A method for customer lifetime value ranking — Combining the analytic hierarchy process and clustering analysis

Ya-Yueh Shih

is a doctoral candidate at the Institute of Information Management at the National Chiao Tung University and is also based at the Department of Information Management, MingHsin University of Science and Technology. She received her BSc and MSc degrees in information management from the National Yunlin University Science Technology, Taiwan in 1996 and 1998. Her current research interests include data mining, consumer behaviour and electronic commerce.

Chung-Yuan Liu

is Deputy Manager of the Engineering Division, Synergy Integration Tech., Inc. His research interests include data mining and recommendation systems. He received his BSc and MSc degrees in information management from the National Yunlin University of Science and Technology, Taiwan in 1996 and 2001.

Abstract Creating successful transaction actions to retain customers for future repurchasing is extremely important in today's fiercely competitive environments. The best marketing strategy is to obtain customers who are profitable and remain loyal for a lifetime. However, although the recency, frequency and monetary (RFM) technique has been used to predict customer behaviour for over 50 years in direct marketing, few studies have discussed the relative importance of RFM variables via a systematic approach. Therefore, this study applies the analytic hierarchy process (AHP) to determine the relative importance of RFM variables (weighted RFM) in evaluating customer lifetime value (CLV). Clustering techniques are then employed to cluster customers according to weighted RFM value. The simple weighted sum approach is then used to derive CLV ranking and thus customer segments can be identified and compared clearly. The usefulness of the approach is verified by applying it to a hardware retailer, and a useful marketing database is derived to evaluate the proposed method. Finally, this study also discusses three perspectives for validating the proposed method. This study provides a useful method of CLV ranking and can thus assist market practitioners in performing more effective market segmentation.

Ya-Yueh Shih
Institute of Information Management, National Chiao Tung University, 1001 Ta Hsueh Rd, Hsinchu 300, Taiwan.

Tel: +886 3 5721257;
e-mail:
moonshih@iim.nctu.edu.tw

INTRODUCTION

Berry defined relationship marketing as based on building, maintaining and enhancing customer relationships in multiservice organisations.[1] Relationship management is emerging as the core marketing activity for businesses operating in fiercely competitive environments. Additionally, customer relationship management (CRM) is a customer–oriented business philosophy that involves analysing, planning and controlling customer relationships via modern information and communication

techniques. CRM has become a buzz word among business practitioners and consultants since the late 1990s.[2] Therefore, attracting customers is considered to be an intermediate step in the relationship–building process with the ultimate aim of increasing the loyalty of profitable customers in CRM.

All customers are different, however, even those receiving identical products or services, and thus niche marketing is important to obtain a competitive advantage. Niche marketing is a marketing strategy that uses product differentiation to appeal to a focused group of customers.[3] From this perspective, niche marketing emphasises product differentiation and market segmentation. Hughes[4] also indicated that building a database and identifying customer segments are elements that make up the key first step in strategic database marketing. Developing separate marketing programmes for each segment is what it is all about.

Furthermore, the computation and application of customer lifetime value (CLV) has become an important issue for marketing. It was pointed out that CLV is the net profit derived from an individual customer during a certain period.[5] Enterprises can compute CLV for each customer from purchase records, and thus can forecast an individual customer's benefit, distribute promotions and allocate resources to retain customers. Generally, methods for measuring CLV include recency, frequency and monetary value (RFM). A number of authors[6–10] suggest that the RFM method would avoid focusing on less profitable customers and allow resources to be diverted to more profitable ones. In fact, RFM has been used in direct marketing to predict customer behaviour for more than 50 years. It is one of the most powerful

techniques available to database marketing. It does not require any additional data. If a firm has a database of its customers, with their purchase history, the firm can use RFM analysis immediately at virtually no cost.[11] From the behavioural perspective, the RFM method is an important one for assessing the relationship between enterprise and customers.

However, how to evaluate CLV using RFM? In fact, a number of studies have discussed the evaluation of CLV. Hughes[12] proposes a method for RFM scoring, which involves sorting the customer data into quintiles. Meanwhile, Stone hypothesised that RFM variables had different weights depending on industry characteristics.[13] To analyse the value of customers who paid using credit cards, Stone suggested placing the highest weighting on the number of purchases, followed by the period of purchase time, followed by the purchase time, while placing the lowest weighting on the amount of purchases. While various combinations and weightings have been proposed, judging and weighting RFM variables remains subjective.

Therefore, this paper uses an analytic hierarchy process (AHP)[14–16] to evaluate the weight (relative importance) of each RFM variable through the perception of decision makers. K–means, one of the most widespread approaches in clustering methodology, is used to group customers into those with similar lifetime value or loyalty based on the performance value of RFM in this study. Each target market can be further ranked using a simple weighted–sum approach. The case study of a hardware retailer is used for illustration.

