Segmentation and profiling consumers in a multi-channel environment using a combination of self-organizing maps (SOM) method, and logistic regression

Seyed Ali Akbar Afjeh[a] and Maryam Darvishi[b*]

[a]*Professor of Business Administration, Faculty of Management and Accounting, Allameh Tabatabaei University, Tehran, Iran*
[b]*PhD candidate in Business Administration, Allameh Tabatabaei University and Faculty member of Islamic Azad University Omidiyeh, Iran*

C H R O N I C L E	A B S T R A C T
	Market segmentation plays essential role on understanding the behavior of people's interests in purchasing various products and services through various channels. This paper presents an empirical investigation to shed light on consumer's purchasing attitude as well as gathering information in multi-channel environment. The proposed study of this paper designed a questionnaire and distributed it among 800 people who were at least 18 years of age and had some experiences on purchasing goods and services on internet, catalog or regular shopping centers. Self-organizing map, SOM, clustering technique was performed based on consumer's interest in gathering information as well as purchasing products through internet, catalog and shopping centers and determined four segments. There were two types of questions for the proposed study of this paper. The first group considered participants' personal characteristics such as age, gender, income, etc. The second group of questions was associated with participants' psychographic characteristics including price consciousness, quality consciousness, time pressure, etc. Using multinominal logistic regression technique, the study determines consumers' behaviors in each four segments.

1. Introduction

With the advent of electronic commerce via Internet, multi-channel retail business world has been the dominant paradigm (Rangaswamy & Van Bruggen, 2005). During the past few years, people have switched buying goods and services through internet, catalog, etc. and this has changed consumer's purchasing pattern, significantly. Ansari et al. (2008) developed a model of customer channel migration and implemented it to a retailer who markets over the Web and through catalogs. They detected the key phenomena needed to analyze customer migration, demonstrated how these phenomena could be modeled, and proposed a technique for forecasting the model. The technique was unique in its capability to incorporate heterogeneous customer replies to a relatively big number of distinct marketing communications in a dynamic context. They reported that web purchasing was

associated with lower subsequent purchase volumes than when buying from other outlets. In addition, marketing efforts in their survey were associated with channel usage and purchase incidence, offsetting negative Web experience effects; and finally, negative interactions existed between like communications. They also reported that over the four-year period of their data, a Web-oriented "migration" segment emerged, and this group maintained higher sales volume. Their post hoc analysis recommended that marketing efforts and exogenous customer-level played essential roles in forming these segments.

Bhatnagar and Ghose (2004) implemented a latent class modeling technique to segment web shoppers, based on their purchase behavior across various product categories. They also profiled the segments along the twin dimensions of demographics and different advantages and described that they could provide more diagnostic information than mere descriptive demographic profiling. The study shed light on consumer perceptions and behavior with in terms of online commerce. First, consumers were more concerned about web attributes associated with perceived losses than with web attributes that consumers associate with gains. In addition, compared with other online purchase-related attributes, getting the lowest price did not seem to be a very important attribute for web shoppers. The finding may also state that prices on web pages were somewhat similar, and consumers were moving on to other criteria to continue their evaluation process.

Chandon et al. (2000) proposed a model of the multiple consumer advantages of a sales promotion and explained the implications of the multi-benefit and the benefit congruency frameworks for understanding consumer responses to sales promotions. Chen (2003) developed a scheme integrating the chi-square automatic interaction detection method, which expands the current analytical spectrum of travel segmentation. He also used a case study of college students' spring-break trips to describe the necessary steps of segmenting a population into mutually exclusive subpopulations. Darian (1987) investigated the relationship between demographic characteristics of households and in-home shopping rates, based on data from the Institute of Retail Management/Management Horizons Shopping Behavior Data Base. They study reported that housewives and part-time female workers with preschool children, single men less than 40 years old, households where the wife was aged 40–49 yeas, and households in the middle income groups could be the groups who were most likely to be in-home shoppers.

Inman et al. (2004) investigated the moderating role of channel-category associations in consumer channel patronage by contributing to the literature on brand associations to the context of channels. They estimated a model, which builds a link between channel-category associations and consumer geodemographics as well as channel share of volume. The channel-category associations found significant impacts and interaction impacts with channel type and geodemographic factors on channel share of volume, and they accounted for most of the explained variance (72%) in channel share of volume.

Kumar and Venkatesan (2005) developed a conceptual model, which determined the customer level details and supplier factors associated with buying behavior across multiple channels by proposing that multichannel shoppers could provide some advantages as measured by several customer-based metrics. They built an empirical analysis based on the customer database of a high technology hardware and software manufacturer and reported that customers who purchase across multiple product categories could initiate more contacts with the firm. These customers had some past experiences with the supplier through the online channel, maintained longer tenure, bought more frequently, were larger and receive communication from the supplier through multiple communication channels, especially through highly interpersonal channels. Konuş et al. (2008) considered shopping a dynamic process, which includes search and purchase phases, in which the total utility of shopping process was determined by the perceived consumer utility toward channel use and they were driven by consumer characteristics.

