

The Use of Customer Value Changing Trends in Business Analysis

Sorush Niknamian

Military Medicine Department, Liberty University, USA

Abstract

With the development of competition in business and also today's rapid changes in competitive market, understanding these changes are the key factor of being more efficient in markets. CRM which is known as basic structure for describing customer's needs for efficiently understand customer's behavior and finely get the maximum of market share and profit. There are major differences between B2B and B2C businesses such as long term purchase cycle, purchase interests and the amount of the transactions. These differences needs more interactive strategies. The knowledge that gets from CRM is extremely related to market changes. In recent years data mining increasingly help organizations to get and understand customer's behavior. But with the rapid changes in market these procedures must be change too. Change mining as the higher order of data mining tries to get knowledge by analysis patterns instead of data. In this paper we attempt to calculate customer's value by using RFM model and K-MEANS clustering method and then analysis changes in deferent time periods. We tries to find out cluster transitions and most frequent customer value changing trend for proactive decision making. For this purpose we use customer purchase transactions in insurance industry which are gathered in 3 years.

Keywords: Data mining, customers relationship management, change mining, clustering, segmentation, RFM model.

1. Introduction

With the development of customer-oriented behavior in business, growing attention has been paid to customers and their needs as one of vital factors to gain higher profits (Cheng & Chen, 2009). Customer relationship management (CRM) seeks to identify customer needs and facilitate the interaction between customers and businesses (Ling & Yen, 2001). Given the changing customer behavior and needs in today's market, businesses must make decisions in keeping with these changes. Most researches on customer value and clustering-based segmentation do not consider changes over time and there is a paucity of researches on customer behavior based on association rules (Böttcher, Spott, Nauck, & Kruse, 2009; Chen, Chiu, & Chang, 2005; Huang, 2012). The integration of CRM and analytic structures like data mining has been one of the important issues in recent years. This vision can help businesses develop new strategies or verify and correct their current strategies (Bose & Chen, 2009; Davenport, Harris, & Morison, 2010). Nowadays, With the application of data mining technology in CRM, techniques like decision trees, clustering algorithms, genetic algorithms, and association rules in different areas like commerce have been used to solve customer problems and formulate new strategies (Berson, Smith, & Thearling, 1999; Bramer, 2013; Turban, Sharda, Aronson, & King, 2008). In this research, attempts have been made to apply clustering techniques and mining changes to customer value over time to bring a new approach to the customer segmentation. The rest of this paper is organized as follows: definition of customer relationship management is provided in Section 2. Section 3 offers a review of customer value analysis. Temporal data mining defined in section 4. Our proposed model is presented in Section 5, the case study based on proposed model is represented in section 6 and the conclusions are in Section 7.

* Corresponding author.

E-mail address: xxxx@xxxxxx.edu (First Author)



2. Customer Relationship Management

Over the past few years, companies interact with customers has changed considerably and customers can change their providers easily than before. Because of this, Organizations must properly recognize the customer and anticipate their needs. Customer relationship management, let organizations to have better understand of their customers and the differences among them. CRM is a comprehensive strategy and process of acquiring, retaining and partnering with selective customers to create superior value for the company and the customer (Parvatiyar & Sheth, 2001). Customer retention is key factor in this strategy. Feinberg and Kadam, profits increase by 25–80% when customer retention rates increase by five points (Feinberg & Kadam, 2002).

