流通科学大学論集—流通・経営編—第 31 巻第 2 号，1-19（2019）

Customer Clustering for A New Method of Marketing Strategy Support wthin the Courier Business

Jutamat Jintana*, Takayuki Mori †

This paper presents the application of a data analytics technique to enable courier businesses to better serve customers and conduct customer segmentation more effectively. The proposed methodology introduces the concept of RFM and k-means clustering algorithms for the classification and identification of profitable customers. Accordingly, a set of recommendations is provided to the business on how to best implement effective customer relationship management, particularly in the support of ideas for a new business decision:. Finally, customers are organized into two different groups that help identify market segments more clearly and help decision-makers to develop a suitable marketing strategy.

Keywords: Customer Segmentation, Logistics Service Provider; K-means Clustering, RFM Concept

1. Introduction

In 2017, Thailand had the second highest growth rate in B2C e-commerce sales in Southeast Asia, behind Indonesia (yStats.com, 2017). The growth of e-commerce markets in 2016 was worth up to 2.52 trillion baht (an increase from 2.2 trillion baht in 2015) (Ministry of Digital Economy and Society, Thailand, 2016). The impact of this new form of e-commerce has ushered in new opportunities, as well as challenges, for many courier and delivery service companies in the country. Many courier companies offer different services, such as those who provide specific services such as haulage transport, warehouse or subcontracts to airlines or carriers, called "LSP," or those who provide a full logistics service called "3PL." LSP tends to be developed by offering more service functions to meet the demands of different customers. Therefore, there is an increasing trend within LSP to become 3PL. Foreign and joint ventures are competing along with Thai operators in this sector. For example, Thailand Post has the largest market share of postal and courier services in Thailand. They serve almost 95% of the entire country with a comprehensive network that has grown by 18% since 2016. 50% of their revenue comes from the e-commerce business service (Thansettakij, 22 August 2017). On the other hand, Kerry Express is a courier provider that is owned by Thai and Hong Kong shareholders. In 2017, Kerry Express had the

*Department of Industrial Engineering, Chiang Mai University, Chiang Mai 50200, Thailand
† University of Marketing and Distribution Sciences,3-1 Gakuen Nishi-machi, Nishi-ku,Kobe, 651-2188, Japan

（2018 年 8 月 20 日受理）
©2019 UMDS Research Association

second-largest market share in the sector with 200 branches nationwide. They also generated the highest number of daily deliveries of 250,000 items from their regular daily shipment average of 10,000 items. This was achieved with 3,000 employees in a network with more than 4,000 trucks. Nimexpress is another major courier service provider in the north of Thailand. They currently have more than nine hubs in Bangkok and small branches around the country. Nimexpress plans to expand their business with air cargo in Cambodia, Laos, Myanmar and Vietnam (CLMV). Another new provider in Thailand is SCG Express, who in 2016 went into partnership with Yamato Asia Ltd., a major Japanese LSP company. Investment was provided to the value of around 633 million Baht by providing parcel delivery services for B2B, B2C and C2C (Prachachat online, 2017). These developments are the result of the adaptation and extension of LSP in the search for new opportunities to support more refined requirements.

The company used in the case study of this research is the biggest bus company in the north of Thailand: Chaipattana Transport Co., Ltd. or the Green Bus Company. Over the years, the company has faced strong competition from other forms of public transportation such as taxis (for short journeys) or trains (long journeys) and especially from budget airlines. The boom in budget airlines has affected a huge number of market segments due to their competitive prices and promotions, which have reduced customer numbers and profits. To respond to this threat, the Green Bus Company has developed two new business units. The first of these is a charter service, while the second is a courier service that uses the space available on routed buses. Since 2010, the company has provided a port-to-port courier service along the same routes, which covers 115 service ports in 22 provinces from the north to the south of Thailand. The courier service also includes a packing service, short-term warehousing and, in some areas, a door-to-door service.

The courier service has become an important aspect of the business due to high demand from a wide range of customers, and created a growth in revenues of 17.03% in 2016. The strengths of this area of the business are the high frequency of the buses and low costs as the existing passenger service buses are used. Moreover, development of the business is facilitated by the company’s database. Every sales activity is recorded in the transaction database, but the data is not analyzed and exploited. Instead, it provides the company with ideas on how to develop new business models and a door-to-door service, as well as to support the growth in e-commerce and to generate new sources of revenue.

The customers of the courier business vary in terms of geography, psychology, behavior and general demographic i.e., B2B, B2C and C2C, with each customer exhibiting a different behavior in terms of factors such as the frequency and quantity of demand. It can therefore be seen that combining different segmentations will allow even more specific niche markets to be reached and will create maximum impact. This tactic can be beneficial in both small business start-ups and global companies within all industries. It can therefore be seen that efficient utilization of customer segmentation will help businesses achieve greater focus and reach the people that they want to be talking to, ultimately driving their revenue (Bell, 2016). Hence the Green Bus Company needs to make decisions in term of customer segmentation to provide a suitable service, as well as to conduct businesses while minimalizing investment risks. The first thing the company needs to address when making new business decisions is the characteristics of their customers. Questions, such as the following, therefore need to be asked:

- Who are the most valuable customers for the business? What is the pattern of their behavior?
- What is the segmentation of each customer?

