Big Data Analytics toward Intelligent Mobile Service Provisions of Customer Relationship Management in e-Commerce

Luyao Xu[1], and Hai-Cheng Chu[2*]

[1] Department of Bilingual Teaching and Research at Business School, Sias International University

Xinzheng, 451150, Henan, China

xluyao@yahoo.com

[2*] Department of International Business, National Taichung University of Education

Taichung, 40306, Taiwan

hcchu@mail.ntcu.edu.tw

Received 18 June 2015; Revised 5 September 2015; Accepted 6 October 2015

Abstract. The rapid development of wireless technology has accelerated the use of mobile devices and mobile services between customers and business. Hefty amount of data are unprecedentedly generated within a second. Under the big data era, making full use of the benefits of mobile devices will enable the co-creation of value to realize the increased customer retention. Hence, the capability to process large-scale data poses an interesting but critical challenge for contemporary business organizations. It becomes urgent research issues result to efficiently and effectively conducts data mining mechanisms with respect to massive amount of data to meet the organizational strategic objectives. This paper will introduce multi-agent systems and their applications from a data mining aspect, followed by the value of data mining from a Customer Relationship Management (CRM) aspect. Finally, we propose a three-step data-mining model to support mobile service, which can help business organizations to dig out potential value to manage CRM optimally.

Keywords: Big Data, Multi-Agent System, Mobile Service, Customer Relationship Management.

1 Introduction

With the increasing use of mobile devices and development of Internet, search engines hosting diversified streams of information have created unprecedented scattered data. Instead of traditional data-computing devices, massive information requires a superior data-processing technology to achieve high-speed data transferring and integration. Such a wave of information that has overcapacity on data mining and integration compared with conventional database management system is big data.

Big data is also called "massive data". Big data processing has higher requirement on speed. Big data has four main features: Volume, Variety, Velocity, and Value. People are eager to look for regular characteristics through examining all relevant data. During the big data era, everyone is a big contributor of data. The expansive resources of data include, but are not just limit to the following: social media, such as Weibo, WeChat, email, video, audio, web scanning, GPS, traffic monitoring systems, and other media. Users are expanding their digital footprint and creating more data whenever they use these media. Personal information, consuming habits, preferences, and even relevant social networking will be identified under big data era. Undoubtedly, such a data collection and mining are beneficial for business through providing valuable clues for future strategic management.

Customers are expanding their online footprints extensively, which makes it hard to extract data value through data collection and data mining. Due to the distributed databases embedded based on heterogeneous platforms, business organizations are facing problematic challenges. How to continue data mining accurately has challenged enterprises to probe for new technologies with stronger processing power to find the value that hidden inside data. The Multi-Agent System (MAS) makes it easier to retrieve the most suitable results for users by processing massive data into an explicit reasoning engine. A M.A.S. is a computerized system composed of multiple interacting intelligent agents within an environment. Comparing with Intelligent Agent (IA), M.A.S. accelerates the cooperation and communication between IAs as the information processor. Users can retrieve accurate target information in a timely manner via the assistance of new technology.

Enterprises can have a better understanding of customers through integrating and analyzing various data such as historical transaction records, geography, frequency of web scanning and other data. Retail can use big data processing and analysis to achieve the demand forecasting, price and merchandizing optimization; Manufacturing can use big data analysis to work towards product customization, new product development, and supply chain management; Medical agencies can use data processing to implement disease management and preliminary diagnosis; Government can use data analysis to achieve crime prevention, fraud detection, and revenue optimization; Business can customize service or product with the highest return on investment through targeting on potential customers accurately. All of these situations stated above can be attributed to benefits brought by effective Customer Relationship Management (CRM) through big data mining. It is no wonder that effective CRM is exceedingly important for any business based on its potential economic benefits.

Some literature suggests that CRM is understood as a customer-oriented management approach in which information systems provide information to support operational, analytical, and collaborative CRM processes and thus contribute to customer profitability and retention [3]. Various data resources are mushrooming. Collecting customers' information and analyzing the information using data mining techniques are the primary processes of CRM. Currently, commonly used data mining techniques include: classification, outlier analysis, k-means algorithm, regression analysis, and clustering. Also, models using Bayesian networks, decision trees, artificial neural networks and association rules have been advocated in many enterprises to serve for CRM through analyzing valuable customers' information. On the basis of these achievements, this paper will focus on the discussion about another data mining model to extract data value effectively. This data mining approach has three steps including using K-means to cluster massive data. In addition, we generalize data to focus on relevant attributes via using a user-defined threshold for the information loss calculation, and information gain ratio calculation method to make decision trees for extracting potential valuable knowledge purpose.