Finally, three viewpoints to validate the proposed method are also discussed, including using non-weighted RFM (relative importance among RFM variables are equal) to cluster and thus to

observe the CLV ranking, applying this in another empirical case study for model appropriation and using a predictive model with the decision tree algorithm to help validate the proposed method. According to the decision tree algorithm, classification rules are determined for each cluster using the decision tree algorithm C5.0. The classification rules are employed to predict other potential customers regarding their groups and loyalty ranking. The remainder of this study is arranged as follows. First, the following section reviews related works in evaluating CLV, including clustering techniques and RFM evaluation. Then the following section outlines the methodology used herein. The next section illustrates an empirical case to elucidate the proposed method and also presents three viewpoints to validate the proposed method. Finally, the conclusion summarises the contributions of this study and outlines areas for further research.

RELATED WORK

This section reviews the concepts used to evaluate CLV ratings. Clustering analysis is used to segment the target market then related works are reviewed to evaluate CLV.

Clustering analysis for customer segmentation

Clustering is one of the data mining tools used to discover knowledge processes.[17] Clustering aims to maximise variance among groups while minimising variance within groups. The variance represents distance, distance is measured by such methods as Euclidean, correlation or mutual information. In clustering many algorithms have been developed continuously, such as K-means, hierarchical, fuzzy c-means

approaches and so on.

The clustering method can be divided into hierarchical and nonhierarchical.[18] Judging which of these approaches is best and how to determine when to apply it appropriately is difficult. Punj and Stewart[19] compared the advantages, disadvantages and outliers affecting clustering to several other approaches, and the comparison revealed that average link and Wards methods were better. Punj and Stewart also noted that researchers used the average linkage or Ward's methods to identify the number of clusters, removed the outliers and then clustered by nonhierarchical clustering approaches. K-means is one of the most widespread approaches; it is used in this study.

The K-means clustering algorithm[20] is shown as follows: (1) beginning with an initial partition of the data set into the specified number of clusters, the centroid of these clusters is calculated; (2) each data point is allocated to the cluster whose centroid is the nearest to this data point; (3) the new centroid of the clusters is computed after all data are reallocated; (4) steps 2 and 3 are repeated until data points no longer change at each step.

RFM evaluation

RFM terms were explained by Bult and Wansbeek[21] as follows: (1) R (recency): time period since the last purchase, and the lower the value is, the higher the probability of the customer making a repeat purchase; (2) F (frequency): number of purchases made within a certain time period; higher frequency indicates higher loyalty; (3) M (monetary): the amount of money spent during a certain time period; increasing monetary contribution from customers indicates increased focus on the company supplying the products.

Table 1: The scoring method for each R, F and M

Variables	The scoring of each criterion	The weighting level
R (Recency)	Last 3 months = 24; last 6–9 months = 6 Last 3–6 months = 12; last 9–12 months = 3 More than one year = 0	Medium
F (Frequency)	Number of purchases multiplied by 4	Highest
M (Monetary)	The amount of purchases multiplied by 10 per cent (the highest is 9)	Lowest

A widely-used method for evaluating RFM was proposed by Hughes.[22] Each RFM is divided into quintiles. The latest purchase time of 20 per cent customers is set to 5, and is inferred to other customers. Meanwhile, a score of 1 indicates that the most recent transaction was a long time ago. Purchase frequency and monetary value are ranked using the same system. Finally, the RFM scores are obtained for each customer, with that of the best customer equalling 555 while that of the worst equals 111, and different marketing strategies are developed for different customers accordingly.

Stone[23] hypothesised that the weights of RFM variables vary depending on industry characteristics. To analyse the value of customers with credit cards, Stone proposed that the number of purchases be assigned the highest weighting, followed by purchase timing and finally purchase size. Table 1 lists the scoring architecture.

In practical applications, RFM variables need to measure different weights in different industries (see, eg Stone[24]). Stone, however, determined the RFM weightings subjectively, without using a systematic approach or evaluation to determine the RFM weightings. This study employs AHP to evaluate each weight (relative importance) among RFM variables, and specifically asks decision makers to make intuitive judgments about ranking order to produce pairwise comparisons.

METHODOLOGY FOR CLV RANKING

The proposed method primarily utilises AHP and clustering techniques, as shown in Figure 1. This work employs the AHP to evaluate the weighting (relative importance) of each RFM variable, and specifically asks decision makers to make intuitive judgments about ranking order to make pairwise comparisons. K–means clustering is then employed to group customers with similar lifetime value or loyalty, according to weighted RFM. Finally, the simple weighted sum of normalised RFM values are used to rank each cluster to identify the profitable customers.