3. Customer Value Analysis

Customer value analysis is an analytical structure for interpreting customer behavior from the vast source of otherwise meaningless data. There are different definitions for customer value. Kotler (2001) defines it as the profit of net present value (NPV). Hwang et al. (2004) describes customer value as the sum of revenues gained by customers over lifetime of transitions. Thus, customer values are based on past and potential profit as well as defection probability. Later, Kim et al. (2006) proposed a new customer value model which was based on current value, potential value and loyalty. One of the most effective customer segmentation models based on customer value is the RFM model (Verhoef, Spring, Hoekstra, & Leeflang, 2003), which was introduced by Bauer (1988) and later developed by Hughes (1994). In recent studies, the RFM model has been adopted in different industries and under various conditions by adding extra parameters. Stone (1995) introduced RFM model with unequal value for each R, F, and M parameter. Tsai and Chiu (2004) attempted to weigh the parameters based on the characteristics of industries. Liu and Shih (2005) proposed WRFM, which is the defining weight for each parameter, using the analytic hierarchy process (AHP). Chang and Tsay (2004) proposed LRFM model that incorporates customer retention length (L) to the RFM model. Chang & Tsai (2011) tried to incorporate the characteristics of purchased products into their analysis by introducing GRFM (group RFM). Later Hosseini et al. (2010) included another factor called “period of product activity” in the RFM model (RFML) for vehicle manufacturing industries.

4. Mining Changes over Time

Data mining is the statistical analysis of data based on the information of models, which is appropriate for the time period. Recent studies suggest that data changes over time and therefore the results would not be helpful (Roddick & Spiliopoulou, 2002). Today’s world is in a constant state of flux which makes the previous results out-of-date. The new approach is to mine these changes over time periods. In this approach, time is split into intervals $T = \langle t_0, \dots, t_n \rangle$ where for each t_i there is a M_i model. These models are based on statistical methods of data mining. Here, the main purpose is to analyze the evolution of $\langle M_1, M_2, \dots, M_n \rangle$ that creates the final model M (Böttcher, Höppner, & Spiliopoulou, 2008). In this context, most approaches are in association rules. Chen et al. (2005) attempted to integrate customer behavior variables, demographic variables and transaction database to present a method of mining changes in customer behavior. Böttcher et al. (2009) presented a system of customer segmentation based on the discovery of frequent item-sets and the analysis of their change over time. Hu, Huang & Kao (2013) used RFM analysis in the sequential pattern mining process. As can be seen, there is a paucity of studies on customer value analysis based on clustering techniques.

Clustering is one of the key method in pattern recognition and machine learning. Over the past decades, many studies have been done on these methods. Monitor the cluster evolution is very important in many real-world applications. The changes that found in clusters over time, can be associate with real world changes. For this reason, tracking cluster evolution over Time can be useful in many areas, including marketing, fraud detection and economy. The cluster evolution, can be defines into two different types, internal and external transition. In this study we working on external transition of clusters.

4.1. Cluster transition types

There are different types of cluster transition, but based on cluster transitions that represent in literature (Oliveira & Gama, 2012; Spiliopoulou, Ntoutsis, Theodoridis, & Schult, 2006), the main external transitions are Birth, Death, Split, Merge and Survival. Table 1 bring more details about these transitions.

Table 1 Main Types of Cluster External Transition

Transition Type	Mathematical Notation	Description
Birth	$\emptyset \rightarrow Y(t_2)$	From time interval t_1 to t_2 new cluster Y emerge
Death	$X(t_1) \rightarrow \emptyset$	From time interval t_1 to t_2 existing cluster X removed
Split	$X(t_1) \rightarrow^c \{Y_1(t_2), \dots, Y_p(t_2)\}$	Cluster X in time interval t_1 split into p cluster Y_1 to Y_p in time interval t_2
Merge	$\{Y_1(t_1), \dots, Y_p(t_1)\} \rightarrow^c X(t_2)$	P cluster Y_1 to Y_p in time interval t_1 are merge and make cluster X in time interval t_2
Survival	$X(t_1) \rightarrow Y(t_2)$	From time interval t_1 to t_2 cluster X survive as cluster Y

All of the transitions in table 1 are based on the minimum threshold survival τ and minimum threshold split λ (Spiliopoulou et al., 2006). For example in Death transition if members of cluster X in time interval T_1 less than minimum threshold τ_{Death} in time interval T_2 then we found that a Death occurred.

4.2. Interpretation of Cluster Transition in Business

We defined 5 external cluster transition types in section 4.1, the meaning of these transitions and the strategies based on them, can be useful as proactive strategies to businesses. For example if a split occurred, that means a group of our customers are going to make new cluster with differences value, business behavior characteristics and so on.