2 Related Works

2.1 The Cooperation Mechanism of Multi-Agent Systems (MAS)

M.A.S. is a computerized system composed of multiple interacting intelligent agents within an environment. The intelligent agent is a part of Distributed Artificial Intelligence (DAI), which carries out a goal-directed behavior through sensors-based observation and actuators-orientated action as the information processor. By means of contemporary interrelated networking systems, an IA is able to communicate, collaborate, or even to negotiate with other IAs with semantic reasoning capability [2, 5, 6, 9]. M.A.S. has improved a lot on data transferring and integration compared with IA. M.A.S. can accelerate the cooperation and communication among different intelligent agents [27]. For M.A.S., how to ensure the autonomous characteristic of every single IA as a coherent behavior has become a challenge. Based on the related researches, social autonomy is believed to be one of the most important behaviors concerning the interactions among the agents in M.A.S. [10, 18]. In other words, adoption of goals is a reflection of social autonomy. On the contrary, it is hard for individual IA to integrate massive data in a timely manner to provide users the most useful information accurately based on its limited capability. Fortunately, the M.A.S. can be the intelligent director as the information processor due to its capability of directing the cooperation or negotiation between different IAs [26].

In other words, the M.A.S. can call for different IAs to work together for problem solving in an efficient way. One IA can delegate its task or responsibility to the other IAs based on cooperation purpose or specialized function. Also, IA can choose its intelligent partner based on interactive platforms and databases support. Of course, these IAs act like a real problem-solving team. There maybe some conflicts between IAs during this cooperation process. It is time-consuming to deal with negotiation between different IA platforms. One intermediary system that acts as the communicator and decision maker among IAs is exceedingly important and meaningful under the big data era. Luckily, the M.A.S. acts as such a role to direct IAs to cooperate and negotiate with each other to provide users the most useful information in a timely manner. Under the big data era, high accuracy and high efficiency are main points for data mining technology.

2.2 Big Data and Data Mining in CRM

With the increasing use of mobile devices, people are expanding their footprints extensively, from web scanning, keywords input in search engines and click stream within the site, to online transaction, online video, and point of entry to the site. Such a massive data requires a superior data-processing technology to achieve high-speed data transferring and integration.

Intent here, firms have been investing billions of dollars in order to enhance the hardware and software platforms for mobile commerce to create more personalized services for customers. Mobile applications that target a service provision will be in line with the core competency of business. However, literately, time criticality has been identified as a value-added characteristic of the mobile channel [20, 22]. Consequently, mobile services under the big data era has put higher requirement on information processing speed. A wave of information that has overcapacity on data mining and integration compared with conventional database management system is big data. The big data era advocates higher requirements on data integration and analysis, which push the development of more advanced data analysis technologies. The process of extracting valuable information through probing for regular patterns of data based on effective data analysis technologies is data mining.

Data mining is a computational process that aims to process uncertainty through discovering regular patterns from amounts of data based on different technologies such as artificial intelligence, statistics and other methods. In addition, from the CRM aspect, based on customer's changing need and habit behaviors, how to extract useful data to identify target customers has become the main point of business. For business, data mining can be regarded as the technology that can be used for categorizing customers based on different attributes such as purchasing behavior, attitudes to serve for future effective business strategy and CRM strategy [13].

As far as the CRM in business, how to set up a customer profile through finding patterns in a customer database is the key point. Businesses can achieve targeted marketing strategies through targeting of specific promotions to existing and potential customers. Also, businesses can continue market-based analysis accurately. With data mining, retailers can determine which products to stock and how to display them within a store. In addition, businesses can hold customer retention by determining characteristics of a customer who is likely to leave for a competitor. Furthermore, some financial institutions such as banks and insurance company can achieve fraud detection through identifying potentially fraudulent transactions. The achievement of the above purposes cannot survive without effective data mining functions. The commonly used data mining methods include，but are not just limited to classification, clustering, regression, association rules, the abnormal detection and other methods or models.

