CONSTRUCTION OF MARKETING INFORMATION MANAGEMENT SYSTEM OF LIQUOR-MAKING ENTERPRISES UNDER CLUSTERING ALGORITHM

M.Y. LI

College of Business Administration, Liaoning Technical University, Fuxin, China.
NelsongCAb@yahoo.com

Abstract — Large supermarket is the strategic commanding point of liquor marketing channels. For liquor, large supermarket is the combination of linear and offline channels. It is the gathering place for middle and high end liquor, which is the strategic commanding point for brand building. This paper uses the high-end liquor brand A as an example to excavate the consumption data of a large supermarket from January 1, 2015 to December 31, 2015. This paper builds a customer cluster by combining with the status quo of liquor marketing. The algorithm, which uses the analytic hierarchy process to calculate the weight of each index, obtains the key customer groups for the marketing of liquor companies and helps to optimize and integrate the marketing resources of liquor companies.

Keywords — offline business supermarket; liquor marketing; huge amount of data; optimization.

I. INTRODUCTION

It is well-known that the development of science and technology and economic progress make the application of modern information technology more and more extensive in various fields. Marketing management is no exception (Noel et al., 2017). Information technology plays an increasingly important role in the daily operation and decision-making of enterprises.

As the traditional essence of the Chinese nation, liquor has occupied an important seat in the Chinese consumer market, and the informatization construction of the liquor marketing industry has received more and more attention (Bruijn et al., 2016). In recent years, the marketing team's informatization construction has been initially applied in many large and medium-sized enterprises in the liquor industry (Stautz et al., 2016).

The information management system provides a good platform for the company's own operation and management, realizes the collaborative work of the team, and brings high efficiency of operation. Therefore, the construction of enterprise software system will be changing with each passing day in the liquor industry.

II. STATE OF THE ART

At present, as a research direction of Distributed Artificial Intelligence (DAI), the research and application of Agent is actively being carried out within the international and domestic scope. As we all know, the application of computers is becoming more and more widespread and the problems that need to be handled are becoming more and more complicated (Prescott et al., 2015). As the information, data, and data involved in problem solving become more and more difficult to centralize, the process of computer solving becomes increasingly difficult to centralize (Angmo et al., 2016).

The distribution and concurrent processing of all these data or knowledge has brought great potential for the development of AI (Artificial Intelligence) and also brought about various difficult problems to be solved. Research on multi-agent systems has become a hot topic in artificial intelligence research. The multi-agent system mainly studies the coordination of intelligent behavior among autonomous intelligent agents (Hoffman et al., 2015). For a common global goal, or different targets, knowledge of related problems and solutions is shared, solving problems collaboratively (Green et al., 2015).

Some colleges and universities and research institutes in the Mainland are conducting relevant research, such as related theories and methods of the National University of Science and Technology of the People's Republic of China in carrying out Agent-based distributed integration environment and multi-agent cooperation model; Nanjing University is developing software engineering and security research for Agent technology; Tsinghua University and Zhejiang University are engaged in research on the application of multi-agent structure in concurrent engineering and network management. Although Agent-oriented technology plays an important role in the development and implementation of complex systems, Agent technology itself has some inherent flaws. First, the entire system developed with Agent technology will be uncertain (Jeong et al., 2015).

III. METHODOLOGY

A. Liquor Marketing Clustering Algorithm

Clustering is a process of dividing a data set into groups or classes. It makes the data objects in the same group
have a high similarity, while the data objects in different groups are not similar (Érica et al., 2015). A cluster is a collection of objects that resemble each other. Objects in different clusters are usually dissimilar. Similar or dissimilar measures are determined based on the values of the data object description attributes (Lenardson et al., 2015). Cluster analysis relies on the understanding of the proximity (distance) or degree of similarity between observations. Defining different distance metrics and similarity measures can produce different clustering results. There are many cluster analysis methods. This paper mainly applies the fast cluster analysis algorithm. The following uses the fast clustering analysis algorithm to cluster the number of business target samples.

Multivariate data forms a data matrix (see Table 1). In this data matrix, there are n samples: \( x_1, x_2, \ldots, x_n \) (column) and p indicators (row-wise).

There are two types of clustering analysis: Clustering by sample or clustering by variable (indicator).