EXPERIMENTAL RESULTS

A data set is used to elucidate the proposed methodology. The case concerns a hardware retailing company that manufactures wheels, casters, platforms and hand trucks for industrial, medical, hospital and institutional use. This company presently produces over 3,000 products. Its decision makers must target customer groups and develop market strategies to satisfy customer needs and thereby increase the market share of the company. Two years of data on consumer transactions, approximately 70,000 rows, have been collected. The data set is preprocessed to extract customer transactions. Unreasonable records such as those of customers who have a non–zero amount of purchases but have never made any transactions are also
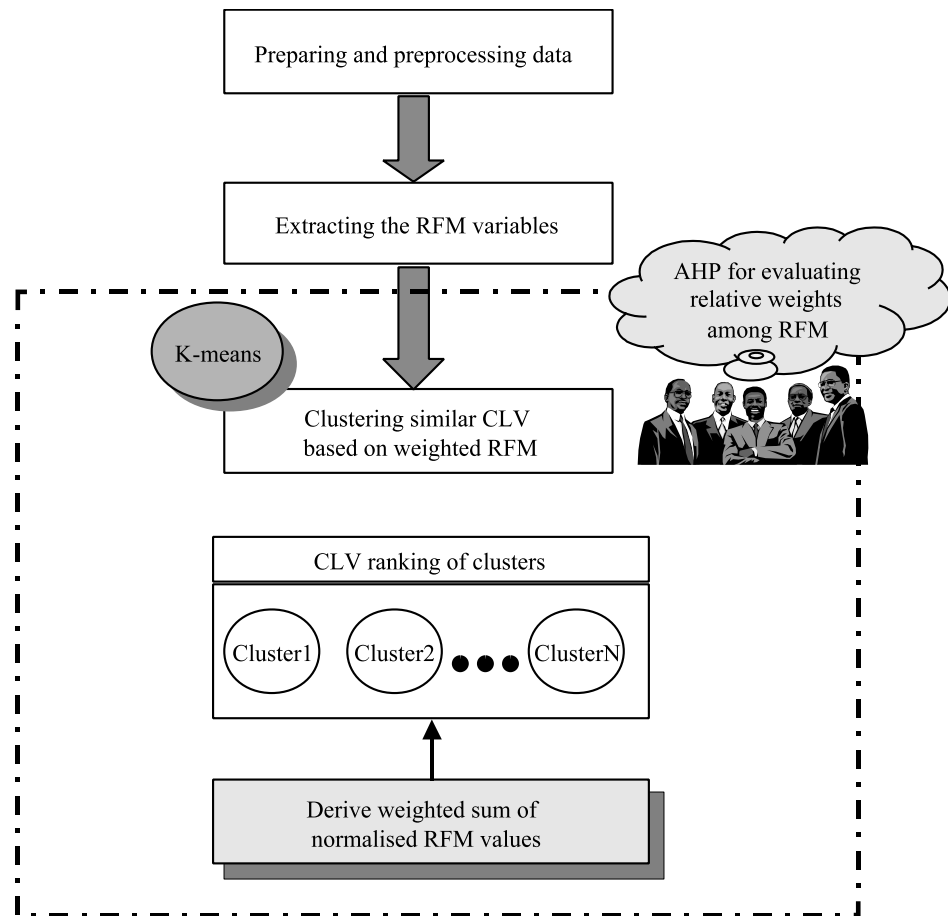
Figure 1: Architecture for CLV ranking

removed. RFM values of the 895 customers are extracted from the marketing database to measure the customers' CLV. (70,000 rows represent 70,000 transaction records in the database. Each row records a product that a customer bought. In this particular marketing database, there are 895 customers.) Table 2 shows some CLVs expressed in terms of RFM.

In measuring the CLV, the relative importance of RFM values may vary with the characteristics of the product or industry. AHP is used to assess the weighting of RFM variables by asking decision makers to make intuitive judgments about ranking order to yield pairwise comparisons, as detailed below. Furthermore, K–means clustering is used

to group customers with similar lifetime value, based on their weighted RFM value. The resulting groups (clusters) represent market segments to be targeted with different marketing strategies. Analysis of variance (ANOVA) is performed to test whether RFM significantly discriminates between these market segments. The section headed 'Clustering customers with similar lifetime value' details the clustering of customers.