3 Research Analyses

3.1 Data Mining Model from CRM Respect

The mushrooming use of multimedia will generate more information, include both structured and unstructured data, and these data will continue to increase the exponential growth. Diverse data resources and nature has made it to be a challenge to analyze these data for extracting potential useful information. Collecting customer information and analyzing the information using data mining techniques are the primary processes of CRM.

CRM optimizes profitability, revenue, and customer satisfaction by organizing around customer segments, fostering customer-satisfying behaviors, and implementing customer-centric business models [17]. CRM is understood as a customer-oriented management approach in which information systems provide information to support operational, analytical, and collaborative CRM processes and thus contribute to customer profitability and retention [3]. On the other hand, consumer behavior analysis is concerned with the study of interactions among the consumers, products and operations such as purchasing, brand choice, etc. How to analyze these data through clustering, integrating, and modeling is the main concern of this paper.

In view of the importance of CRM in business, this paper will propose a model to elaborate its contributions for business. This model is a three-step process, including using K-means to cluster massive data; Generalizing data to focus on relevant attributes via using a user-defined threshold for the information loss calculation; Calculating information gain ratio to make decision trees for extracting potential valuable knowledge purpose.

K-means Algorithm-Based Clustering

Clustering is a process for grouping a set of objects into classes (or clusters) to find groups of customers with similar behavior within a group, but is very dissimilar to objects in other clusters. Clustering methods proposed in different literature include: hierarchical clustering, divisive clustering, density-based clustering with DBSCAN algorithm, grid-based clustering, and model-based clustering. In this paper, for our concern, we propose to calculate the similarities within a group using the equation (1) shown below.

$$d(x_a, x_b) = \sqrt{\sum_{m=1}^{d}(x_{am} - x_{bm})^2} \, . \tag{1}$$

As one of the most typical clustering methods, the goal of K-means is grouping data into K clusters based on parameter K and target data. After picking up several data randomly, we can image that each data stands for one average number of one cluster. Then, grouping other data based on their distance to these identified average number. Again, repeat this process to get a new average number for each cluster, until the criterion function is converged to get the best clustering result. Compared to other clustering methods, K-means method is more efficient based on its simple application principle and fast speed.

The function (2) is proposed to identify the terminate number of clusters. Here, n is the number of objects, and K is the number of clusters. After clustering, to evaluate its effectiveness E, we propose to use function (3) shown below to evaluate.

$$K \approx \sqrt{n/2} \, . \tag{2}$$

$$E = \sum_{i=1}^{k} \sum_{a \subset A} \left\| a - q_i \right\|^2 . \quad (3)$$

Here, q^i is the average number of the A cluster, and letter a here is one target data in a cluster. After repeating this computing process, massive data can be categorized into different clusters.

Attribute Selection

In fact, within a group of customers (cluster), not all original attributes are pertinent (or informative) and some of them should be discarded. Here again, we developed an entropy-based approach to decide whether or not an attribute is important with regard to a given group of customers using a user-defined threshold for the information loss. As a measure for the "information" conveyed by an attribute, we will use the entropy based on the frequency of the values taken by that attribute within a given cluster. By applying the clustering approach to a given attribute within a given group of customers, we obtain a set of clusters each embedding values for that attribute that are very similar; and therefore considered as being a single value occurring as many times as the size of the corresponding cluster. Now, we can define the frequency of the value of an attribute within a given cluster as follows:

$$Y_{c,f}^{g} = \frac{P_{c,f}^{g}}{N_f} . \quad (4)$$

Here, c denotes an attribute. Nf is the number of individuals in one group. $Y_{c,f}^{g}$ (g = 1. . . n) is the gth cluster for attribute c within one group. $P_{c,f}^{g}$ is the number of individuals in cluster $Y_{c,f}^{g}$.Thereby, we define the entropy of attribute c with respect to group wi by:

$$E_{w_i}(c) = -\sum_{g=1}^{\sigma_c^f} Y_{c,f}^{g} \log\left(Y_{c,f}^{g}\right), . \quad (5)$$