### Table 1 Data Matrix

<table>
<thead>
<tr>
<th>Sample</th>
<th>( x_1 )</th>
<th>( x_2 )</th>
<th>( \ldots )</th>
<th>( x_j )</th>
<th>( \ldots )</th>
<th>( x_n )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( x_1 )</td>
<td>( x_{11} )</td>
<td>( x_{12} )</td>
<td>( \ldots )</td>
<td>( x_{1j} )</td>
<td>( \ldots )</td>
<td>( x_{1n} )</td>
</tr>
<tr>
<td>( x_2 )</td>
<td>( x_{12} )</td>
<td>( x_{22} )</td>
<td>( \ldots )</td>
<td>( x_{j2} )</td>
<td>( \ldots )</td>
<td>( x_{n2} )</td>
</tr>
<tr>
<td>( \ldots )</td>
<td>( \ldots )</td>
<td>( \ldots )</td>
<td>( \ldots )</td>
<td>( \ldots )</td>
<td>( \ldots )</td>
<td>( \ldots )</td>
</tr>
<tr>
<td>( x_p )</td>
<td>( x_{ip} )</td>
<td>( x_{2p} )</td>
<td>( \ldots )</td>
<td>( x_{jp} )</td>
<td>( \ldots )</td>
<td>( x_{np} )</td>
</tr>
</tbody>
</table>

Let \( d_{ij} = d(x_i, x_j) \), \( D = (d_{ij})_{n \times n} \) form a distance matrix, where

\[
D = \begin{bmatrix} 0 & d_{i1} & \cdots & d_{in} \\ d_{1i} & 0 & \cdots & d_{in} \\ \vdots & \vdots & \ddots & \vdots \\ d_{ni} & d_{ni} & \cdots & 0 \end{bmatrix}
\]

(3)

Based on the above basic theory of cluster analysis, the specific steps in the rapid cluster analysis algorithm include the steps of: randomly selecting the cluster seed points or center points, allocating each observation sample to the nearest seed, re-aggregating each cluster, using the center point in the seed as a seed and repeating the above process until the change in the seed is small enough.

### B. Agent-based data mining system

The mobile agent, abbreviated as ma. The travel plan details the movement and path of the ma. The ma mobile mode can be divided into three types: sequence, selection, and distribution. The moving path is a list of identifiers of the peripheral node to be travelled. The task body is composed of a series of operations that complete the tasks undertaken by the ma and expressed in the Java language. The identifier of the mobile agent consists of the host node identifier that creates the agent and the serial number inside the node. The platform can perform distributed data mining from different data sites. Mainly consists of seven parts: user interface agent UIA, user information base, Agent server, global knowledge base, global data mining manager, data management server, and database access interface.
mining agent DMA (Data Mining Agent), and distributed database. The DMA is dynamically created by the system when there is a mining request. Although the DMA may be bound to a specific operating system, the UIA, the Agent server, and the system's definition and communication part are all encoded in Java and are independent of the operating system platform. According to the functional design of the Mobile Miner platform above, the Mobile Miner platform we want to implement is mainly composed of the following Java packages:

<table>
<thead>
<tr>
<th>The package name</th>
<th>Describe</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mobile Miner Ma</td>
<td>Mobile Miner M core class library</td>
</tr>
<tr>
<td>directory</td>
<td>Related classes that provide directory services</td>
</tr>
<tr>
<td>server</td>
<td>Management classes for MA servers</td>
</tr>
<tr>
<td>resource</td>
<td>Related classes that provide resource management</td>
</tr>
</tbody>
</table>

In the Mobile Miner system, all Mobile Agents are inherited from a unified Agent abstract class. This abstract class is a template class that defines the universal features of MA, implements the Serializable interface and Cloneable interface. The definition of the form is as follows: public abstract class Mobile Miner MA implements Serializable, Cloneable {...}. The properties of the Agent abstract class mainly include: a system-wide unique and unchanging identifier, a constant code base URL, and a protected, serializable object that stores the running status of the agent.

C. Customer Definition

According to the differences between three values of the consumer in the three indicators of the SPC model and the mean value of the average, greater than or equal to are remarked as "+" and less is remarked as "-". The e-commerce consumer can be divided into eight types of consumers. According to each consumer, the change of the cluster's indicators analyzes the nature of this type of consumer and defines the type of consumer. They are SPC-, S-P-, S-P+C-, S-P+C+, and S&P-, C-, focusing on retaining consumers (S+P+C+), and high-value customers (S+P-C+). Each type of customer has its own characteristics.

The maintenance method is also different. For example, S+P-C+. This kind of consumer has a large amount of consumption and large potential value. It is a loyal consumer of the enterprise and can be regarded as a company. It is recommended that companies increase their understanding of the important development consumers. If necessary, companies should temporarily reduce the consumer's profit requirements, such as the distribution of gift certificates for increasing their consumption indicators and consumer's profit index. S + P + C +, this type of consumer not only can bring high profits for enterprises, but also has a higher profit margin, which is the company's most important high-quality consumers. It is recommended that companies invest in major resources to analyze such consumers, fully understand their needs, maintain and expand relationships with consumers to increase the loyalty of the consumer.