In fiercely competitive environments, different marketing strategies should be aimed at different customer groups (market segments). Accordingly, identifying CLV or loyalty ranking of customer segments will greatly help decision makers to target markets more

Table 2: RFM values for each customer[25]

Customer no	Recency (days)	Frequency	Monetary (New Taiwan dollars)
1108001	65	434	1,252,430
1108003	411	7	37,930
1108006	496	2	8,780
1108008	68	50	287,845
1108011	65	128	916,210
…	…	…	…
1860003	159	87	313,763

Table 3: Relative degree of importance for pairwise comparisons

Comparative importance	Description	Explanation
1	Equal importance	Two activities contribute equally to the objective
2	Intermediate between equal and weak	Experience and judgment slightly favour one activity over another
3	Weak importance of one over another	
4	Intermediate between weak and strong	Experience and judgment strongly favour one activity over another
5	Essential or strong importance	
6	Intermediate between strong and demonstrated	An activity is strongly favoured and its dominance is demonstrated in practice
7	Demonstrated importance	
8	Intermediate between demonstrated and absolute	The evidence favouring one activity over another is of the highest possible order of affirmation
9	Absolute or extreme importance	

clearly and make more effective strategies. The section headed 'CLV ranking' describes the CLV ranking of market segments. Finally, three viewpoints to validate the proposed method are discussed in the section headed 'Validating the proposed method'.

AHP approach

The AHP[26–28] is used to determine the relative importance (weights) of the RFM variables, w_R, w_F, and w_M, respectively. The three main steps of the AHP are as follows.

Step 1: Perform pairwise comparisons. This step asks evaluators (decision makers) to make pairwise comparisons of the relative importance of RFM variables using the scale shown in Table 3. (Pairwise comparison is a widely used method for evaluating relative importance among critical incidents. Three-way comparison could be easier to use than pairwise comparison since there

are only three variables needed in evaluation. If, however, there are more than n critical incidents, pairwise comparison can compare C_2^n ways whereas three-way comparison could not easily deal with such a case. Therefore, the authors believe that pairwise comparison outperforms the three-way comparison.)

Step 2: Assess the consistency of pairwise judgments. Evaluators may make inconsistent judgments when making pairwise comparisons. For example, 'I like apples more than oranges', 'I like oranges more than bananas' and 'I like bananas more than apples'. Before the weights are computed based on the pairwise judgments, the degree of inconsistency is measured by the inconsistency index. Perfect consistency implies a zero inconsistency index. Perfect consistency, however, is typically not achieved, since human beings are often biased and inconsistent when making subjective judgments. Therefore,

Table 4: AHP questionnaire for RFM

Criteria	Comparative importance									Criteria
	9:1	7:1	5:1	3:1	1:1	3:1	5:1	7:1	9:1	
Recency	9	7	5	3	1	3	5	7	9	Frequency
Frequency	9	7	5	3	1	3	5	7	9	Monetary
Monetary	9	7	5	3	1	3	5	7	9	Monetary

Table 5: Example of RFM pairwise comparison matrix

	Recency	Frequency	Monetary
Recency	1	5	7
Frequency	1/5	1	3
Monetary	1/7	1/3	1

an inconsistency index of less than 0.1 is acceptable. If the inconsistency index exceeds 0.1, then the pairwise judgments may be revised before the weights of RFM are computed.

Step 3: Computing the relative weights. This step determines the weight of each decision element according to the pairwise comparisons. This work employs eigenvalue[29] computation to derive the weights of RFM.

In this study, the following three groups of evaluators judge the RFM weightings: (a) three administrative managers, (b) two business managers in sales, and one marketing consultant and (c) five customers who have made at least one purchase. These groups were invited to evaluate the relative importance of the RFM variables. Data were gathered by interviewing the evaluators. Interviews were conducted using a questionnaire (Table 4), and the answers were expressed in the form of a pairwise comparison matrix (Table 5).

According to the assessments obtained by the AHP, the relative weights of the RFM variables are 0.7306, 0.1884 and 0.081 respectively. The implication of the RFM weightings is as follows. Recency is the most important variable, since the unit price of hardware products is relatively low, and thus evaluators mainly concentrate on whether customers purchase continually. If some customers perform no transactions for a long period, the customers may have been lost or have transferred to new vendors.

Clustering customers with similar lifetime values

In this section, customers with similar lifetime values, in terms of weighted RFM, are clustered using the K-means method. The method must specify the number of clusters, *k*, in advance. The parameter k is set to 8, since eight $(2 \times 2 \times 2)$ possible combinations of inputs (RFM) can be obtained by assigning ↓ or ↑, according to whether the average R (F, M) value of a cluster is less than or greater than the total average R (F, M). The RFM values of customers are normalised as follows. The profit form, $x' = (x - x^S)/(x^L - x^S)$, is used to normalise the F (frequency) and M (monetary) values, since F and M positively influence CLV or loyalty. The cost form, $x' = (x^L - x)/(x^L - x^S)$, is used to normalise R (recency) value, since R negatively impacts CLV. Notably, x' and x represent the normalised and original R (F, M) values, respectively, while x^L and x^S represent the largest and smallest R (F, M) value of all customers,