Here, and wi is a computed group of customers. Consequently, we define the entropy of group w_i as being a summation of the entropy of all attributes, i.e.:

$$E(w_i) = \sum_{c} E_{w_i}(c) \cdot \quad (6)$$

In order to quantify the importance of an attribute, say c, with respect to a group of customers, say w_i, we will define the information loss, noted ILwi (c), as being the "quantity of information" lost when a is removed from the initial set of attributes. It is defined as:

$$IL_{W_i}(c) = 1 - \Delta w_i(c) . \quad (7)$$

Here, Δ_{Wi} (c) is the entropy variation when attribute a is removed and computed by:

$$\Delta wi(c) = \frac{\left(E(wi) - E_{wi}(c)\right)}{E(wi)} . \quad (8)$$

Here, E (w_i) is the entropy of group w_i computed by Eq. (6); and Ew_i (c) is the entropy of attribute c computed by Eq. (5). Intuitively speaking, an attribute is important for a group of customers if it is essential for grouping these customers together. Stated otherwise, this attribute is a common characteristic for the individuals within this group. The customers within the same group will have very similar values for an attribute that is important.

Learning Customer Model

In this third and final step, our main objective is modeling the knowledge embedded in each already reduced cluster (or group of customers). Several approaches have been proposed in the literature including neural networks, association rules, self-organizing Feature Maps and other approaches. In our approach, we will use decision trees in order to extract the most important features (characteristics) for each group of customers on the basis of the first two steps stated above.

3.2 Experiment

For our research purpose, we can assume that we were given a series of data about customers' potential purchasing behaviors, on the basis of the clustering and attribute selection process stated above, we can assume that the selected most informative attributes for clustered groups are A, B, C, D, E, and each attribute include some different values. For instance, attribute A may include X1, X2, and X3… B may include Y1, Y2, and Y3…. the same, other informative attributes also include different factors. For the convenience of our research, we can assume here that all kinds of variable factors under A attribute value could be categorized into three general ranges: a1, a2, a3, the same, other variable factors under different attributes also could be categorized into different ranges: attribute B could include b1, b2, b3… C could include c1, c2, c3…D could include d1, d2, d3…. E could include e1, e2, e3….

Then, it is time to use information gain and information entropy calculation method to make decision trees for extracting potential valuable knowledge purpose (See Table 1). This paper will use this assumed database as an example to illustrate the application of proposed model.

Table 1. Attributes value of customers' sample

	A	B	E	D	C
1	X1	Y1	E1	Z1	C1
2	X2	Y2	E1	Z2	C2
3	X3	Y3	E2	Z3	C1
4	X4	Y4	E2	Z3	C2
5	X5	Y5	E1	Z4	C2
6	X6	Y6	E1	Z5	C2
7	X2	Y7	E2	Z6	C2
8	X7	Y8	E1	Z7	C2
9	X8	Y9	E1	Z8	C2
10	X9	Y2	E2	Z9	C2
11	X10	Y8	E1	Z10	C2
12	X11	Y3	E1	Z11	C2
13	X12	Y9	E2	Z12	C2
14	X13	Y3	E1	Z13	C1
15	X14	Y2	E1	Z14	C2
16	X15	Y1	E1	Z15	C2
17	X16	Y3	E1	Z16	C2
18	X17	Y3	E2	Z17	C1
19	X18	Y1	E2	Z18	C1
20	X19	Y5	E1	Z19	C2
21	X20	Y7	E1	Z20	C2
22	X21	Y4	E2	Z21	C1
23	X22	Y9	E1	Z22	C2
24	X23	Y8	E1	Z23	C1
25	X24	Y9	E2	Z24	C1

Use attribute value of "A" shown in Table 1 as an example, assume that these massive data under this cluster can be attributed to three different groups: a1, a2, and a3. The same, we can generalize other attributes such as "B", "E", "D" and "C" (See Table 2).