Based on these result, businesses can define different strategies such as because of resulting clusters are more manageable and have more homogeneous characteristics, we let this split be stable, otherwise we make proper strategies to get back these clusters together, because this split made group a of customers that are going to churn. Other cluster transitions can be describe like above. In the section 5 we will propose new model by using these cluster transitions.

5. Propose Model

As we mentioned before, to have more proactive decision making system we have to use models that run over time interval and mine the changes over these time intervals. In Fig 1 we propose new model based on RFM customer value calculator model combined with cluster transitions in order to cover customer value changing trends.

All procedures are described separately below:

Step 1- **Data preparation and cleansing:**

In step 1, the data is cleaned by removing irrelevant values, noisy and incorrect data as well as the fields of data inappropriate to our research and we prepare data structure for implementing RFM analysis like choosing R, F and M parameters.

Step 2- **Incorporating data into time intervals:**

As mentioned earlier, we try to include time into the RFM model. For this purpose, we split the time period into z equal time intervals (equation 1).

$$T = \sum_{i=1}^{i=z} T_i, T = \{T_1, T_2, \dots, T_z\}, |T_1| = |T_2| = |T_z| \quad (\text{equation 1})$$

Step 3- Implementing customer value calculation model (RFM model) for each time interval:

After preparing data in step 3, customer segmentation based on one of the customer value calculation model (RFM model) and data mining clustering technique like K-means technique (Han 2005) were implemented. Because of using RFM model in Current model. To control the effect of each parameter on other parameters, we normalize R, F and M parameters between (0, 1) (Bramer, 2013). Here, because of the independency of time intervals in analysis, each time interval must be normalized separately. If the whole time is normalized and then split into time intervals, there would be some values that their influence exceeds time intervals and affect the normalization of whole data. Given the importance of the independent of analysis of each time interval, this will be inaccurate.

Step 4- finding cluster transitions in whole duration

Based on provided definitions in section 4.1 and results of step 3, external transitions based on minimum threshold λ and τ are known. Based on these transition and the data with them companies can made proper decisions.

Step 5- Determining most frequent customer value changing trends

After implementing customer value calculation in step 3, now we have value changing trend for each customer. By using these trends we can found the value changing attitude of our customer and what the most frequent customer changing trend is. Fig 2 present more information.

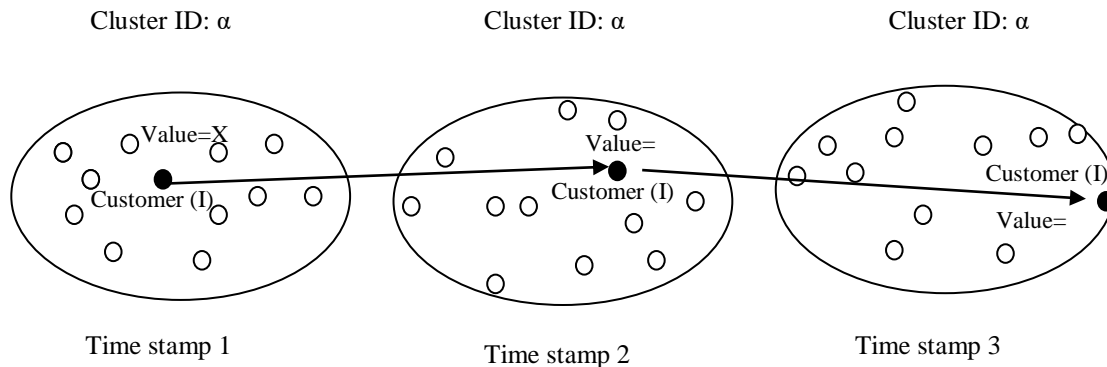


Figure 2 customer value changing trend

Fig 2 shows the value changing trend of customer (I) in 3 time stamp that in time stamp 1 our customer have the value of X in time stamp 2 its value change to Y and in time stamp 3 the customer (I) have value of Z. now we can show that our customer value changing trend among these 3 durations are : $X \rightarrow Y \rightarrow Z$ based on increasing or decreasing attitude of this trend we can found the status of our customer, is it going to be more profitable? Or it is churning our business.