IV. RESULT ANALYSIS AND DISCUSSION

A. Data Collection and Preprocessing

A liquor company entered the large-scale commercial super system in 2015, which strongly promoted the overall sales of the marketing center in the northeastern market of the country. The sales volume of Supermarkets has increased by more than 90% every year, which has become a flag for the increase in the sales of A liquor companies. A liquor company entered the high-speed Hualian system in the northeast region in 2016 so that the drivers from across the country can effectively see the A liquor brand. They have an impression of A liquor and think that A liquor is famous.

The social benefits of advertising effects are far greater than actual. The economic benefits of sales. Therefore, this article uses the consumption data of a large-scale retailer from January 1, 2015 to December 31, 2015 as the research object. Through preliminary statistics on the consumption data of A liquor consumers, there are a total of 12,294 consumption records generated by 10105 consumers with an average of 1.217 consumptions per consumer. The total consumption was 70,389.19 yuan, and the average cost per purchase was 572.55 yuan. The total number of consumptions was 16,114 times on average.

According to the analysis of the data, we can see that the brand's consumer distribution is in the shape of a pyramid. The number of consumers with high consumption, consumption and frequency of consumption is small. It can be seen that there are differences between the consumer behaviors of the brand's different consumers, and the brand has some loyal consumers. Therefore, it is necessary to subdivide the customers to identify the different customer types of the company and provide data for the company's marketing strategy and customer management stand by.

In general, data preprocessing techniques include
data extraction, data cleansing and data conversion. According to the model built in this paper, using C1 as an example with the extreme difference method to illustrate the use of excel to process the data standardization process.

Sorting the amount actually paid by buyers can get a maximum of 10,841 yuan and a minimum value of 31 yuan for consumer spending. Substitution into the analytic hierarchy process formula can be based on the standardization results, in which the value of column D is obtained by subtracting the minimum value of 31 from column B.

The value of column E, that is, the value of C1 finally obtained is got from the column D to the max-min column. In the same way, the consumption data of the eight program levels can be standardized.

Substitute the data after the completion of standardization into the business over consumer segmentation model under the established mobile environment, assign each indicator with corresponding indicators and convert this indicator into the SPC subdivision model established in this article 3.2. The conversion can convert the value of the six indicators into the final value of the subdivision model established by the three indicators of consumption indicators to profit indicators and consumer indicators.

<table>
<thead>
<tr>
<th>Buyer's name</th>
<th>S</th>
<th>P</th>
<th>C</th>
</tr>
</thead>
<tbody>
<tr>
<td>chumphy</td>
<td>0.2351</td>
<td>0.1749</td>
<td>0.2742</td>
</tr>
<tr>
<td>pumpy221</td>
<td>0.1729</td>
<td>0.1524</td>
<td>0.3564</td>
</tr>
<tr>
<td>suzue</td>
<td>0.1604</td>
<td>0.1263</td>
<td>0.3194</td>
</tr>
<tr>
<td>ssdq</td>
<td>0.1581</td>
<td>0.1364</td>
<td>0.501</td>
</tr>
<tr>
<td>f71662.</td>
<td>0.1564</td>
<td>0.1364</td>
<td>0.5718</td>
</tr>
<tr>
<td>teresa850</td>
<td>0.1535</td>
<td>0.14</td>
<td>0.4517</td>
</tr>
<tr>
<td>Hongpang</td>
<td>0.1532</td>
<td>0.1364</td>
<td>0.501</td>
</tr>
<tr>
<td>Sunnai</td>
<td>0.147</td>
<td>0.1412</td>
<td>0.3645</td>
</tr>
<tr>
<td>200</td>
<td>0.1437</td>
<td>0.1011</td>
<td>0.3233</td>
</tr>
<tr>
<td>1234567cat</td>
<td>0.14</td>
<td>0.2628</td>
<td>0.501</td>
</tr>
<tr>
<td>123highpig</td>
<td>0.1399</td>
<td>0.0625</td>
<td>0.224</td>
</tr>
<tr>
<td>s_jinniu</td>
<td>0.1391</td>
<td>0.1524</td>
<td>0.1928</td>
</tr>
<tr>
<td>hienmaar</td>
<td>0.1378</td>
<td>0.0984</td>
<td>0.2586</td>
</tr>
<tr>
<td>feiter tu</td>
<td>0.1358</td>
<td>0.1074</td>
<td>0</td>
</tr>
<tr>
<td>hg020203</td>
<td>0.1325</td>
<td>0.1289</td>
<td>0.3474</td>
</tr>
<tr>
<td>hanzleng</td>
<td>0.1313</td>
<td>0.1064</td>
<td>0.501</td>
</tr>
<tr>
<td>h05640998</td>
<td>0.1306</td>
<td>0.0814</td>
<td>0.3737</td>
</tr>
</tbody>
</table>