In order to effectively address these business concerns, the techniques of data mining (DM) and customer relationship management (CRM) are creating new opportunities within relationship marketing. In recent years, the k-means clustering algorithm is being widely integrated within the RFM concept for the recognition of patterns and relationships hidden within the data. This is being done not only to develop a product/service that will satisfy customers, but also to track customer purchase behavior and present distinct products/services for each segment. RFM, which stands for Recency, Frequency and Monetary, is a marketing technique used for analyzing customer behavior. It can determine factors such as how recently a customer has made a purchase (recency), how often a customer makes a purchase (frequency), and how much a customer spends (monetary). The integration of the RFM concept and the k-means clustering algorithm provides useful informations on both current and new customers. It is therefore a useful method for improving customer segmentation by dividing customers into various groups for consideration for future personalization services, as well as to identify customers who are more likely to respond to a particular marketing strategy. Extensive research has been proposed on RFM and clustering techniques for different sectors, such as for the online retailer industry (Chen et al.,2012; Daoud et al., 2015), banking (Aggelis et al., 2005), and patient analytics within dental clinics (Wu, Lin and Liu, 2014)

This paper presents a method of customer segmentation through the use of the RFM concept and k-means clustering techniques to classify customers for a courier business, a relatively new, but rapidly growing entrant to the field. To support the idea of a door-to-door service, the analytics of customer demand and customer characteristics are the priority. The main purpose of this research is to help the business better appreciate customer requirements and therefore conduct customer-segmentation more effectively.

The study is organized as follows: Section 2 reviews related literature on RFM, DM and the k-means clustering technique, as well as its application in CRM. Section 3 describes the research methodology. Section 4 is an empirical case study. Finally, Section 5 draws conclusions on the study.

2. Literature Review

This section presents a review of the literature relevant to the fields of customer segmentation, the concept of RFM and k-means clustering techniques. This includes many issues in relation to the main cores of study, which are as follows:

2.1 RFM Concept

RFM is an abbreviation for the words “Recency, Frequency and Monetary.” It was first introduced by Bult and Wansbeek (1995) and has been used ever since in the field of marketing. RFM is used to analyze customer behavior, such as the length of time since the last purchase (recency), the number of purchases within a certain time period (frequency), and the amount of money spent during a certain time period (monetary). Variables relating to customer behavior can improve customer segmentation by observing customers’ attitudes toward a product, brand, benefits, or even loyalty, and recording these on a database (Wei et al., 2010). RFM analysis is cost-effective in acquiring important data regarding customer behavior, and also provides easy to quantify customer behavior analysis (Kahan, 1998; Miglautsch, 2000), in which customer and transactional data can be stored in an accessible electronic form (Lumsden et al., 2008). This allows decision-makers to easily understand the application of an RFM model by summarizing data into a small number of variables (McCarty and Hastak 2007; Wang, 2010).

In recent years, the integration of RFM analysis and data-mining techniques have been proposed for different areas, such as identifying customers and analyzing customer profitability. Aggelis and Christodoulakis (2005) applied RFM in a customer database and measured customer profitability to enhance banking services. Khajvand et al. (2011) proposed a customer lifetime value (CLV) for the customer segmentation of a health and beauty company by using a RFM marketing analysis method in order to segment customers and extend the RFM analysis method by one additional parameter “Count Item.” The results compared calculated CLV for different segments and was used to explain marketing and sales strategies. Chen, Sain and Guo (2012) developed RFM and the k-meanss clustering algorithm to help the business better understand its customers and conduct customer-centric marketing.

Research cases of DM with RFM variables that have been applied with data-mining techniques, include a self-organizing map (Li et al., 2008; Daoud et al., 2015), a neural network and a decision tree (Olson et al., 2009), rough set theory (Cheng and Chen, 2009), CHAID (McCarty and Hastak, 2007), a genetic algorithm (Chan, 2008), sequential pattern mining (Chen et al., 2009; Liu et al., 2009), and a classification and association algorithm model for the prediction of customer behavior and for making recommendations for target customer profiles (Birant, 2011). Several studies have discussed different versions of RFM analysis, for example, the length, recency, frequency, and monetary (LRFM) model proposed by Wu, Lin and Liu (2014). The LRFM model further considers the relationship length between the organization and customers, where L is defined as the number of time periods (such as days) from the first purchase to the last. Alternatively, Hosseini, Maleki and Gholamian (2010) proposed a WRFM-based method with the use of a k-meanss algorithm that classifies customer product loyalty via the concept of B2B. A weighted form of RFM, or WRFM, is another version of RFM that calculates the weight of RFM from the multiplication of wR, wF and wM, according to its relative importance, to make intuitive judgments about ranking and ordering.

2.2 Clustering Rule

Data mining (DM) is a set of techniques used to extract patterns from large datasets by combining methods from statistics and machine learning with database management (Ngai, Xiu and Chau, 2009). These techniques include the association rule, cluster analysis, classification, and regression. Applications include mining customer data to determine segments that are most likely to respond to an offer, mining human resources data to identify characteristics of the most successful employees, or market basket analysis to model the purchase behavior of customers.

The Clustering Rule is the process of grouping physical data. A cluster is a collection of data objects that are similar within the same cluster, but are dissimilar to the objects in other clusters (Han & Kamber, 2001). K-meanss, introduced by J.B. MacQueen in 1967, is a well-known algorithm f or clustering. This algorithm is very sensitive to the choice of a starting point when partitioning the items into K initial clusters. The performance of different clustering methods is compared using the intraclass method, in which the number of fixed clusters that have the value of k is defined (Michaud, 1997; Shina & Sohnb, 2004; Seyed el at., 2010).

