Asian Journal of Information and Communications 2017, Vol. 9, No. 2, 58-66

KISDI KOREA INFORMATION SOCIETY DEVELOPMENT INSTITUTE Asian Journal of Information and Communications

Customer analysis of the securities companies based on modified RFM model and simulation by SOM neural network

Juan Cheng*
Shanghai International Studies University

Xiongwei Zhou**
Everbright Securities Co, Ltd

Bixuan Fu***
King's College London

Abstract

The customers of the securities company have been classified according to the balance of their accounts for a long time. The method cannot measure the value of customers dynamically. Some securities companies apply RFM model to classify customers in recent years. But only recency, frequency and monetary cannot evaluate customers of the securities companies well. The commission of the trading varieties should be taken into account when calculating recency, frequency and monetary as it will affect the value of customers. This paper proposes a modified RFM model in order to precisely evaluate the customers of the securities companies. Every variety is weighted according to the commission in the modified RFM model. In order to evaluate the modified RFM model, the paper classifies 5,000 randomly chosen customers by self-organizing feature map neural network. The results show the customers are classified into eight categories and the number of general customers and valueless customers increase significantly which reflects that the modified RFM model can classify customers strictly and precisely. So, the modified RFM model is helpful for the securities company to develop corresponding service strategies.

Keywords: securities companies, customer analysis, RFM model, SOM neural network, simulation

1. Introduction

Customer relationship management is the central to an enterprise. Customer classification is the base of customer analysis, which is central to customer relationship management. Customer classification can distinguish customers into star customers, general customers, valuable customers and etc. Enterprises adopt different marketing strategies for different customers in order to concentrate the limited resources on the high- value customers. Furthermore, enterprises can establish differentiated service system to maintain high-value customers and maximize the profit based on the classification.

Homogenization of business and service in securities industry has existed for many years. The problem the securities companies face is how to improve the competitiveness and expand the market shares. Developing customer-oriented service strategy based on customer analysis is a solution to the problem.

* Joint first author: 1110210@xdsisu.edu.cn
** Joint first author: zhouxw@ebscn.com
*** Corresponding author: bixuan.fu@kcl.ac.uk

RFM(recency, frequency, monetary) model is a widely applied method to analyze customers, which has been used in industries including bank, telecom and etc. If the model is applied in securities companies without any adjustment, the results tend to reflect the activeness rather than the value of customers. Some researches added other variables such as age and income into the RFM model when analyzing the customers of the securities companies. But the results reflect the potential value of the customers. To fully measure the value of the customers, we must consider the time, the variety and the volume of the transactions. This paper introduces the effect of the varieties of transactions on the value of the customers through the RFM model by weighting the varieties when calculating the recency. At the end of the paper, we simulate the classification by SOM neural network. The results of the simulation prove that the modified RFM model can measure the contribution to the profit of the securities companies rather than the activity of the customers. Therefore, the modified RFM model can truly realize the purpose of categorization of the customers according to their value and provide support for developing customer-oriented strategies.

2. Related works

RFM model is an important tool to measure customer's value and profitability. The existing researches about customer classification mainly focus on refining the three factors based on the research industry. For example, the RFM model was modified by adding age and gender in customer classification of the securities companies (Xiaoyan and Chunying, 2008; Moeini and Alizadeh, 2016). For the customer classification of the bank industry, R, F, M were subdivided into three categories in order to divide customers into different subsections (Bin and Jindong, 2013). In telecommunications industry, the recency is replaced by the customers' last payment time since customers use telecom services every day and the recency is zero. Also, the payment time, frequency and monetary are given different weights (Sheng and Xu, 2006). Some researches use average monetary to replace monetary in customer classification since sometimes frequency and monetary change in the same direction (Ziyi, 2014).

Many algorithms are used in customer classification on the basis of RFM model. SOM neural network was used relatively early in customers classification (Chenbo and Liang, 2004). Since SOM algorithm does not provide accurate clustering information after classification, SOM neural network is used in combination with K-mean clustering. First, to get the number of clustering and the center points by SOM neural network. Then, to classify customers by K-mean clustering with the determined number of clustering and the center points (Huan, Guangming and Gaoyu, 2010).