Table 2. Generalized attributes value of sampled customer group

	A	B	E	D	C
1	a1	b1	e1	d1	c1
2	a2	b2	e2	d3	c2
3	a3	b3	e1	d2	c1
4	a1	b1	e2	d3	c2
5	a2	b3	e1	d3	c2
6	a1	b2	e1	d1	c2
7	a2	b1	e2	d1	c2
8	a1	b2	e1	d3	c2
9	a3	b3	e1	d2	c2
10	a2	b1	e2	d2	c2
11	a1	b1	e1	d1	c2
12	a2	b2	e1	d2	c2
13	a2	b2	e2	d2	c2
14	a1	b2	e1	d1	c1
15	a2	b1	e1	d3	c2
16	a1	b2	e1	d2	c2
17	a2	b1	e1	d1	c2
18	a1	b2	e2	d3	c1
19	a2	b1	e2	d1	c1
20	a1	b3	e1	d3	c2
21	a2	b1	e1	d1	c2
22	a2	b2	e2	d3	c1
23	a2	b2	e1	d3	c2
24	a1	b3	e1	d3	c1
25	a3	b2	e2	d3	c1

3.3 Explaining Attributes-Decision Trees

Here we use decision trees method, because we don't have to make any changes on the target data set if the data set belongs to "IF-THEN" format. Decision trees can make data analysis processes easier by showing the mapping relationship between object properties and object values in one visual and intuitionistic way. Consequently, compared to other data models, this method is easier to understand, and can clearly explain the mapping relationships among data within a relative short time.

On the basis of the two steps stated above, decoding the qualitative and quantitative information that are encoded on table is another main task. This paper mainly proposes the use of decision trees to explain the value hidden in big data. This paper recommends the use of the information gain-ID3 calculation method. The specifics are shown as following:

From the Table 2, it is easy to know that there are two different values for the C attribute: so we identify m=2. Here we can identify these two values are C1, and C2. So 17 samples are included in C1, and 8 samples are included in C2. Our goal is probing for the most important test attribute of decision trees as the root node. In view of information-gain and information gain rate-based calculation. This root node can be obtained through calculating and comparing the information gain rate among different attributes. We propose the following function about information entropy's calculation. The specifics are as follows:

1. Based on the information entropy calculation as follows:

$$H(x_1, x_2 ... x_n) = -\sum_{i}^{n} p_i \log_2(p_i) \text{ (i=1 , 2,…n)} . \tag{9}$$

   The information entropy can be got as follows:

$$H(x_1, x_2) = H(17,8) = -\frac{17}{25}\log_2\frac{17}{25} - \frac{8}{25}\log_2\frac{8}{25} = \frac{17}{25}\log_2\frac{25}{17} + \frac{8}{25}\log_2\frac{25}{8} = 0.9044 . \tag{10}$$

2. Calculating the entropy of every attribute, including attribute A, B, E and D (Calculating the entropy of attribute A as a detailed example shown as follows):

   When $A \subset a_1$ ： x11=9 ， x21=2，

$$H(x_{11,} x_{21}) = -\frac{9}{11}\log_2\frac{9}{11} - \frac{2}{11}\log_2\frac{2}{11} = 0.6840 . \tag{11}$$

   When $A \subset a_2$: x12=8，x22=4，

$$H(x_{12,} x_{22}) = -\frac{8}{12}\log_2\frac{8}{12} - \frac{4}{12}\log_2\frac{4}{12} = 0.9183 . \tag{12}$$

   When $A \subset a_3$： x13=0，x23=2，

$$H(x_{13,} x_{23}) = -\frac{2}{3}\log_2\frac{2}{3} = 0.3899 . \tag{13}$$

   So the information entropy of attribute A sample is:

$$E(A) = \frac{11}{25}H(x_{11}, x_{21}) + \frac{12}{25}H(x_{12}, x_{22}) + \frac{2}{25}H(x_{13}, x_{23}) = 0.7929 . \tag{14}$$

   With the same method, it is easier to know the information entropy of other attributes:

$$E(E) = \frac{16}{25}\left(-\frac{13}{16}\log_2\frac{13}{16} - \frac{3}{16}\log_2\frac{3}{16}\right) + \frac{9}{25}\left(-\frac{4}{9}\log_2\frac{4}{9} - \frac{5}{9}\log_2\frac{5}{9}\right) = 0.8024 . \tag{15}$$

$$E(B) = \frac{9}{25}\left(-\frac{5}{9}\log_2\frac{5}{9} - \frac{4}{9}\log_2\frac{4}{9}\right) + \frac{11}{25}\left(-\frac{7}{11}\log_2\frac{7}{11} - \frac{4}{11}\log_2\frac{4}{11}\right) + \frac{5}{25}\left(-\frac{5}{5}\log_2\frac{5}{5} - \frac{0}{5}\log_2\frac{0}{5}\right) = 0.7729 . \tag{16}$$