This research uses the clustering rule for the grouping of customers who are more likely to use a service. K-meanss is used to assign each record in the dataset to only one of the initial clusters. Each record is assigned to the nearest cluster (the cluster to which it is most similar) using a measure of distance or of similarity, as per the Euclidean Distance Measure Eq.(1) or the Manhattan/City-Block Distance Measure: Eq.(2).

$$Euclidean\ distance\ measure; \qquad d = \sqrt{(a_1 - b_1)^2 + (a_2 - b_2)^2 + \cdots + (a_n - b_n)^2} \qquad (1)$$

$$Manhattan\ distance\ measure; \qquad d = |a_1 - b_1| + |a_2 - b_2| + \cdots + |a_n - b_n| \qquad (2)$$

A suitable number of groups of a dataset is called K-optimal, and this is determined by K-meanss clustering. One method to validate the number of clusters is the elbow method. The idea of the elbow method is to run k-meanss clustering on the dataset for a range of values of k (say, k from 1 to 10 in the examples above), and for each value of k. The average distance with centroid is then calculated using the following Eq.(3), where n is the number of the dataset point, and d is the distance between the clustering centroid and the dataset point. (see Fig. 1)

$$Average\ with\ in\ centroid\ distance = \frac{d_{1c} + d_{2c} + d_{3c}}{n} \tag{3}$$

Following the calculation, a line chart of the average with centroid distance is plotted for each value of k. If the line chart resembles an arm, then the "elbow" of the knee is the optimal number value of k (see Figure 2). The idea is similar to the sum of a square error (SSE). Where a small SSE is preferable, the SSE tends to decrease toward 0 as k is increased (The SSE is 0 when k is equal to the number of data points in the dataset. This is because each data point is its own cluster; and there is no error between it and the center of its cluster). The goal is therefore to minimize the value of k that still has a low average with in-centroid distance. The elbow usually represents where to start diminishing returns by increasing k.

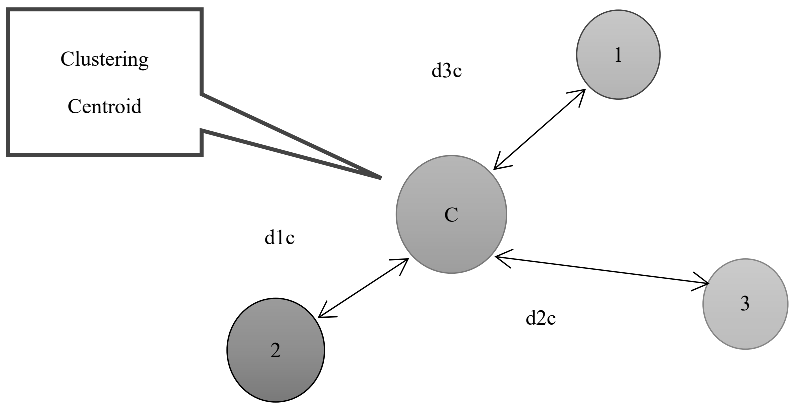

Figure 1. Clustering Centroid Distance

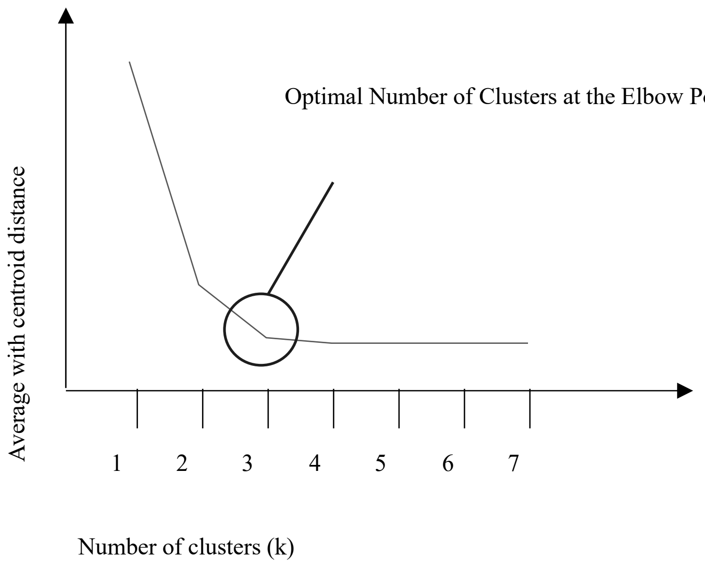

Figure 2. Find to optimal number of clusters (k) following by Elbow Point

3. Research Methodology

The proposed methodology based on the Cross Industry Standard Process for Data Mining (CRISP-DM) is divided into four steps, as shown in Figure 3 below. First, business understanding and data understanding, this step needs to understand business background and data characteristics of case study. Second, Data preparation, begin with data pre-processing technique such as data cleaning, data transformation and redundant some not necessary attribute. Customer dataset were extracted in 3 three dimensions (R, F and M). Then, Modeling step, this step purpose value for R, F and M were firstly determined optimal numbers of clusters, based on K-optimal method and customers are segmented into similar clusters according to their RFM variables by K-mean technique. Finally, evaluation and deployment, a group of customer segmentation can divided by concept of customer pyramid and putting right service to targeted customers group and hence improvement of marketing strategies efficiency.