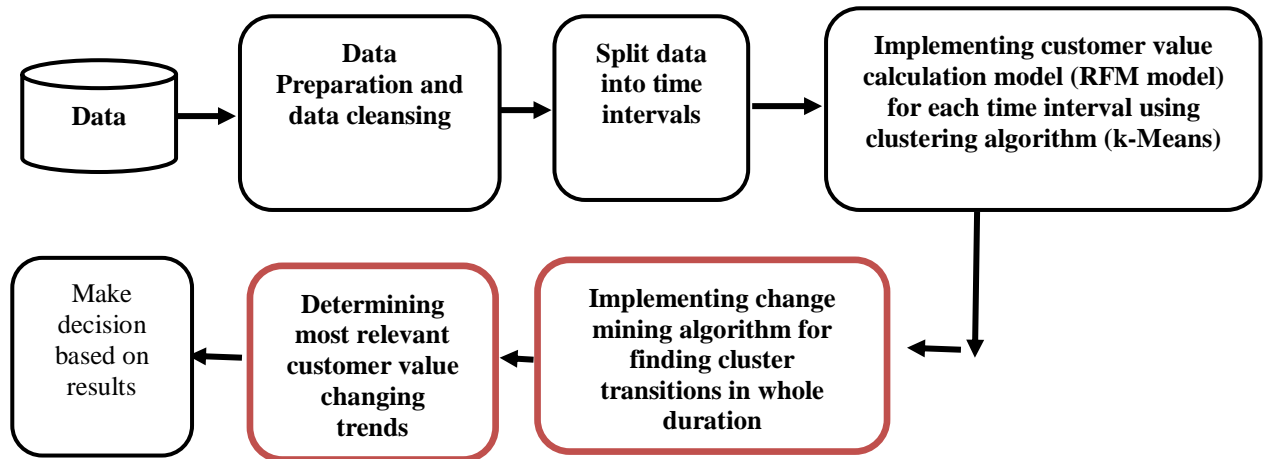


Fig. 1 Purpose Model

6. Empirical Implementation

Our data consist of patient’s prescription of an insurance company. Table 2 represent data fields

Table 2 data Field Description

ID	Field name
1	Patient ID
2	medicines
3	Price of medicines
4	Prescriber Doctor
5	Pharmacy Name
6	Date of prescription

Based on propose model in section 5 the implementation steps are describe below:

1. First, the data is cleaned by deleting irrelevant values, noisy and incorrect data as well as the fields of data inappropriate to our research. Then, the data is organized in such a way to extract parameters of customer value calculator model that in this paper is RFM model. Table 3 shows the data ready for analysis.
2. As we mentioned for finding cluster evolutions we must divided data into time intervals. Here we divide data into 4 time stamp that each time stamp includes 3 month. Because of having different durations we add new field to our data as duration that indicate duration of analysis. Table 3 shows the final data ready for analysis.
3. We use K-means data mining technique to cluster customers in each time period. Table 5 shows the results of customer clustering for two consecutive time periods. The important point is that the cluster number is similar in two time intervals. Also, given the independent analysis of each time interval, there is no relationship between them.
4. Based on the results in step 3 we can find cluster transitions in whole time intervals. For instance, Table 5 shows survival of cluster 1 in whole time intervals. In table 5, cluster 1 in duration 1, survived to cluster 5 with overlap on 0.89 which means more than 89% of cluster 1 customers in time interval 1, survived into cluster 5 in time interval 2. This stability means that the customers inside cluster 1 are strongly homogeneous. Other transitions of this cluster can be seen in table 5.