   So the information gain can be calculated as follows:

$$Gain(E) = H(x_1, x_2) - E(E) = 0.9044 - 0.8024 = 0.102 . \tag{17}$$

$$Gain(A) = H(x_1, x_2) - E(A) = 0.1115 . \tag{18}$$

$$Gain(B) = H(x_1, x_2) - E(B) = 0.1315 . \tag{19}$$

$$Gain(D) = H(x_1, x_2) - E(D) = 0.0269 . \tag{20}$$

Then, this paper proposes to use C4.5 calculation to get the gain ratio of each attribute to probe for the root node. Gain ratio of one attribute can be got through dividing gain of each attribute into its information entropy, which is shown as follows:

Gain Ratio (A)=Gain (A)/E (A)=0.0713/0.8192=0.087. **(21)**

Gain Ratio (B)=Gain (B)/E (B)=0.1236/0.7669=0.1612. **(22)**

Gain Ratio (D)=Gain (D)/E (D)=0.0375/0.853=0.0419. **(23)**

Gain Ratio (E)=Gain (E)/E (E)=0.1075/0.783=0.137. **(24)**

Based on the above calculation result, "B" has the highest information gain to be the root node of decision tree. (See Figure 1)

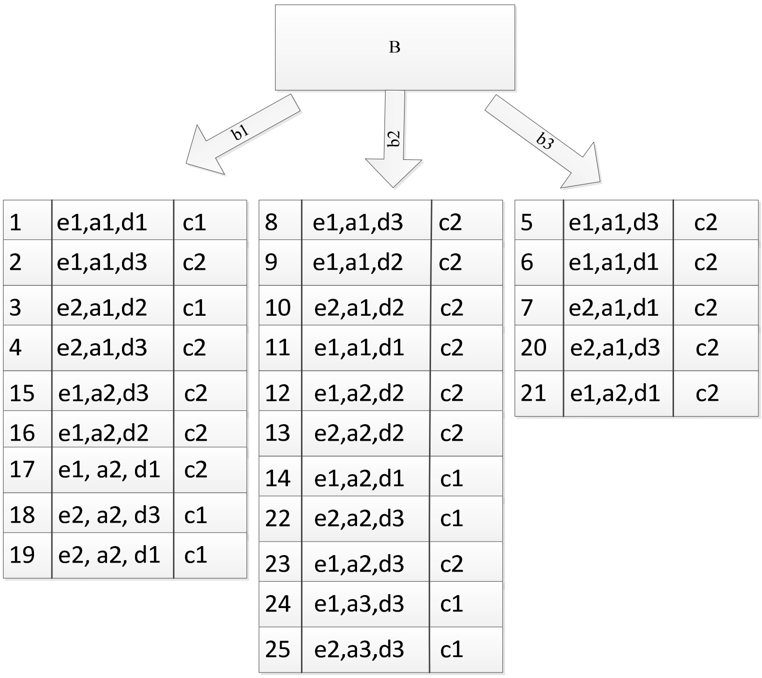

Fig. 1. Decision tree model

Here, we can get three new subsets: P1={1,2,3,4,15,16,17,18,19}，P2={8,9,10,11,12,13,14,22,23,24,25}, P3={5,6,7,20,21}. It is easier to calculate the information gain of other attributes in these subsets using the same calculation method that is stated above. For P1,

$$Gain(A)=I(S_{1,}S_2)-E(A)=0.007 \quad \textbf{(25)}$$

$$Gain(D)=I(S_{1,}S_2)-E(D)=0.103 \quad \textbf{(26)}$$

$$Gain(E)=I(S_{1,}S_2)-E(E)=0.23 \quad \textbf{(27)}$$

4 Experiment Results and Discussion

Based on the above example and discussed techniques, it is easy to find that attribute "E" has the highest information gain, so this attribute will be regarded as another new node for further establishment of decision trees. After dealing with P2 using the same method, the final decision tree is formed as follows:

Now, it is easier for business to find regular pattern through analyzing the potential valuable information. From the above decision tree, we can find the following patterns: 1. The C2 is more popular among customers in d3 group. 2. C1 is more popular among d2 group customers and d1 group customers. 3. Among different attribute B values, C2 is more popular for a1 group, but C1 is more popular for a2 group customers and a3 group customers.