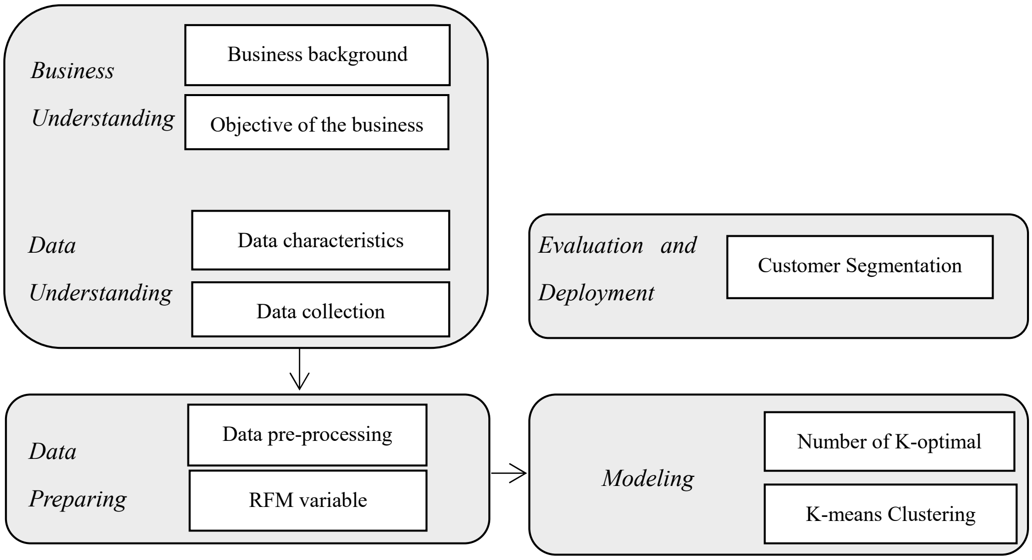

Figure 3. Methodology framework for customer segmentation based on the RFM model and k-means clustering techniques

4. Empirical Study

The methodology given above was implemented for the Green Bus Company. The process can be shown according to the following steps:

Step 1: Business Understanding and Data Understanding

The company has a database system, and the data has been collected since 2013. The database is updated as follows: when a customer uses a service, a member of staff registers a member ID for each customer. Each service activity made by each customer is recorded as a sales transaction as basic selling Excel data. This provides the opportunity of creating a value from the database by mining the knowledge that possibly supports the introduction of a door-to-door service. However, to do this, it is necessary to know what services the customers are likely to use, and the suitable marketing strategies for each customer. To do this, it is necessary to answer the following questions:

- Who are the most valuable customers for the business? What is the pattern of these customers' behavior?
- What is the segmentation of each customer?

Step 2: Data Preparing

In this step, the data from the sales transaction database is taken from the 1 January to the 31 December, 2016. After this was done, 496,433 records were obtained. The main attributes of these are the Shipper/Consignee Name, Date Income, Time Income, Cargo Category Name, Cargo Type Name, Cargo Weight, Station Begin/End and Cargo Fee. (shown in Table 1). The data preparation process is as follows:

1. Firstly, the database is cleaned. Records that include missing values or inaccurate values are deleted.
2. Secondly, the redundant attributes are eliminated, and the datum is transformed into a format that will be more easily and effectively processed for clustering customer value. Appropriate variables of interest are then selected from the sales transaction database. In this case, sender, date income, price and sum of service using the calculated RFM value are selected, as shown in Tables 2 and 3.
3. Three essential aggregated variables are created, i.e. recency, frequency and monetary. The values of these variables are calculated per each customer. (See Tables 2 and 3)

- For recency, the attribute “Dateincome” is used to determine the period time (day) a customer has recently used the service. For example, Customer A used the service on the 31 May 2016, the 1 December 2016 and the 20 December 2016. The last date of customer service is deducted with the last date of the year 2016 (31 December 2016). The recency value of Customer A is 11 days.
- For frequency, the sum of the service using (time) for each customer is determined. For example, in 2016, Customer A used the service on the 3 June 2016 twice and 9 July 2016 once. The frequency value of Customer A is therefore 3 times.
- For monetary, the attribute “Price” is used to determine the sum of each customer’s spending. The attribute “Price” is calculated from the cargo weight and quantity. The unit is baht.

Table 1. An Example of a Customer Transaction Dataset (496,433 records)

Sender	Member Apply	Consignee	Payment	Sale Station	Station Begin	Station End	Date income	Time income	Cargo id	Cargo type name	Quantity (Item)	Weight (Kg.)	Price (Bath)
A40000	1/1/2010	บรศ ปลาเงิน	CASH	BKK.;Maochit;	BKK.;Maochit;	Chiang Mai	**1/1/2016**	4:45:06	3377153	Document	1	1	0
A40000	1/1/2010	[illegible]	CASH	Phare	Phare	Chiang Mai	**1/1/2016**	5:50:26	3377154	Document	1	1	0
0979908682	12/9/2015	[illegible]	CASH	Chiangchom	Chiangchom	Chiang Mai	**1/1/2016**	6:30:26	3377155	Clothes Garment	1	7	85
bu02303	7/5/2015	สำนักงานการประปา สาขา 9	CASH	Maesaey	Maesaey	Chiang Mai	**1/1/2016**	6:46:00	3377157	Document	1	2	60
Mitoko2499	5/18/2013	[illegible]	AGENT	Samyekkasetsuk	Samyekkasetsuk	Chiang Mai	**1/1/2016**	6:59:41	3377158	Agriculture Products	1	23	137
3500700564	1/2/2013	[illegible]	CASH	Chiang Mai(HO)	Chiang Mai	Mae Jan	**1/1/2016**	7:01:00	3377159	Agriculture Products	2	84	423
…	…	…	…	…	…	…	…	…	…	…	…	…	…

Table 2. Transformation of the Data

Initial Data	Transformed Data
The last service use date	Recency (Day)
The total number of times the service is used	Frequency (Time)
Total money spent	Monetary (Baht)

Table 3. An Example of the Transformation Data

Customer ID	Recency (R)	Frequency (F)	Monetary (M)*
BU00341	1	5,969	871,886
cu04917	0	2,975	641,600
A35198	0	2,724	255,916
bu01926	0	2,516	154
bu00181	1	2,133	146,495
…	…	…	…

*Thai Baht (THB)

Step 3: Modeling

The clustering method is used for customer segmentation. The modeling approach to customer clustering applies RFM parameters and the k-meanss technique. In the k-meanss clustering technique, the number of clusters should be determined to identify the optimal value of k. Various metrics can be employed, which are defined in Eq. (3). Figure 4 shows a number line plotting each point in the dataset. An elbow chart shows the average centroid distance after running k-meanss clustering for k ranging from 1 to 20. The elbow of the arm k = 6 is shown in Fig. 5, indicating that 6 is the best number for clusters. RapidMiner 6.5 software is used for the implementation.