Table 6: The results of K-means clustering among 984 customers

Customer segmentation	Number of customers	Recency (days)	Frequency	Monetary (New Taiwan dollars)	Pattern
Potential crisis	19	228.53	101.31	551,210.20	R↑F↑M↑
Valueless	136	170.22	16.58	60,793.83	R↑F↓M↓
Valueless	54	283.85	8.05	35,479.72	R↑F↓M↓
Gold	30	67.23	254.70	1,142,476.00	R↓F↑M↑
Gold	136	70.66	102.01	357,753.70	R↓F↑M↑
Potential	467	80.25	24.99	93,086.88	R↓F↓M↓
Valueless	94	417.37	19.15	96,256.32	R↑F↓M↓
Valueless	48	620.69	2.40	17,169.58	R↑F↓M↓
Total average	123	163.57	40.38	159,481.10	

respectively. The normalised RFM values of each customer are then multiplied by the relative importance of RFM variable, w_R, w_F and w_M, which are determined by the AHP. The K-means method is then applied to cluster the customers into eight groups, according to the weighted RFM values.

Table 6 lists the clustering result, listing eight clusters, each with the corresponding number of customers and average R, F and M values. The last row also shows the total average, that is the overall average RFM values over all customers. The average RFM values of each cluster are compared with the total averages. If the average R (F, M) value of a cluster exceeds the total average R (F, M), then an upward arrow ↑ is shown; otherwise, a downward arrow ↓ is shown. The last column of Table 6 shows the RFM pattern for each cluster.

Each cluster represents a market segmentation. Customers in clusters 4 and 5 have the same characteristic (with the pattern R↓F↑M↑), both have average recency below the total average (163.57); along with frequency and monetary exceeding the total average (40.38 and 159,481.1). Consequently, customers in clusters 4 and 5 would be considered to be loyal customers who purchase frequently and contribute a lot of monetary value. These customers are called gold customers. Gold customers are the biggest contributors to company profitability, and some may deal with just one company. Therefore, marketing strategies were proposed to maintain customers' activity by providing special customer services and thus helping to ensure customer loyalty.

Cluster 6 displays the pattern R↓F↓M↓ and may represent new customers who have recently visited the company to make a purchase. These customers try to develop closer relationships with the company. Reasons for purchase frequency and purchase amount being lower than average may be that these customers have suffered from operating problems or depressed businesses. The purchase activity of these customers indicates that they have the potential to become gold customers. These customers are called potential customers. Although their monetary contribution, purchase frequency and recency were all below the total average, this indicates that these customers were beginning to develop a closer relationship with the company and were also increasing their purchase ability gradually. From the marketing strategy perspective, potential customers must be made to realise that if they increase their purchase frequency, the company will give them improved services, similar to gold customers. The company will also

attempt to establish a long-term relationship with such customers.

Cluster 1 displays the pattern R↑F↑M↑, and represents customers who may once have enjoyed a good relationship with the company. The purchase frequency and quantity of such customers exceeds the average purchase frequency and quantity. However, these customers have not made transactions recently, possibly because of relocating their businesses or suffering bankruptcy. Defections may occur if a company ignores these customers. This study called such customers potential crisis customers. This study must reconsider whether inattentiveness towards customers causes customers defections, and also should consider how to retain these customers. After all, retaining an existing customer is cheaper than developing a new one.

Finally, Clusters 2, 3, 7 and 8 display the pattern of R↑F↓M↓. These customers almost never visit and make very few transactions. Furthermore, these customers generally only make purchases during sales. Enterprises can reduce prices to attract these customers, but will suffer reduced margins as a result. Since few enterprises achieve profits from these customers, they are classed as valueless customers. Valueless customers only care about price.

Consequently, such customers will readily defect to other suppliers with lower-priced products. A company can only retain such customers by offering low prices, but the profits that can be obtained from such customers are very limited. However, offering low prices can still achieve a win–win situation for companies and valueless customers.

Analysis of variance is used to determine whether RFM variables can be used to distinguish the eight clusters with statistical significance. The analysis rejects the null hypothesis H_0 because

the p-values are significant ($p < 0.05$). The result confirms that these eight clusters can be distinguished by recency, frequency and monetary. (Statistical significance in the sample is not due merely to random sampling variation, but reflects an actual difference or relationship in the population. However, practical importance is a substantial difference or relationship that is important psychologically, biologically, clinically, etc. In fact, p-value significance does not determine the practical importance of the result but helps to support the decisions being made.)

CLV ranking

The CLV ranking is derived to help market practitioners develop more effective strategies for retaining customers and thus clearly identify and compare market segments. The CLV ranking of clusters proceeds as follows. The RFM values of each customer are normalised, as described above. Table 7 shows the average normalised RFM values of each cluster, denoted as C_R^j, C_F^j and C_M^j, respectively, for $j = 1$ to k (the number of clusters). C_R^j, C_F^j and C_M^j are computed by averaging the normalised RFM values of customers in cluster j. Let C_I^j be the integrated rating of cluster j. C_I^j is computed as the weighted sum of C_R^j, C_F^j and C_M^j; that is,
$C_I^j = w_R \star C_R^j + w_F \star C_F^j + w_M \star C_M^j$, where w_R, w_F and w_M are the relative importance of the RFM variables, as determined by the AHP. Finally, the CLV ranking of the clusters is derived according to their integrated rating. The ranking indicates that cluster 7 has the highest ranking, followed by cluster 5; cluster 8 has the lowest ranking. Customers in a cluster with a higher rank are more loyal. Accordingly, decision makers can make different strategies to retain customers according to CLV ranking or loyalty.