Based on the above analysis, business can sense that the target group of C2 is a1 group customers with d3 attribute, and the target groups of good C1 should be a2 group customers and a3 group customers. Consequently, it is the right time to accelerate advertisement about newly C2 among a1 group customers; especially among a1 members with d3 attribute and promotes newly C1 good among d2 group customers and a3 group customers.

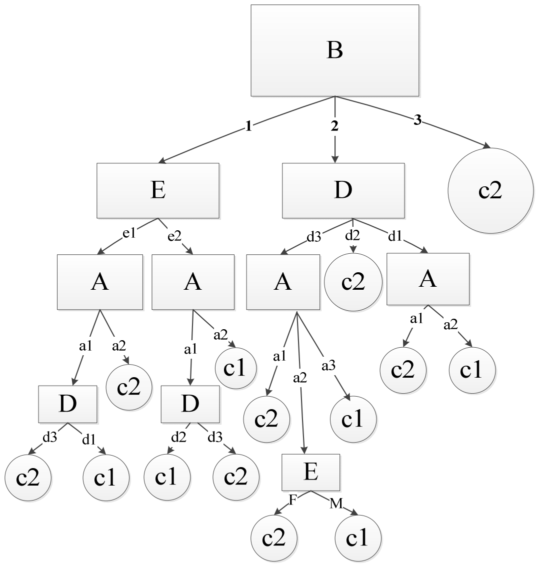

Fig. 2. Decision tree-based clustering model

Based on such a data-mining model, businesses can extract valuable information to achieve accurate marketing strategies. Businesses can promote personalized marketing strategies through finding inner regular patterns on the basis of data clustering and integration. It is no doubt that business can lower the operation cost, and optimize the customers' retention rate for better CRM.

5 Conclusions

Based on the above discussion about this proposed three-step data mining model including using K-means to cluster massive data; Generalizing data to focus on relevant attributes via using a user-defined threshold for the information loss calculation; Calculating information gain ratio to make decision trees for extracting potential valuable knowledge purpose. More importantly, such a data-mining model simplifies the data analysis process, and provides logical basis for analysis. Unarguably, data mining model can bring amazing value for CRM, no matter for business, government, or for public institutions. CRM will work effectively with the speedy data integration and data mining. This model also sets up a good foundation for future data research under "big data era".

Acknowledgment

This work was supported in part by National Taichung University of Education, Taiwan, under Grant NTCU-F104201.

References

[1] D. Anthony, H. Oscar, J. Gandy, "All that glitters is not gold: digging beneath the surface of data mining," *Journal of Business Ethics*, Vol. 40, No. 4, pp. 373-386, 2002

[2] I. Bittencourt, E. Costa, M. Silva, E. Soares, "A computational model for developing semantic web-based educational systems," *Knowledge Based Systems*, Vol. 22, No. 2, pp. 302-315, 2009

[3] A. Bueren, R. Schierholz, L. Kolbe, W. Brenner, "Customer Knowledge Management: improving performance of customer relationship management with knowledge management," in *Proceedings of The 37th IEEE Hawaii International Conference on System Sciences*, IEEE Computer Society Press, 2004.

[4] C. Hung, and C. Tsai, "Market segmentation based on hierarchical self-organizing map for markets of multimedia on demand," *Expert System*, pp. 780–787, 2008

[5] D. Chen, N. Vachharajani, R. Hundt, X. Li, S. Eranian, W. Chen, W. Zheng, “Taming hardware event samples for precise and versatile feedback directed optimizations,” *IEEE Transactions on Computers*, Vol. 62, No. 2., pp. 76-389, 2013

[6] X. Chen, Z. Zheng, X. Liu, Z. Huang, H. Sun, “Personalized QoS-aware web service recommendation and visualization,” *IEEE Transactions on Services Computing*, Vol. 6, No. 1, pp. 35-47, 2013

[7] C. Huang, M. Chen, C. Wang, “Credit scoring with a data mining approach based on support vector machines,” *Expert Systems with Applications*, Vol. 33, No. 4, pp. 847–856, 2007.

[8] G. Fung, L. Mangasarian, “Proximal support vector machine classifiers,” in *Proceedings of Knowledge Discovery and Data Mining (KDD)*, pp. 77–86, 2001.