-

Step 4: Evaluation and Deployment

The results show differences in the behavioral values of RFM. There is one customer in Cluster 3, 4 customers in Cluster 4 and 29 customers in Cluster 0. Cluster 3 has the highest monetary value. This is due to the sole customer having strong business relationships due to having the highest volume of cargo with the highest monetary value, in other words a high recency value but only a moderate frequency value. For Clusters 4 and 0, there are groups of customers who have the highest frequency, very high recency and monetary value respectively. Cluster 3 seems to contain regular and/or loyal customers of the courier business. These customers have the characteristics of high recency service use, and always using the service with high volume and quantity. These groups of customers are probably entrepreneurs or business customers, not individual retail customers.

Cluster 5 relates to 116 consumers. Their behavior is also indicated by high recency, frequency and monetary values. This is in contrast to Cluster 2 of 636 customers with lower recency, frequency and monetary values respectively. The behavior of these customers is typified by high service use and the volume of delivery being quite large. These two groups of customers are also important.

The majority of the customers are grouped in Cluster 1, which contains 102,024 customers, representing 99% of the entire population. The use of the service by these customers is indicated with very low frequency, low monetary and high recency values. This cluster can be called ‘irregular’ or ‘uncertain’ customers. The behaviors of these customers indicate seldom use of the service with low quantity. The results are presented in Tables 4-6.