K-mean clustering and k-central clustering are widely used in customer classification (Moeini and Alizadeh, 2016; Chiungi and Juichih, 2017). Genetic algorithm and binary-tree classifier are often used in customer classification. (Bin and Jindong, 2013). These researches above do not fully reflect the abilities of the customers to make profit for the securities companies.

In order to measure the value of the customers of the securities companies, the paper modifies RFM model by weighting the recency of different varieties and uses SOM neural network to stimulate the classification based on the modified RFM model.

3. Customer classification based on the modified RFM model

RFM model is a prevalent method for classifing customers. R is short for recency which is the time of the last transaction of the customer. The shorter the recency, the more active the customer is. F is short for frequency which is the number of transactions in a period of time. Customers who trade frequently are more valuable. M is short for monetary which is the amount of the last transaction. Short recency, high frequency and great monetary mean superior customers.

RFM model is a very practical method since the three factors can be easily obtained and provide a full description of the customers. However, in the RFM model, only the time and the amount of the last transaction are not enough. There exists a problem that customers with profitable transactions in previous periods rather than in recent periods are regarded as valueless customers. In order to avoid the problem, a modified RFM model is designed to evaluate the customers of the securities companies.

In the modified RFM model, *R*, F*, M** are used to denote the modified R, F , M. Then,

$$R^* = \sum_{i-1}^{n} \frac{1}{w_i} R_i \qquad F^* = F \qquad M^* = \sum_{i=1}^{F} M_i$$

ω_i is the weight of variety *i*. It depends on the commission of variety *i*. The weight equals to the ratio of the commission of variety *i* to that of the basic variety such as A-share. A great weight means a great commission and a little modified recency. R_i is the recency of the variety i calculated as the number of days from the current time. A smaller R^* contributes to a higher value customer which has the same rule with *R*. M_i is the volume of the variety *i* of the analysis period. M^* is the total volume. F^* is the same as before.

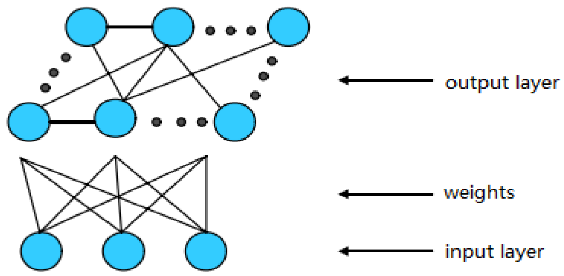
4. Modeling and learning method of SOM neural network

SOM neural network is widely used in digital image processing, system assessment and categorization. SOM can improve the flexibility and efficiency of image segmentation when applied to digital image processing. SOM can also be applied in system assessment such as evaluation of water quality. When applied to text categorization, SOM shows good classification ability as the correct rate is up to 95%. Compared with other algorithms, the most essential advantages of the SOM algorithm are its simple mechanism and quick response. Also, a model for the SOM algorithm is not necessary when applying the algorithm. Therefore, SOM algorithm is suitable to solve complex problems whose mechanism of actions is still unknown. As for the securities companies, there are many variables in customer classification. The securities companies do not know the relationship between these variables and also know little about the influence of these variables on the value and behavior of customers. SOM is more suitable because of the complexity of customer classification of the securities companies.