**Figure 1.** The final result after weight conversion part of the data show.

**B. Data Mining**

This article refers to the use of the CRISP-DM model for the process of excavating consumer spending data, as shown in Fig. 2.

**Figure 2.** CRISP-DM data mining process.
After the basic operation steps of data type setting and data filtering, K-means cluster analysis method was applied. When the number of clusters was 5, the clustering quality was well up to 0.8, so this clustering can be used. The cluster distribution is as follows. By analyzing the final results with SPSS Modeler software, it can be clearly seen which category each consumer is in, the proportion of each cluster, and the weight of each cluster’s index. The data mining flow chart in this paper is shown in Fig. 2. The final clustering result is substituted into the SPC model of supermarket consumer in the mobile environment established by Shang Zhang. The mean value of C is 0.4116, the mean value of B is 0.2982. The mean value of M is 0.4274. The Fig. 4 shows the effect of clustering three types of customers.

**Figure 4. Results of cluster analysis (part).**

We can see that customers are well clustered according to the SPC model and the edges between the categories are relatively clear. Due to limited space, the final segmentation results of consumer SPC are not listed. By clustering we finally get five types of customers:

The Cluster 1 (S+P+C+) is a high-value consumer. They are a loyal customer of the company with the smallest percentage (4.9%) and 502 consumers. The characteristics of the consumer group are the big amount of consumption, frequent times, more profit and valuable potential. They are a loyal brand advocates for the J brand. Cluster 3 (S+P+C-) is an important consumer. This type of consumer is also a loyal customer of the company. The percentage is small (10.2%), and the number of consumers is 1035.

The consumer group is characterized by the amount spent. In the five categories of consumers with large amounts and frequent consumption, enterprises can obtain more profits, but the potential value of consumers may be relatively small, which can be regarded as important to maintain consumers. The cluster 2 (S+PC-) is an average consumer, and the proportion of such consumers is also small (5.7%). The number of consumers is 572.

The characteristics of the consumer group is that the consumption amount is relatively large, but the consumer profit is relatively low. Cluster 4 (SP-C+) is a general development consumer. The proportion of such consumers is large (19.9%), and the number of consumers is 2018. This type of consumer spending is relatively small, and the profits that the company can obtain are not large.

This shows that this type of consumer does not have enough awareness and love of the J brand, but such consumers have greater potential value and the company can appropriately invest in such consumers. Cluster 5 (SPC-) is the consumer to be observed. This type of consumer accounts for the largest proportion (56.1%) and the number of consumers is 5,671. This kind of consumer’s consumption in the J brand is less profitable. The potential value is low, so it can be
regarded as the consumer to be observed. If the consumer does not change for a while, the company can give up him.

V. CONCLUSIONS

The current consumption environment of liquor is undergoing a major transformation. For example, the official consumer market is leading to the business consumer market. The mass consumption is changing to personal preference consumption, which places more professional and more refined demands on channels. Therefore, this paper is rooted in the main sales channels of liquor - offline supermarkets to optimize the integration of liquor marketing resources and build an intelligent liquor marketing model. First of all, this paper builds a customer clustering model combined with the current status of marketing in the liquor industry. Second, this paper uses the analytic hierarchy process to calculate the weights of each indicator. Then, this article has obtained the customer consumption data of a large commercial supermarket in liquor brand A in 2015. After data preprocessing, five important customers were obtained: high-value consumers (S+P+C+), important keepers (S+P+C-), general consumers (S+P-C-), general development consumer (SP-C+) and observed consumer (SPC-).

Finally, this article thinks that based on the clustering algorithm marketing model of liquor, marketing resources should be focused on high-value consumers and important maintain consumers, that is, for these two types of consumers for specific marketing, but for the average consumer and general consumers, is more biased in the side marketing, that is to achieve a certain range of marketing coverage in the process of the main marketing object. For the consumers to be observed, there is no need to invest too much marketing resources.

REFERENCES


Received: December 15th 2017
Accepted: June 30th 2018
Recommended by Guest Editor
Juan Luis García Guirao