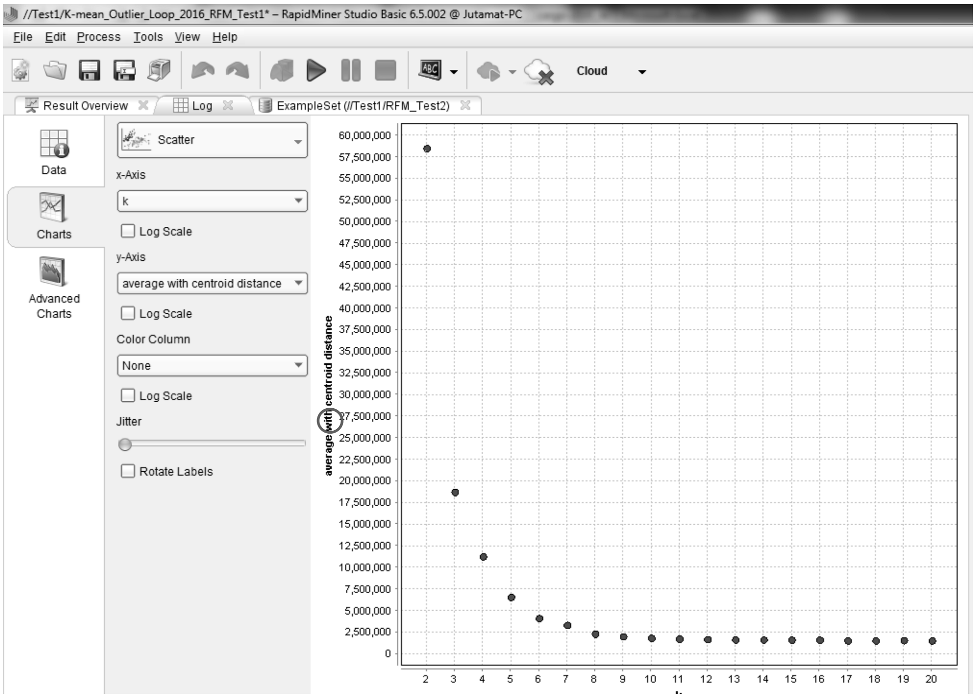

Figure 4. Elbow and k-optimal

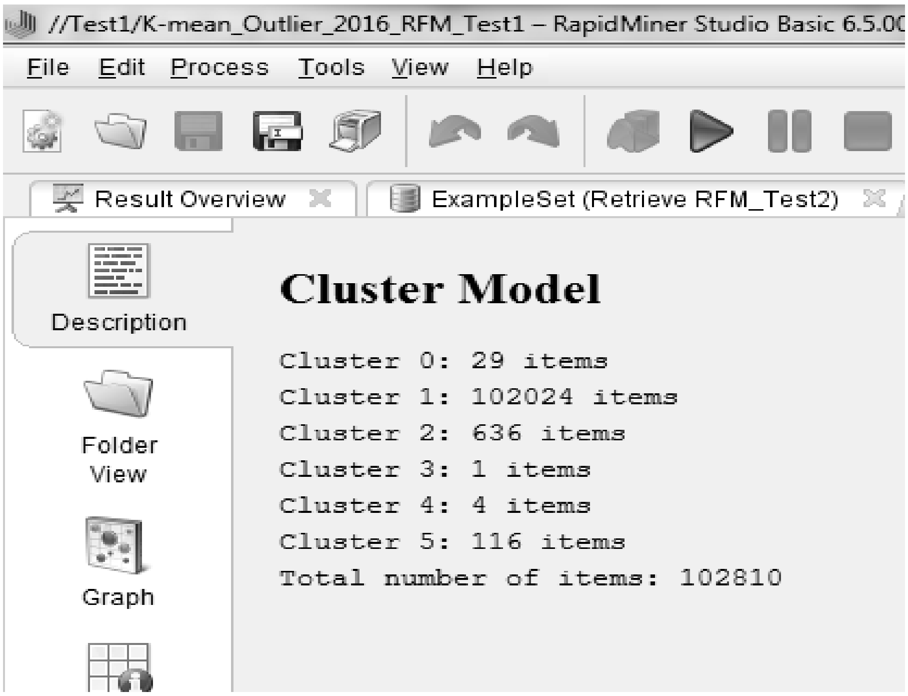

Figure 5. Numbers of the Cluster from the experiment

The majority of the customers are grouped in Cluster 1 with 102,024 customers, composed of 99% of the whole population. These customers use the service with very low frequency, low monetary value and high recency values. This cluster can be called irregular or uncertain customers. The low quantity values of these customers indicate their seldom use of the service. The results are presented in Tables 4-6.

Table 4. Descriptive Statistics of Recency, Frequency and Monetary Values.

Variable	Maximum	Minimum	Average	Standard Deviation
Recency (R)	365	0	145.52	107.907
Frequency (F)	5,969	1	4.59	36.263
Monetary (M)	2,389,532	4	726.96	10,107.842

Table 5. Statistics Relating to the Six Cluster Based on k-means Clustering.

Cluster	Number of Customers (Total: 102,810)	Recency (R) (Day)	Frequency (F) (Time)	Monetary (M) (Baht)
Cluster 0	29	3.6	808.1	247,263.3
Cluster 1	102,024	146.6	3.2	389.4
Cluster 2	636	13.1	106.5	22,174.6
Cluster 3	1	2.0	380.0	2,389,532.0
Cluster 4	4	0.8	2,826.5	644,120.0
Cluster 5	116	3.9	364.0	75,643.7
Average	-	145.5	4.6	727.0

Table 6. Statistics Relating to the Six Clusters Based on k-means Clustering.

	Maximum	Minimum		Maximum	Minimum
Cluster 0			Cluster1		
Recency	67	0	Recency	365	0
Frequency	2,724	53	Frequency	2,516	1
Monetary	402,002	167,193	Monetary	11,259	4
Cluster 2			Cluster 3		
Recency	221	0	Recency	2	2
Frequency	718	1	Frequency	380	380
Monetary	48,556	4	Monetary	2,389,532	2,389,532
Cluster 4			Cluster 5		
Recency	2	0	Recency	0	102
Frequency	5,969	667	Frequency	9	2,133
Monetary	871,886	515,107	Monetary	49,250	151,946

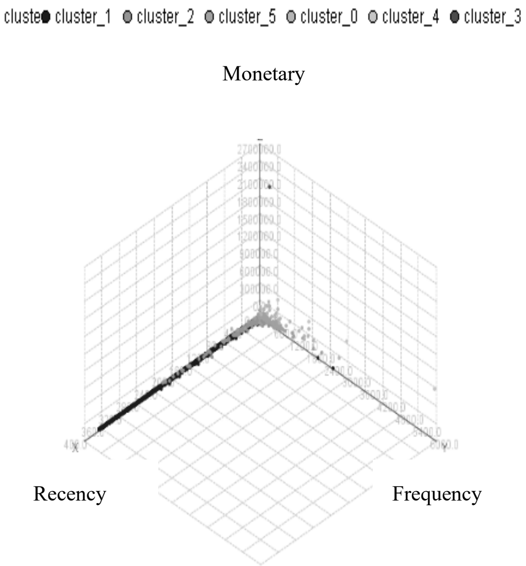

Figure 6. 3D scatter plot of Recency, Frequency and Monetary

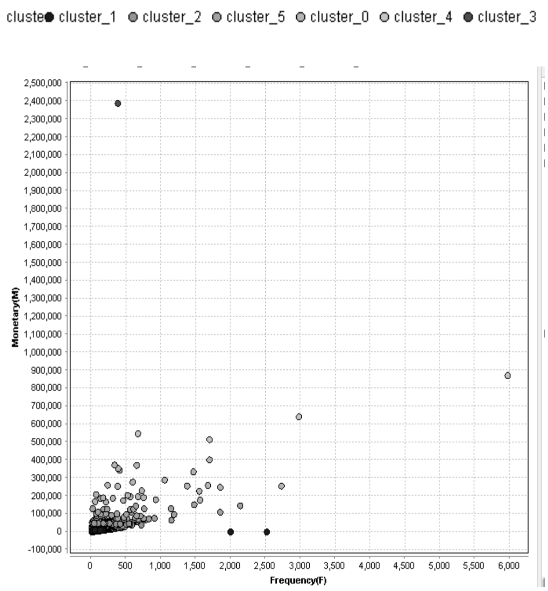

Figure 7. Customer Cluster of Frequency and Monetary

Step 5: Deployment

The Green Bus Company aims to create a new market by providing a door-to-door service. However, the company has limited resources as the cost of such services can be high, and so they cannot offer this service to all customers. Therefore, a study of customer behavior through this research can provide an overview of the customers and allow the classification of them in a profitable manner. As specified in the work of Zeitham, Rust & Lemon (2001), it is important to understand the needs of customers at different levels of profitability, and to adjust services based on those differences. This is a more critical analysis of the company than what has been previously performed. Specifically, in examining customers in terms of profitability and understanding the key elements of costs and revenue aspects of the profit equation, it is possible to actually increase the current and future profitability of all customers in the firm's customer portfolio.