4.1 Modeling of SOM neural network

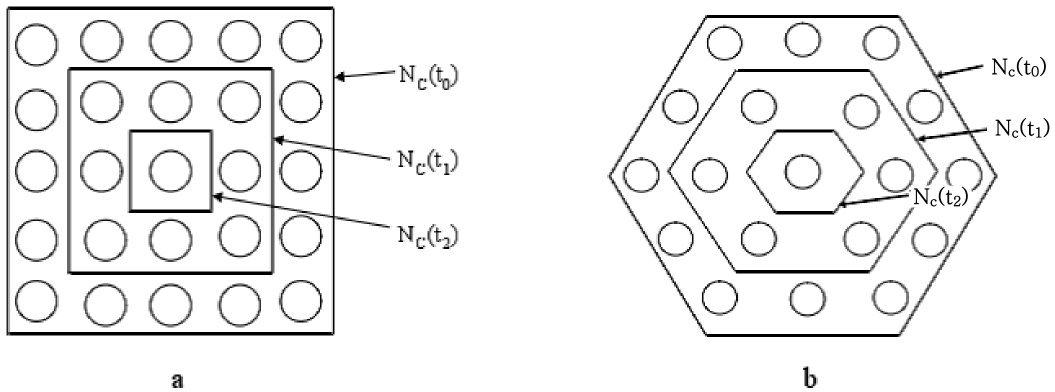
SOM consists of an input layer and an output layer. Every nerve cell in the input layer connects with every nerve cell in the output layer by the weights. Figure 1 below shows the connection between the input layer and the output layer.

Figure 1. The input layer and the output layer



SOM neural network is a self-organizing learning network which can continuously modify the weights during the learning progress and form a centre of excitation. The nerve cells around the centre of excitation have varying degrees of excitation and the nerve cells beyond the centre of excitation are in inhibitory state. The shape of the region constructed by all nerve cells with varying degrees of excitation can be square or hexagon. $N_c(t)$ is used to denote the region.

Figure 2. The shape of the excited region



The two graphs show the changing process of $N_c(t)$ which is becoming smaller as the increase of t and $N_c(t)$ will stop changing when there is only one nerve cell or a group of nerve cells. The last remaining nerve cells represent the feature of the type of the input samples.

SOM neural network determines the final output according to the comparison selection mechanism. Given that the input is $X \in R_n$ and the weight of the connection between the input cell and output cell is $W_i \in R_n$, the output of nerve cell i is O_i which satisfies $O_i = W_i \times X$. The last output is decided by the "winner-take-all" principle. So the real output is $O_k = \max\{o_i\}$. SOM network classification searches for an output nerve cell for every input nerve cell.

4.2 Learning method of SOM neural network

During the learning process, SOM neural network uses Euclidean distance $d(X, w_j) = \| X - w_j \|$ to measure the matching degree of the input vector X and the weight vector $W_j(j = 1, 2 \ldots n)$. SOM searches for a perfect match of the input vector and the weight vector after learning about all input samples in order to get a winning-nerve cell. The nerve cell matching the input vector X best is denoted as C,

Then

$$d(X, w_c) = \min_j d(X - w_j).$$

That is

$$\| X^g - w_c \| < \| X^g - w_j \|, \forall j < m, j \neq c.$$

In our simulation, we first construct a SOM neural network and randomly array with 8×3 dimension based on the initial weights between the input layer and the output layer. Then, let SOM modifies the weights through self-learning. When the feature mapping does not change after 500 times of self-organizing learning, the learning process finishes.

5. Simulation and customer analysis by SOM neural network

In this section, the paper simulates the customer classification by SOM neural network in terms of the modified RFM model. To simplify the analysis, we only choose five primary varieties: the A-shares, the B-shares, treasure bonds, Exchange Traded Funds and funds.

The weight of the A-shares is set as one and the weights of other four varieties are the ratio of their commission to that of the A-shares.

The weights of the five varieties are listed in Table 1.