Table 3 Data ready for analysis

R	F	M	Patient ID	DURATION
17	2	20150	6624327	1
1	2	18030	67832258	1
8	3	94280	6807114	1
3	1	19300	6837796963	1
8	1	8050	68500644	1
27	2	42200	6860787	1
8	1	22880	69514032	1
10	2	34310	69536751	1
1	1	43300	69545023	1
37	1	14250	6960138	1

M: Unit: IRR

R: Unit: Days, interval between customer’s last purchase and the last day of the time period

F: Unit: number of items

Table 4 Customer cluster Number for 2 consecutive time intervals

Cluster ID in second time interval	Cluster ID in first time interval	Customer ID
1	1	6624327
2	1	67832258
4	3	6807114
5	2	6837796963
1	4	68500644
1	5	6860787

Table 5 Survival of Cluster 1

	duration1	duration2	duration3	duration4
cluster1	1	5	4	2
value	0.36	0.5	0.79	0.43
overlap		0.89186	1	0.53204

Another example of cluster transition in our data is represent in table 6. In table 6, in duration 2 a split and in duration 3 a merge occurred. This means that customers in cluster4 are split into 2 different clusters with different characteristics. But based on company strategies or Environmental conditions in duration 3 these customers get together again. Suppose that when split occurred in duration 2 one of the two cluster 2 or 5 are customers with specific characteristics that company want to recognize them. But in the next duration these customers merge into others and can’t be identifies. With our model companies can identify and manage these groups and let them to make new segment of manage them into get back into their previous segment. Fig 3 shows the bipartite graphs for Table 5 and Table 6 examples.

Table 6 Transition of Cluster 4

	duration1	duration2	duration3	duration4
cluster4	4	2,5	4	2
Value	0.52		0.79	0.43
overlap			0.9,1	0.53204

In fig. 3 black lines show overlap of transitions in example table 5 and the orange lines show overlap of transitions in example table 6.

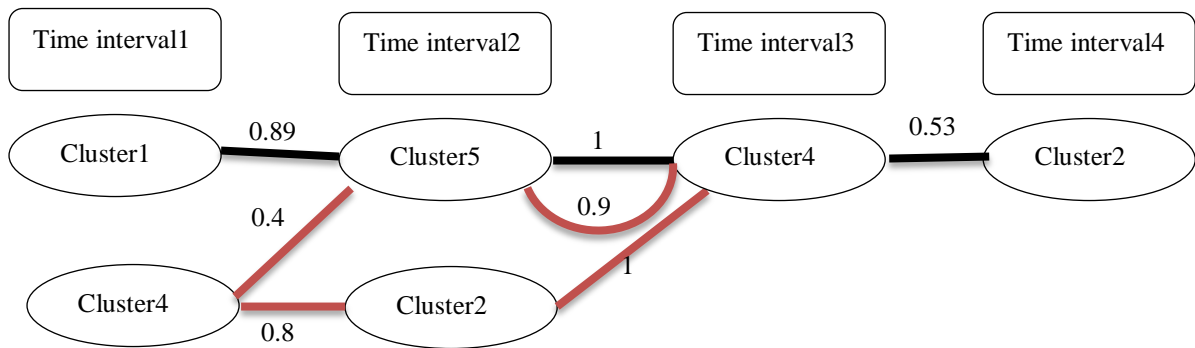


Figure 3 bipartite graphs for Table 5 and Table 6 examples

5. Based on results of step 3 we also can indicate most frequent value changing trend of customers.

Table 7 shows top frequent customers changing trend in 4 durations.

Table 7 most Frequent customers value changing trends

Trend among clusters	Number of customers with this trend	Value changing trend
1->4->4->4	1249	0.36->0.53->0.79->0.42
1->5->4->4	1103	0.36->0.505->0.79->0.42
1->5->4->3	929	0.36->0.505->0.79->0.35
1->4->4->1	922	0.36->0.53->0.79->0.12
1->4->4->3	913	0.36->0.53->0.79->0.35

In table 7 we can see most frequent trends of our customers in 4 duration. This information can be useful because we can find out what is most frequent trend among customers and based on that we can propose different strategies. Fig 4 shows value changing trend in table 7.

