012.
- [14] A. Beheshtian-Ardakani, M. Fathian, and M. Gholamian, "A novel model for product bundling and direct marketing in e-commerce based on market segmentation," *Decision Science Letters*, vol. 7, no. 1, pp. 39-54, 2018.
- [15] L. Calvet, A. Ferrer, M. I. Gomes, A. A. Juan, and D. Masip, "Combining statistical learning with metaheuristics for the Multi-Depot Vehicle Routing Problem with market segmentation," *Computers & Industrial Engineering*, vol. 94, pp. 93-104, 2016.
- [16] S. Dutta, S. Bhattacharya, and K. K. Guin, "Data mining in market segmentation: a literature review and suggestions," in *Proceedings of Fourth International Conference on Soft Computing for Problem Solving*, 2015, pp. 87-98: Springer.
- [17] D. Chen, S. L. Sain, and K. Guo, "Data mining for the online retail industry: A case study of RFM model-based customer segmentation using data mining," *The Journal of Database Marketing & Customer Strategy Management*, vol. 19, no. 3, pp. 197-208, 2012.
- [18] P. A. Sarvari, P. A. Sarvari, A. Ustundag, A. Ustundag, H. Takci, and H. Takci, "Performance evaluation of different customer segmentation approaches based on RFM and demographics analysis," *Kybernetes*, vol. 45, no. 7, pp. 1129-1157, 2016.
- [19] J. R. Miglautsch, "Thoughts on RFM scoring," *The Journal of Database Marketing*, vol. 8, no. 1, pp. 67-72, 2000.
- [20] J.-J. Huang, G.-H. Tzeng, and C.-S. Ong, "Marketing segmentation using support vector clustering," *Expert systems with applications*, vol. 32, no. 2, pp. 313-317, 2007.
- [21] F. Hamka, H. Bouwman, M. De Reuver, and M. Kroesen, "Mobile customer segmentation based on smartphone measurement," *Telematics and Informatics*, vol. 31, no. 2, pp. 220-227, 2014.
- [22] S. H. Han, S. X. Lu, and S. C. Leung, "Segmentation of telecom customers based on customer value by decision tree model," *Expert Systems with Applications*, vol. 39, no. 4, pp. 3964-3973, 2012.
- [23] L. Xu and H. Chu, "Big Data Analytics toward Intelligent Mobile Service Provisions of Customer Relationship Management in e-Commerce," 2015.
- [24] F. M. Bianchi, A. Rizzi, A. Sadeghian, and C. Moiso, "Identifying user habits through data mining on call data records," *Engineering Applications of Artificial Intelligence*, vol. 54, pp. 49-61, 2016.
- [25] J. Han, J. Pei, and M. Kamber, *Data mining: concepts and techniques*. Elsevier, 2011.
- [26] C.-H. Chen, R.-D. Chiang, T.-F. Wu, and H.-C. Chu, "A combined mining-based framework for predicting telecommunications customer payment behaviors," *Expert Systems with Applications*, vol. 40, no. 16, pp. 6561-6569, 2013.
- [27] R.-S. Wu and P.-H. Chou, "Customer segmentation of multiple category data in e-commerce using a soft-clustering approach," *Electronic Commerce Research and Applications*, vol. 10, no. 3, pp. 331-341, 2011.
- [28] R. Keralapura, A. Nucci, Z.-L. Zhang, and L. Gao, "Profiling users in a 3g network using hourglass co-clustering," in *Proceedings of the sixteenth annual international conference on Mobile computing and networking*, 2010, pp. 341-352: ACM.
- [29] X. Zhang, C. Wang, Z. Li, J. Zhu, W. Shi, and Q. Wang, "Exploring the sequential usage patterns of mobile Internet services based on Markov models," *Electronic Commerce Research and Applications*, vol. 17, pp. 1-11, 2016.
- [30] S.-h. Liao, Y.-j. Chen, and H.-h. Hsieh, "Mining customer knowledge for direct selling and marketing," *Expert Systems with Applications*, vol. 38, no. 5, pp. 6059-6069, 2011.
- [31] C. Ordóñez, "Data set preprocessing and transformation in a database system," *Intelligent Data Analysis*, vol. 15, no. 4, pp. 613-631, 2011.
- [32] A. J. Cuadros and V. E. Domínguez, "Customer segmentation model based on value generation for marketing strategies formulation," *Estudios Gerenciales*, vol. 30, no. 130, pp. 25-30, 2014.
- [33] T. Caliński and J. Harabasz, "A dendrite method for cluster analysis," *Communications in Statistics-theory and Methods*, vol. 3, no. 1, pp. 1-27, 1974.
- [34] G. W. Milligan and M. C. Cooper, "An examination of procedures for determining the number of clusters in a data set," *Psychometrika*, vol. 50, no. 2, pp. 159-179, 1985.
- [35] S. Burks, G. Harrell, and J. Wang, "On Initial Effects of the k-Means Clustering," in *Proceedings of the International Conference on Scientific Computing (CSC)*, 2015, p. 200: The Steering Committee of The World Congress in Computer Science, Computer Engineering and Applied Computing (WorldComp).