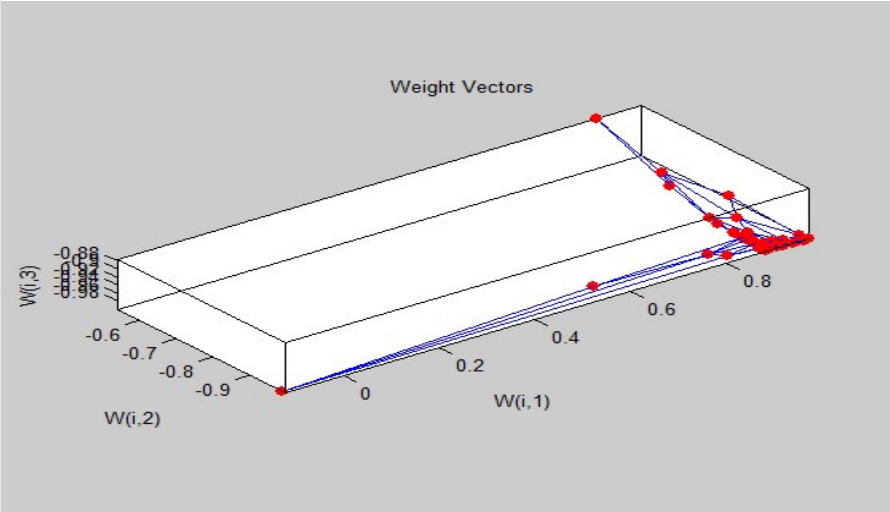
Table 1. The weights of varieties

	No.1	No.2	No. 3	No. 4	No. 5
Varieties	A-shares	B-shares	treasure bonds	Exchange Traded Funds	funds
Commission	4.6‰	9.6‰	0.28‰	0.37‰	0.31‰
Weights	ω_1=1	ω_1=2.086	ω_3=0.061	ω_4=0.08	ω_5=0.217

First, we select 5,000 customers and collect data of their time of the last trade, frequency and accumulative amount of volume during 44 trading days to calculate the modified R, F and M. Then, we use the modified R, F and M of the 5,000 customers to form an array with dimensions 5,000×3 as the input signal of the simulation.

We choose Matlab SOM toolbox to simulate the classification. The first step of simulation is to build a SOM neural network by “net” function and set the output layer to be 6×6 framework. The second step of simulation is to train SOM neural network by “train” function. We train the network 500 times until the weights between the input layer and the output layer do not change. Figure 3 is the framework of the trained SOM neural network.

Figure 3. The framework of the trained SOM neural network



The third step of simulation is to input the sample vectors to the trained SOM neural network. Finally, SOM neural network outputs 8 categories according to comparison selection mechanism. The results of the simulation are listed in Table 2.

Table 2. The results of the simulation

Array Editor - som_num

Stack: Base

	1	2	3	4	5	6	7	8
1	<1x114 double>	<1x168 doubl...	<1x469 doub...	<1x364 doub...	<1x812 doubl...	<1x346 dou...	<1x273 doubl...	<1x2454 doub...
2								
3								

Each cell in Table 2 shows the number of customers of the type. Double clicking the cell displays the corresponding customer types.

We calculate the means of R^*, F^*, R^* of each category and that of all customers and compare the means of R^*, F^*, R^* of each category with that of all customers. $\uparrow$denotes the result that the means of each category is more than the means of all customers. $\downarrow$denotes the opposite result.

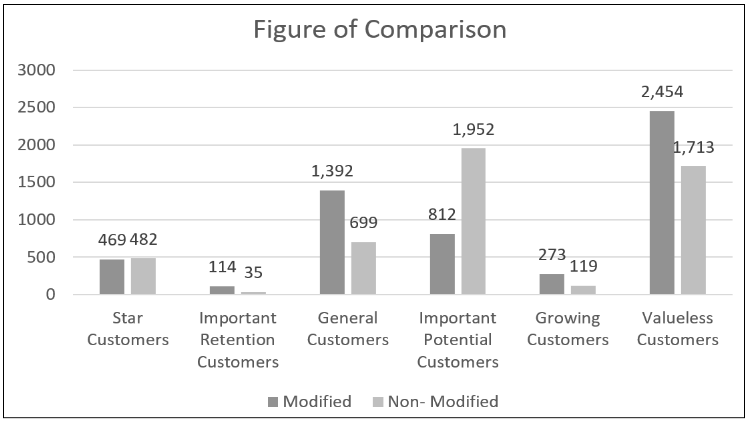
The results of comparison are listed in the third column in Table 3. We summarize the characteristics of these categories and develop strategies for them listed in Table 3.