7. Conclusion

Change mining as mining the changes in data over time can be one of the most effective methods in today’s fast moving businesses. In this paper we introduce the use of change mining in customer segmentation by split data into time intervals and analysis them separately. The process of monitoring customer segments changing over time help us to deal with changes over time and make relevant strategies based on these changes. As future work we try to discuss the shape of survived clusters as internal cluster changes which means, how characteristics of survived clusters change

over time.

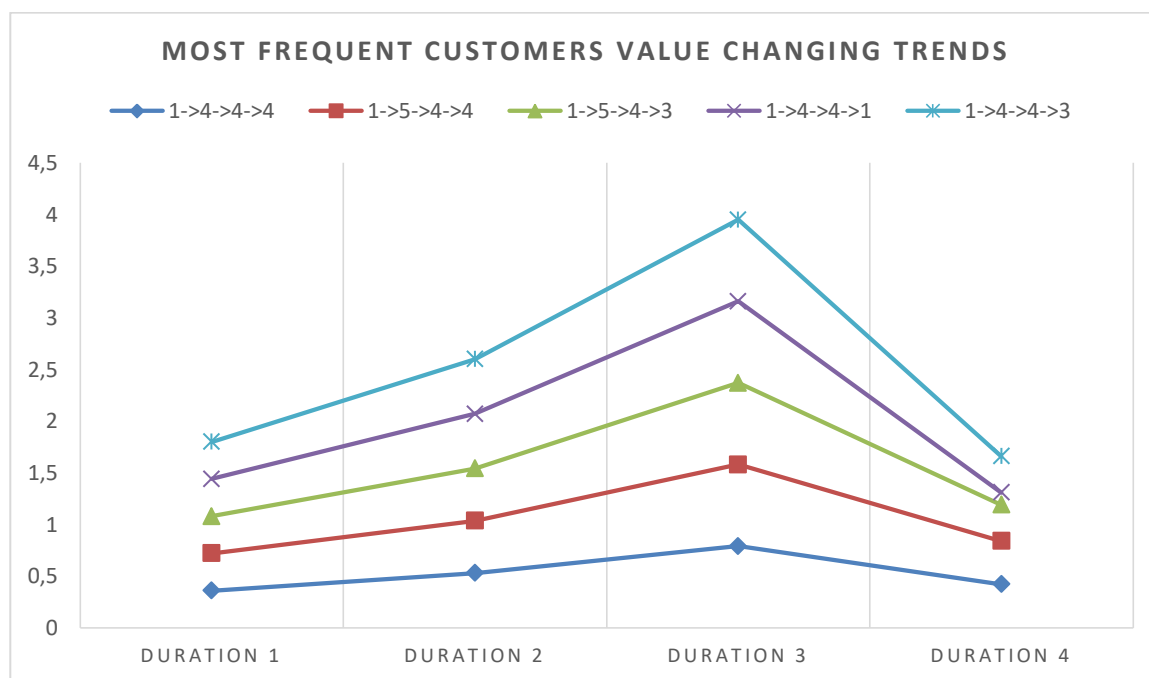


Figure 4 most Frequent customers value changing trends

References

- Bauer, C. L. (1988). A direct mail customer purchase model. *Journal of Direct Marketing*, 2(3), 16–24. doi:<http://dx.doi.org/10.1002/dir.4000020305>
- Berson, A., Smith, S., & Thearling, K. (1999). *Building Data Mining Applications for CRM* (1st ed.). McGraw-Hill Professional.
- Bose, I., & Chen, X. (2009). Quantitative models for direct marketing: A review from systems perspective. *European Journal of Operational Research*, 195(1), 1–16.
- Böttcher, M., Höppner, F., & Spiliopoulou, M. (2008). On exploiting the power of time in data mining. *ACM SIGKDD Explorations Newsletter*, 10(2), 3–11.
- Böttcher, M., Spott, M., Nauck, D., & Kruse, R. (2009). Mining changing customer segments in dynamic markets. *Expert Systems with Applications*, 36(1), 155–164.
- Bramer, M. (2013). *Principles of data mining*. Springer.
- Chang, H. H., & Tsay, S. F. (2004). „Integrating of SOM and K-Mean in data mining clustering: an empirical study of CRM and profitability evaluation,“. *Journal of Information Management*, 11(4), 161–203.
- Chang, H.-C., & Tsai, H.-P. (2011). Group< i> RFM</i> analysis as a novel framework to discover better customer consumption behavior. *Expert Systems with Applications*, 38(12), 14499–14513.
- Chen, M.-C., Chiu, A.-L., & Chang, H.-H. (2005). Mining changes in customer behavior in retail marketing. *Expert Systems with Applications*, 28(4), 773–781.