Table 7: The average rating for all market segmentation by $R \times w$

Customer segmentation	Recency	Frequency	Monetary	Integrated rating (v)	Loyalty ranking
1 Potential crisis	0.71	0.39	0.47	0.63	4
2 Valueless	0.81	0.06	0.04	0.61	5
3 Valueless	0.61	0.02	0.02	0.45	6
4 Gold	1.00	1.00	1.00	1.00	1
5 Gold	0.99	0.39	0.30	0.82	2
6 Potential	0.98	0.09	0.07	0.74	3
7 Valueless	0.37	0.07	0.07	0.29	7
8 Valueless	0.00	0.00	0.00	0.00	8
Total average	0.68	0.25	0.25	—	—

Validating the proposed method

The validating approach to the proposed method is threefold: (1) using non-weighted RFM (relative importance among RFM variables are equal) to repeat an experiment according to the proposed architecture (Figure 1); (2) applying it in another empirical case study for model appropriation; (3) using a decision tree algorithm to construct predictive models based on equal/unequal relative importance among RFM values, to understand the prediction. These three viewpoints are detailed in the following subsections.

Non-weighted RFM

The experimental results CLV ranking based on non-weighted RFM is listed in the right-hand column of Table 8. The fourth cluster displays the highest loyalty, followed by the fifth cluster. It appears especially clear that clusters 2, 3, 7 and 8 are the disloyal groups of customers, with group 8 being the worst, and thus the company should pay less attention and distribute less resources to them.

There is some difference on cluster 1 (potential crisis customers) and 6 (potential customers) and cluster 6 (potential customers) between weighted RFM (see Table 7) and non-weighted RFM. That seems reasonable. For cluster

1 in Table 7, the frequency (0.39) and monetary (0.47) are higher than the total average F (0.25) and M (0.25). The time period since last purchase (recency) is very long, indicating that customers in this cluster may have been lost or have transferred to other vendors. For cluster 6 in Table 7, although the frequency (0.09) and monetary (0.07) were lower than the total average, the recency (0.98) indicates that they have recently been active. Marketers should devote greater effort to retaining customers in this cluster than those in cluster 1. If, however, individual RFM weights are not considered, as the results in Table 7, cluster 1 ranks ahead of cluster 6. The comparison implies that the proposed approach may be a better method for evaluating the CLV ranking. More important is that this result is consistent with the decision makers in the empirical case.

Application in another empirical case study for model appropriation

In this study, the unit price of the hardware retailer's products is relatively low and thus decision makers only care about whether customers purchase continuously or not. In other words, some customers are only mildly profitable, but they are very loyal. They continue to stay with the firm for years

Table 8: The average rating for all of clusters by normalisation in each RFM

Customer segmentation	Recency	Frequency	Monetary	Integrated rating (v)	Loyalty ranking
1 Potential crisis	0.71	0.39	0.47	1.58	3*
2 Valueless	0.81	0.06	0.04	0.91	5
3 Valueless	0.61	0.02	0.02	0.65	6
4 Gold	1.00	1.00	1.00	3.00	1
5 Gold	0.99	0.39	0.30	1.69	2
6 Potential	0.98	0.09	0.07	1.13	4*
7 Valueless	0.37	0.07	0.07	0.50	7
8 Valueless	0.00	0.00	0.00	0.00	8
Total average	0.68	0.25	0.25	—	—

*difference between weighted RFM and non-weighted RFM

and years. Some customers, however, are highly profitable, but disappear after a single purchase. No matter how satisfied these buyers, it is almost impossible to maintain a profitable long-term relationship with them. But in other empirical contexts, eg luxury goods such as new televisions, highly profitable customers are more important than mildly profitable but very loyal customers. For this, the relative importance of RFM values is monetary > recency and frequency. According to the authors' assumption about relative importance among RFM values (0.1, 0.1 and 0.8, respectively), there are only two clusters (the most loyal and the least loyal) where the weighted RFM is the same as the non-weighted RFM, the remaining are different. Therefore, this systematic approach would be valuable and useful for either fast-moving consumer goods industries or high consumption industries such as television and car manufacturers.