Table 3. Classification and strategies

Number	The Number of Customer	The Result of Comparison	The Description of Customers	Strategy	Classification
1	114	R*,↑ F*,↑ R*↑	They are valuable customers. But the company faces the risk of losing these customers since they have not traded for a long time.	The company should maintain close tie with these customers and take measures to keep them.	Important retention customers
2	168	R*↑ F*↑ R*↓	No transaction for a long time. Low volume with high frequency.	Regular service	General customers
3	346	R*↑ F*↓ R*↑	No transaction for a long time. Low frequency with high volume.	Regular service	General customers
4	2454	R*↑ F*↓ R*↓	No transaction for a long time and little volume.	No specific service	Valueless customers
5	469	R*↓ F*↑ R*↑	Transaction in recent time. High frequency and great volume.	Build up their loyalty for the company and motivate them to become lifetime customers.	Star customers
6	346	R*↓ F*↑ R*↓	Transaction in recent time. High frequency with low volume.	The transaction volume may be volatile and hard to predict. Just provide regular service.	General customer
7	273	R*↓ F*↓ R*↑	These customers may contribute more profit if the frequency rises in the long run.	Provide professional investment advice to the customers regularly.	Growing customers
8	812	R*↓ F*↓ R*↓	These customers may be new to the companies. They are important for companies to expand the market share.	To understand their preferences and provide customized service.	Important potential customers

In order to evaluate the modified RFM model, we make a comparison between the result by RFM model and that by the modified RFM model. Figure 4 below shows the result of comparison.

Figure 4. The result of comparison



The number of customers who are classified as valueless customers according to the modified RFM model is greater than that according to non-modified RFM model. It is the same with the category of general customers. The comparison result reflects the modified RFM model is stricter than the non-modified RFM model as customers who contribute little commission are seen as valueless customers despite their recent trading. The number of customers who are classified as important potential customers according to the modified RFM is less than that according to non-modified RFM model. The reduction of the important potential customers will be helpful for the company to focus on developing potential valuable customers. The numbers of star customers are approximately the same in terms of the two models, which reflects that the modified RFM does not "make mistakes" of treating valuable customers as worthless customers.

Besides, the percentage of valueless customers and general customers who both create little profit for the securities companies is 76.92% according to the modified RFM model and that according to the non-modified RFM it is only 48.24%.This proves that the classification of the modified RFM is in line with Pareto Principle which says 20% of the customers create 80% of the profit.

6. Conclusion

The paper simulates the customers' classification of the securities companies by SOM neural network based on the modified RFM model. The results prove that the modified RFM is more precise in customer analysis. So, the classification results can support customer centered service and professional marketing. But, SOM neural network does not introduce any prior information into the customers' classification. Customers' historical information has not been fully used. Furthermore, the long time of convergence of SOM neural network will lead to low efficiency of the classification.

Another question needed to be further discussed is to develop the strategies for different customers. Especially for the customers who may be lost, early warning model should be established.

References

Bin, L. and Jindong, Z. (2013). Marketing strategies analysis of commercial bank based on RFM model. *Statistic and Decision, 29*(14), 65-67.

Chenbo, C., Liang, B., Zhouyue, B., Linxi, Q. and Zhao, Y. (2004). Application of SOM neural network in customer classification systems. *Engineering-Theory & Practice, 24*(3), 8-14.

Chiungi, C. and Juichih, H. (2017). A Two-layer clustering model for mobile customer analysis. *IT Professional, 19*(3), 38-44.

Huan, Z., Guangming, L. and Gaoyu, Z. (2010). SOM+K-means two-phase clustering algorithm and its application. *Modern Electronics Technique, 14*(16), 113-116.

Moeini, M. and Alizadeh, S. H. (2016). Proposing a new model for determining the customer value using RFM model and its developments (case study on the Alborz Insurance Company). *Journal of Engineering and Applied Sciences, 11*(4), 828-836.

Sheng, L. and Xu, X. (2006). A method of telecom consumer market segmentation, *Journal of Harbin Institute of Technology, 38*(5), 758-760.

Xiaoyan, Z. and Chunying, Q. (2008). Securities clients analysis system based on data mining. *Computer Application, 28*(6), 369-371.

Ziyi, L. (2014). The research on the customer management of securities company. *Journal of Liaoning University of Technology (Social Science Edition), 16*(3), 23-26.