The Customer Pyramid is a tool that enables organizations to utilize differences in customer profitability. Organizations can utilize this tool to strengthen the link between service quality and profitability, as well as to determine the optimal allocation of scarce resources. Figure 8 shows the concept of the customer pyramid. (Zeitham, Rust & Lemon, 2001)

From the concept of the customer pyramid, a group of customer characteristics can divined. By taking these into account, the strategic positioning of customer clusters is grouped in terms of RFM symbols. If the average value is greater than the total average, an upward arrow (↑) is given to that particular symbol, while a downward arrow (↓) is used when the average value is less than the total average. The combinations of RFM model can therefore be classified into 4 groups.

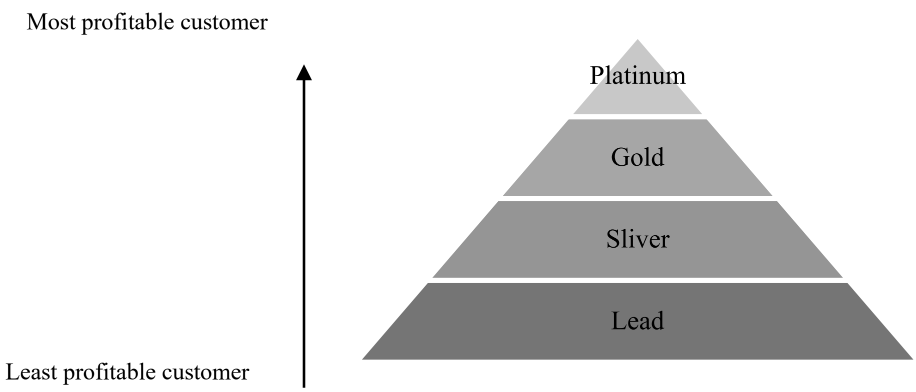

Figure 8. Concept of customer pyramid

- The Platinum level refers to the company's most profitable customers, typically those who are heavy users with very high recency, frequency and monitory values. The loyal customers are R↑F↑M↑
- The Gold level differs from the Platinum level in that profitability levels are not as high, perhaps because the customers prefer price discounts, which limits margins or service quality. The customers in this group might not be as loyal to the firm, and even though they are heavy users in the product category, they might minimize risk by working with multiple vendors rather than just with the local company. This group has at least one variable value that is lower than the average. RFM symbols are R↑F↑M↓, R↑F↓M↑, R↓F↑M↑
- The Silver level contains customers who provide the volume needed to utilize the firm's capacity. But their spending levels, loyalty, and profitability values are not substantial enough for special treatment. RFM symbols are R↑F↓M↓, R↓F↑M↓, R↓F↓M↑
- The Lead level consists of one-time customers or new customers who use the service infrequently. Most of the members of the Lead group are business customers who have used the service only a few times or just once. The members of this group demand more attention than they are due, given their spending and profitability. Members of this group may sometimes complain about the firm and thus tie up the firm's resources. Their low RFM is represented by the symbol R↓F↓M↓

The results from customer clustering are indicated in Table 7.

Table 7. RFM Values and the Customer Levels of the Green Bus Company

Cluster	Number of Customers	RFM value	Customer Level
Cluster 0	29	R↑F↑M↑	Platinum
Cluster 1	102,024	R↓F↓M↓	Lead
Cluster 2	636	R↑F↑M↑	Platinum
Cluster 3	1	R↑F↑M↑	Platinum
Cluster 4	4	R↑F↑M↑	Platinum
Cluster 5	116	R↑F↑M↑	Platinum

The company should assign a specific service for customers at the Platinum level, as these firms are large purchasers. Platinum clients tended to stay with the company, and were willing to try new services and approaches, and so, this group of customers would be open to a door-to-door service. As their service behavior indicates that they use the service every day or every week with a large number of couriers, the provision of a door-to-door service would provide them with help to reduce their transport costs, as well as supply chain costs, increase flexibility and help them to focus on their core competencies. This is seen as an advantage for the company in increasing the quality of service, retaining existing customers and obtaining an additional channel to acquire new customers. Zeithaml, Rust and Lemon (2001) stated that, for remaining loyal customers, a company might develop or offer services that may satisfy the client's needs such as:

- Providing outsourcing
- Becoming a full-service provider
- Increasing brand impact by line extensions
- Offering service guarantees
- Creating structural bonds

Customers in the Lead group is the largest pool but at the lowest level. These customers have only used the service once. For customer at the Lead level, the company should aim on upgrading customers in this group to the Gold and Silver levels. Therefore, the marketing policy should focus on:

- Marketing
- Direct sales promotions
- Improving brand image and service quality
- Becoming a customer expert through technology

Providing quality services can revive a customer's trust in a firm. In addition, the promotion can provide discount or special offers following 10 uses of a service, or the delivering of an item with a higher weight. Frequent promotion programs will improve a strong service program and create customer reliability.

5. Conclusion

Each customer creates a different value. Customer segmentation helps the business to appreciate the value of different customers and their requirements in order to identify different marketing strategies or customer relationship management, such as high-profit customers who are entitled to use a special service. For secondary value customers, businesses may try to encourage their customers to have more business with them. This group is able to add value to higher levels in order to generate more revenue for the organization. Grouping based on customer value is not based on customer-generated revenue alone, but by considering the various factors involved.

In this study, customers are clustered into segments according to RFM and integrated RFM parameters using a k-meanss algorithm. Clustering customers into different groups helps decision-makers identify market segments more clearly and thus develop more effective marketing and sales strategies for customer retention. This is an important starting point for the use of a database to support the making of decisions in the development of new business models by the data analytics technique.