Using the decision tree algorithm to validate the clustering results

To validate the clustering results, the authors adopted the decision tree algorithm to construct a decision tree and rules to calculate classification accuracy rates of existing customers. The predictive model of decision tree

and rules was used to predict for potential customers which segments they belong to and which CLV ranking they map to. It would help decision makers to consider relative marketing strategies for ideal customers. The earliest decision tree algorithm was extended from Concept Learning System (CLS) — Iterative Dichotomizer 3 (ID3), which works by computing a metric known as the information gain ratio.[30] The spirit of the decision tree is achieved by maximising the information gain threshold at each node in the decision tree and the evaluation is based on classification validation. The ID3 has been refined into C4.5 by Quinlan.[31]

This paper uses the latest C5.0 algorithm, proposed in 1998, and the See5 software released by Rulequest.[32] Customers are randomly selected from a database and split 70 per cent for calibration and 30 per cent for validation. The first set is used as a training set to construct the classification rules; the latter is used for testing. The accuracy rate of the confusion matrix is critical to validating the classification result. Figure 2 displays the decision tree with 12 leaves node.

Rulesets are generally easier to understand than trees since each rule describes a specific context associated with a class. Furthermore, rulesets

```
Recency <= 117:
:...Frequency <= 63: 6 (325)
:    Frequency > 63:
:    :...Frequency > 180: 4 (19)
:        Frequency <= 180:
:        :...Monetary <= 1767742: 5 (97/1)
:            Monetary > 1767742: 4 (2)
Recency > 117:
:...Recency <= 224:
    :...Frequency <= 56: 2 (89/1)
    :    Frequency > 56:
    :    :...Monetary <= 189820: 2 (3)
    :        Monetary > 189820: 1 (12/1)
    Recency > 224:
    :...Recency > 502: 8 (33)
        Recency <= 502:
        :...Recency <= 347:
            :...Monetary <= 201080: 3 (36)
            :    Monetary > 201080: 1 (2/1)
            Recency > 347:
            :...Frequency <= 129: 7 (69)
                Frequency > 129: 1 (2)
```

Figure 2: Decision tree

generated from a tree generally have fewer rules than the tree has leaves.

Each rule summarised the performance using the statistics (N/E, lift L) or (N, lift L) where: (1) N denotes the number of training cases covered by the rule; (2) E (if shown) represents the number of training cases covered that do not belong to the rule's class. Meanwhile, the accuracy of the rule is estimated by the Laplace ratio $(N - E + 1)/(N + 2)$; (3) L is the estimated rule accuracy divided by the previous probability of the rule class. Taking rule 9 as an example, if Recency ≤ 117 and Frequency ≤ 63, then cluster 6 contains 325 customers. The accuracy of the rule is estimated to equal 0.997.

Rule 9: (325, lift 2.1) {Recency ≤ 117 and Frequency $\leq 63 \rightarrow$ class 6 [0.997]}

Finally, Table 9 lists the classification result of the calibration and validation samples, 11 rulesets are produced to classify eight different clusters, and the error rate of classification is only 3.4 per cent for calibration samples. According to

the analytical results, the clustering based on weighted RFM should be a reliable step for identifying profitable customers.

DISCUSSION

This work presents a method of CLV ranking that primarily utilises AHP and clustering techniques. Applying AHP to determine the relative importance of RFM variables is important, since the RFM weights may vary with product and industry characteristics. K-means clustering is then employed to group customers with similar lifetime value or loyalty, based on weighted RFM. Finally, the simple weighted sum of normalised RFM values are used to rank clusters and identify profitable customers.

Experimental results indicated that the proposed method can yield a more reasonable CLV ranking than methods that do not consider the relative importance of RFM variables. Not only can the proposed method apply to general consumption industries but it is

Table 9: The confusion table of calibration sample and validation sample

Calibration sample (689 cases) Decision tree size = 12; Ruleset no = 11			Validation sample (295 cases) Decision tree size = 12; Ruleset no = 11		
Cluster	Accuracy/error	Accuracy rate (%)	Cluster	Accuracy/error	Accuracy rate (%)
1	14/0	100.00	1	4/1	80.00
2	91/0	100.00	2	45/0	100.00
3	36/0	100.00	3	16/2	88.89
4	21/0	100.00	4	9/0	100.00
5	96/1	98.96	5	37/2	94.87
6	325/2	99.39	6	136/4	97.14
7	69/1	98.57	7	23/1	95.83
8	33/1	97.06	8	15/0	100.00
Decision tree error = 4 (0.6%) Ruleset error = 4 (0.6%)			Decision tree error = 9 (3.1%) Ruleset error = 10 (3.4%)		

also appropriate to high consumption industries.