[9] C. Hsu, S. Chen, T. Chen, “Message transmission techniques for low traffic P2P services,” *International Journal of Communication Systems*, Vol. 22, No. 9., pp. 1105-1122, 2009

[10] C. Hsu, Y. Chen, H. Kang, “Performance-effective and low-complexity redundant reader detection in wireless RFID networks,” *EURASIP Journal on Wireless Communications and Networking*, Vol. 2008, Article ID: 604747, 2008.

[11] J. Gehreke, “Decision trees,” in *The Handbook of Data Mining,* pp. 3–25, 2003.

[12] Y. Kuoa, L. Ana, W. Wanga, J. Chungbi, “Integration of self-organizing feature maps neural network and genetic K-means algorithm for market segmentation,” *Expert System with Applications*, Vol. 30, No. 2, pp. 313–324, 2006.

[13] B. Lotfi, A. Romdhane, N. Fadhel, B. Ayeb, “Building customer models from business data: an automatic approach based on fuzzy clustering and machine learning,” *International Journal of Computational Intelligence and Application,* Vol. 8, No. 4., pp. 445-465, 2009.

[14] N. Hsieha, “An integrated data mining and behavioral scoring model for analyzing bank customers,” *Expert Systems with Applications*, Vol. 27, No. 4, pp. 623–633, 2004.

[15] P. Giudici, G. Passerone, “Data mining of association structures to model consumer behavior,” *Computer Statistics Data Analysis*, Vol. 38, No. 4, pp. 533–541, 2002.

[16] S. Mitra, S. Pal, P. Mitra, “Data mining in soft computing framework: a survey,” *IEEE Transactions on Neural Networks*, Vol. 13, No. 1, pp. 3–14, 2002.

[17] P. Soukakos, N. Georgopoulos, V. Economou, “Interrelated frame-works proposed for mapping and performance measurement of customer relationship management strategies,” *International Journal of Knowledge and Learning*, Vol. 3, No. 2-3, pp. 299–315, 2007.

[18] S. Thomas, and R. Moore, “E-education 3.0: challenges and opportunities for the future of I-campuses,” in *Proceedings of International Conference on Digital Object Identifier*, pp. 953-958, 2012.

[19] L. Romdhane, N. Fadhel, B. Ayeb, “Building customer models from business data: an automatic approach based on fuzzy clustering and machine learning,” *International Journal of Computational Intelligence and Applications*, Vol. 8, No. 4, pp. 445–465, 2009.

[20] S. Susanne, S. Gerhard, “Designing for mobile value co-creation-the case of travel counseling,” *Electronic Markets*, Vol. 24, pp. 5-17, 2014.

[21] N. Nhat, D. Cavada, F. Ricci, “Trip at device-mobile extension of a case-based travel recommender system,” in *Proceedings of* 2*nd International Conference on Mobile Business (ICMB)*, 2003.

[22] K. Mirella, K. Ruyter, W. Martin, “An assessment of value creation in mobile service delivery and the moderating role of time consciousness,” *Journal of Retailing*, Vol. 83, No. 1, pp. 33-46, 2006.

[23] Y.-S. Wang, H.-H. Lin, P. Luarn, “Predicting consumer intention to use mobile service,” *Information Systems Journal,* Vol. 16, No.2, pp. 157-179, 2006.

[24] Y. Shuk, H. Sai, “The attraction of personalized service for users in mobile commerce: an empirical study,” *ACM SIGecom Exchanges,* Mobile Commerce, Vol. 3, No. 4, pp. 10-18, 2002.

[25] J. Fink, J. Koenemann, S. Noller, I. Schwab, “Personalization into Practice,” *Communications of the ACM*, Vol. 45, No. 5, pp. 41-42, 2002.

[26] H.C. Chu, D.J. Deng, J.H. Park, “Live data mining concerning social networking forensics based on a facebook session through aggregation of social data,” *IEEE Journal on Selected Areas in Communications*, Vol. 29, No. 7, pp. 1368-1376, 2011.

[27] M.H. Guo, H.T. Liaw, D.J. Deng, H.C. Chao, “Cluster-based secure communication mechanism in wireless ad hoc networks,” *IET information security*, Vol. 4, No. 4, pp. 352-360, 2012.