Acknowledgments

This work was jointly supported by the Thailand Research Fund, Chiang Mai University and Chaipattana transport Co., Ltd. (Contract No. PHD58I0068). The authors gratefully acknowledge the Faculty of Commerce, University of Marketing and Distribution Sciences, Kobe and Excellence Center in Logistics and Supply Chain Management (E-LSCM), and Chiang Mai University for their support in the research of this work.

References

Aggelis, V., & Christodoulakis, D. (2005, July). Customer clustering using rfm analysis. In Proceedings of the 9th WSEAS International Conference on Computers (p. 2). World Scientific and Engineering Academy and Society (WSEAS).

Birant, D. (2011). Data Mining Using RFM Analysis. In Knowledge-oriented applications in data mining. InTech.

Bult, J. R., & Wansbeek, T. (1995). Optimal selection for direct mail. Marketing Science, 14(4), 378-394.

Chan, C. C. H. (2008). Intelligent value-based customer segmentation method for campaign management: A case study of automobile retailer. Expert systems with applications, 34(4), 2754-2762.

Chen, D., Sain, S. L., & Guo, K. (2012). Data mining for the online retail industry: A case study of RFM model-based customer segmentation using data mining. Journal of Database Marketing & Customer Strategy Management, 19(3), 197-208.

Chen, Y. L., Kuo, M. H., Wu, S. Y., & Tang, K. (2009). Discovering recency, frequency, and monetary (RFM) sequential patterns from customers' purchasing data. Electronic Commerce Research and Applications, 8(5), 241-251.

Cheng, C. H., & Chen, Y. S. (2009). Classifying the segmentation of customer value via RFM model and RS theory. Expert Systems with Applications, 36(3), 4176-4184.

Daoud, R. A., Amine, A., Bouikhalene, B., & Lbibb, R. (2015, November). Combining RFM model and clustering techniques for customer value analysis of a company selling online. In Computer Systems and Applications (AICCSA), 2015 IEEE/ACS 12th International Conference of (pp. 1-6). IEEE.

Elizabeth Bell (2016). Customer Segmentation: 5 Ways to Divide the Consumer Base. Retrieved March 2, 2018, from http://www.thebridgecorp.com/customer-segmentation/

Han, J., & Kamber, M. (2001). Data mining concepts and techniques San Francisco Moraga Kaufman.

Hosseini, S. M. S., Maleki, A., & Gholamian, M. R. (2010). Cluster analysis using data mining approach to develop CRM methodology to assess the customer loyalty. *Expert Systems with Applications*, *37*(7), 5259-5264.

Kahan, R. (1998). Using database marketing techniques to enhance your one-to-one marketing initiatives. Journal of Consumer Marketing, 15(5), 491–493.

Khajvand, M., Zolfaghar, K., Ashoori, S., & Alizadeh, S. (2011). Estimating customer lifetime value based on RFM analysis of customer purchase behavior: Case study. Procedia Computer Science, 3, 57-63. Chicago

Li, S. T., Shue, L. Y., & Lee, S. F. (2008). Business intelligence approach to supporting strategy-making of ISP service management. Expert Systems with Applications, 35(3), 739-754.

Liu, D. R., Lai, C. H., & Lee, W. J. (2009). A hybrid of sequential rules and collaborative filtering for product recommendation. Information Sciences, 179(20), 3505-3519.

McCarty, J. A., & Hastak, M. (2007). Segmentation approaches in data-mining: A comparison of RFM, CHAID, and logistic regression. Journal of business research, 60(6), 656-662.

Michaud, P. (1997). Clustering techniques. Future Generation Computer Systems, 13(2-3), 135-147.

Ngai, E. W., Xiu, L., & Chau, D. C. (2009). Application of data mining techniques in customer relationship management: A literature review and classification. Expert systems with applications, 36(2), 2592-2602.

Olson, D. L., Cao, Q., Gu, C., & Lee, D. (2009). Comparison of customer response models. Service Business, 3(2), 117-130.

Prachachat (2017). SCG-Yamato Logistics. Retrieved February 12, 2018, from https://www.prachachat.net/news_detail.php?newsid=1473833319

Qiasi, D., Minaei-Bidgoli, B., & Amooee, G. (2012). Developing a model for measuring customer's loyalty and value with RFM technique and clustering algorithms. The Journal of Mathematics and Computer Science, 4(2), 172-181.

Rygielski, C., Wang, J. C., & Yen, D. C. (2002). Data mining techniques for customer relationship management. Technology in society, 24(4), 483-502.

Wei, J. T., Lin, S. Y., & Wu, H. H. (2010). A review of the application of RFM model. African Journal of Business Management, 4(19), 4199.

Wu, H. H., Lin, S. Y., & Liu, C. W. (2014). Analyzing patients' values by applying cluster analysis and LRFM model in a pediatric dental clinic in Taiwan. The Scientific World Journal, 2014.

Seyed Mohammad Seyed Hosseini, Anahita Maleki, Mohammad Reza Gholamian, (2010) "Cluster analysis using data mining approach to develop CRM methodology to assess the customer loyalty," Expert Systems with Applications 37, 5259–5264, 2010.

Shin, H. W., & Sohn, S. Y. (2004). Segmentation of stock trading customers according to potential value. Expert systems with applications, 27(1), 27-33.

Thansettakij (2017). Economic news. Retrieved March 2, 2018, from http://www.thansettakij.com/content/196856.

Zeithaml, V. A., Rust, R. T., & Lemon, K. N. (2001). The customer pyramid: creating and serving profitable customers. California Management Review, 43(4), 118-142.