In addition, this study validates the proposed model using decision tree algorithms by calculating classification accuracy rates and constructing a model for predicting which other potential customers will realise their CLV ranking for future purchases using previous purchasing records. Future studies will address two themes. First, the approach was evaluated experimentally using a marketing database obtained from a hardware retailer. In the future, the approach could be evaluated for other application domains, such as electronic commerce. Secondly, one–to–one marketing should be of paramount importance for tailoring service, friendship and information to satisfy customers. Therefore, it is necessary to consider how to combine this with the CLV ranking approach for more effective personalisation and service. Such an approach would be invaluable for marketing practitioners in the future.

References

1 Berry, L. L. (1983) 'Relationship marketing of services: Growing interest, emerging perspectives', *Journal of the Academy of Marketing Science*, Vol. 23, No. 4, pp. 236–245.

2 Sonja, G. K. and Moedritscher, G. (2002) 'Alternative approaches toward measuring CRM performance', 6th Research Conference on Relationship Marketing and Customer Relationship Management', Altanta, June, pp. 9–12.

3 Linneman, R. E. and Stanton, J. L. (1991) 'Making niche marketing work: How to grow bigger by acting smaller', McGraw-Hill, Inc., New York.

4 Hughes, A. M. (1994) 'Strategic database marketing', Probus Publishing, Chicago.

5 *Ibid*.

6 Ha, S. H. and Park, S. C. (1998) 'Application of data mining tools to hotel data mart on the Intranet for database marketing', *Expert Systems with Applications*, Vol. 15, No. 1, pp. 1–31.

7 Kahan, R. (1998) 'Using database marketing techniques to enhance your one–to–one marketing initiatives', *Journal of Consumer Marketing*, Vol. 15, No. 5, pp. 491–493.

8 Miglautsch, J. (2000) 'Thoughts on RFM scoring', *Journal of Database Marketing*, Vol. 8, No. 1, pp. 67–72.

9 Saarenvirta, G. (1998) 'Data mining to improve profitability', *CMA Magazine*, March, pp. 8–12.

10 Shaw, M. J., Subramaniam, C., Gek Woo Tan and Welge, M. E. (2001) 'Knowledge management and data mining for marketing', *Decision Support Systems*, Vol. 31, No. 1, pp. 127–137.

11 Hughes (1994) *op. cit*.

12 *Ibid*.

13 Stone, B. (1995) 'Successful direct marketing methods', NTC Business Books, Lincolnwood, IL.

14 Saaty, T. L. (1980) 'The analytic hierarchy process', RWS Publications, Pittsburgh, PA.

15 Saaty, T. L. (1990) 'Multicriteria decision making: The analytic hierarchy process', Vol. 1, AHP Series, McGraw-Hill, New York, NY.

16 Saaty, T. L. (1994) 'Fundamentals of decision making and priority theory with the analytic hierarchy process', RWS Publications, Pittsburgh, PA.

17 Chen, M. S., Han, J. and Yu, P. S. (1996) 'Data mining: An overview from a database perspective', *IEEE Transactions on Knowledge and Data Engineering*, Vol. 8, No. 6, pp. 866–883.

18 Johnson, R. A. and Wichern, D. W. (1992) 'Applied multivariate statistical analysis', Prentice Hall, Englewood Cliffs, New Jersey.

19 Punj, G. N. and Stewart, D. W. (1983) 'Cluster analysis in marketing research: Review and suggestions for application', *Journal of Marketing Research*, Vol. 20, pp. 134–148.

20 Anderberg, M. (1973) 'Cluster analysis for application', Academic Press, New York, p. 1.

21 Bult, J. R. and Wansbeek, T. J. (1995) 'Optimal selection for direct mail', *Marketing Science*, Vol. 14, No. 4, pp. 378–394.

22 Hughes (1994) *op. cit*.

23 Stone (1995) *op. cit*.

24 *Ibid*.

25 This paper uses RFM variables to evaluate customer lifetime value (CLV). The original value of RFM variables are summarised in Table 2.

26 Saaty (1980) *op. cit*.

27 *Ibid*.

28 *Ibid*.

29 For each pairwise comparison matrix A, this paper uses the theory of eigenvector, ie $(A - \lambda_{max}I)W = 0$, to calculate the eigenvalue λ_{max} and the eigenvector, where $w = (w_1, w_2, \ldots w_n), n$ is the number of critical incidents. Accordingly, weights can be estimated.

30 Quinlan, J. R. (1979) 'Discovering rules by induction from large collections of examples', in Michie, D. (ed.) 'Expert systems in the micro-electronic age', Edinburgh University Press, Edinburgh.

31 Quinlan, J. R. (1993) 'C4.5: programs for machine learning', Morgan Kaufmann Publishers.

32 Rulequest research (1998) data mining tools See5 and C5.0, http://www.rulequest.com/See5-infor.html.

Furthermore, the relative importance among RFM variables may differ and depend on industry or products. Therefore, this paper used the AHP method to determine the relative importance to decision makers. The results of K-means clustering are shown in Table 7.