- 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.
- Davenport, T. H., Harris, J. G., & Morison, R. (2010). *Analytics at work: Smarter decisions, better results*. Harvard Business Press.
- Feinberg, R., & Kadam, R. (2002). E-CRM Web service attributes as determinants of customer satisfaction with retail Web sites. *International Journal of Service Industry Management*, 13(5), 432–451.
- 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.
- Hu, Y.-H., Huang, T. C.-K., & Kao, Y.-H. (2013). Knowledge discovery of weighted RFM sequential patterns from customer sequence databases. *Journal of Systems and Software*, 86(3), 779–788.
- Huang, T. C.-K. (2012). Mining the change of customer behavior in fuzzy time-interval sequential patterns. *Applied Soft Computing*, 12(3), 1068–1086.
- Hughes, A. (1994). *Strategic database marketing*. Probus Publishing, Chicago.
- Hwang, H., Jung, T., & Suh, E. (2004). An LTV model and customer segmentation based on customer value: a case study on the wireless telecommunication industry. *Expert Systems with Applications*, 26(2), 181–188.
- Kim, S.-Y., Jung, T.-S., Suh, E.-H., & Hwang, H.-S. (2006). Customer segmentation and strategy development based on customer lifetime value: A case study. *Expert Systems with Applications*, 31(1), 101–107.
- Kotler, P. (2001). *Marketing Management*, 10. Pearson Education Canada.
- Ling, R., & Yen, D. C. (2001). Customer relationship management: An analysis framework and implementation strategies. *Journal of Computer Information Systems*, 41(3), 82–97.
- Liu, D.-R., & Shih, Y.-Y. (2005). Integrating AHP and data mining for product recommendation based on customer lifetime value. *Information & Management*, 42(3), 387–400.
- Oliveira, M., & Gama, J. (2012). A framework to monitor clusters evolution applied to economy and finance problems. *Intelligent Data Analysis*, 16(1), 93–111.
- Parvatiyar, A., & Sheth, J. N. (2001). *Conceptual framework of customer relationship management*. New Delhi, India: Tata/McGraw-Hill.
- Roddick, J. F., & Spiliopoulou, M. (2002). A survey of temporal knowledge discovery paradigms and methods. *Knowledge and Data Engineering, IEEE Transactions on*, 14(4), 750–767.
- Spiliopoulou, M., Ntoutsis, I., Theodoridis, Y., & Schult, R. (2006). Monic: modeling and monitoring cluster transitions. In *Proceedings of the 12th ACM SIGKDD international conference on Knowledge discovery and data mining* (pp. 706–711). ACM.
- Stone, B. (1995). *Successful Direct Marketing Methods* (pp. 37–57). NTC Business Books.
- Tsai, C.-Y., & Chiu, C.-C. (2004). A purchase-based market segmentation methodology. *Expert Systems with Applications*, 27(2), 265–276.
- Turban, E., Sharda, R., Aronson, J. E., & King, D. N. (2008). *Business intelligence: A managerial approach*. Pearson Prentice Hall Upper Saddle River.
- Verhoef, P. C., Spring, P. N., Hoekstra, J. C., & Leeflang, P. S. H. (2003). The commercial use of segmentation and predictive modeling techniques for database marketing in the Netherlands. *Decision Support Systems*, 34(4), 471–481.