2. The proposed study

This paper presents an empirical investigation to shed light on consumer's purchasing attitude in multi-channel environment. The proposed study of this paper designed a questionnaire and distributed it among some people who were at least 18 years of age and had some experiences on purchasing goods and services on internet, catalog or regular shopping centers. The sample size is calculated as follows,

$$N = Z_{\alpha/2}^2 \frac{p \times q}{e^2},$$ (1)

where N is the sample size, $p = 1 - q$ represents the probability, $z_{\alpha/2}$ is CDF of normal distribution and finally ε is the error term. For our study we assume $p = 0.5, z_{\alpha/2} = 1.96$ and e=0.05, the number of sample size is calculated as N=660. The study has distributed 800 questionnaires among various kinds of people in order to make sure that the survey maintains sufficient validity. The proposed study of this paper uses self-organizing map (SOM) (Dolnicar & Leisch, 2003), which is a type of artificial neural network (ANN) (Grønholdt & Martensen, 2005) trained using unsupervised learning to generate a two-dimensional, discretized representation of the input space of the training samples, called a map. Self-organizing maps are different from other similar techniques in the sense that they apply a neighborhood function to preserve the topological properties of the input space. Based on the results of our survey, the study uses multinominal logistic regression technique to determine consumers' behaviors in each segment. Next, we present details of our findings.

3. The results

In this section, we present details of our findings on the implementation of SOM technique as well as regression analysis for analyzing consumer's behaviors. Fig. 1 shows distribution of market segmentation.
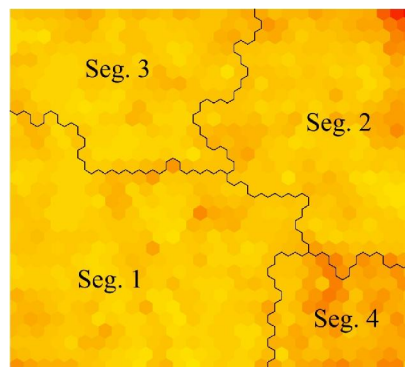

Fig. 1. The summary of market segmentation using SOM technique

According to the results of SOM technique, 37.63% of the participants are located in the first section, 27.38% of them are located in the second part, 23.88% of them are located in the third section and 11.13% of the participants were located in the last part of the market segmentation. Table 1 and Fig. 2 show the results of our findings on market segmentation.

Table 1
The attitude of consumers in each of the different channels

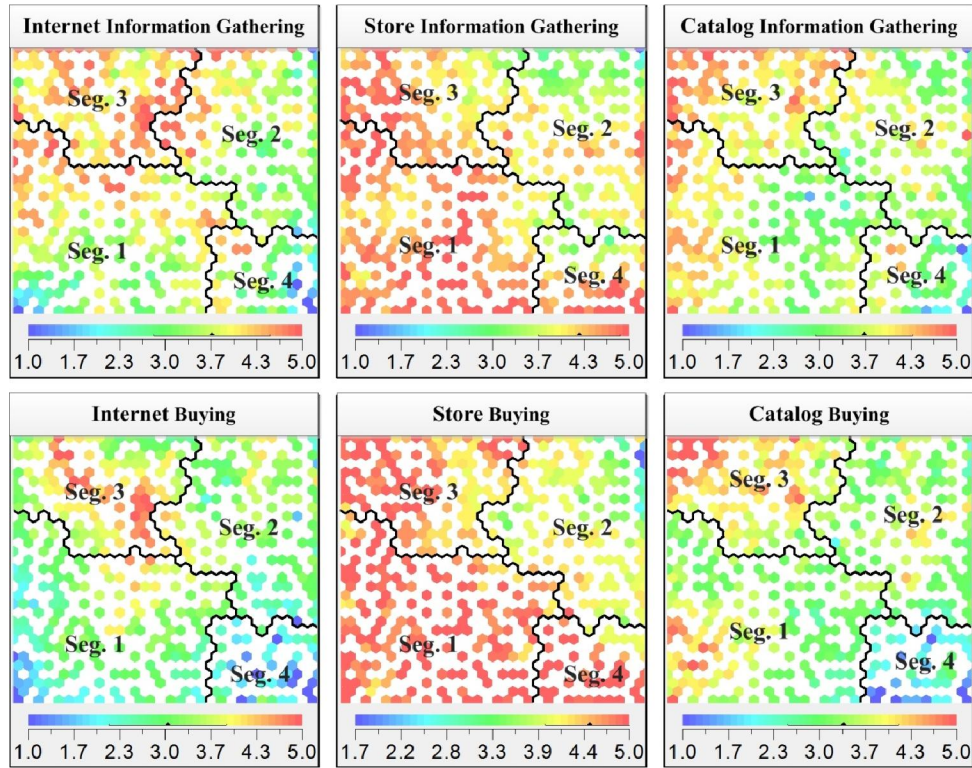| Segment	Information			Purchase		
	Internet	Shopping center	Catalog	Internet	Shopping center	Catalog
1	3.87	4.62	3.89	2.80	4.84	3.39
2	3.52	3.72	3.37	3.14	3.86	3.23
3	4.31	4.30	4.27	3.81	4.62	4.10
4	2.76	4.35	2.83	1.87	4.82	1.85

Fig. 2. Map of the characteristics reported for segmentation variables

The correlation between various variables are measured based on the similarity of the colors. In addition, dot colors represent micro clusters (MC) and each MC contains some consumers with similar purchasing behaviors. The number of MCs depends on the number of neurons on the output layer of the network as well as on the number of samples with very similar characteristics. Fig 3 shows details of consumers' behaviors in various market segmentations.
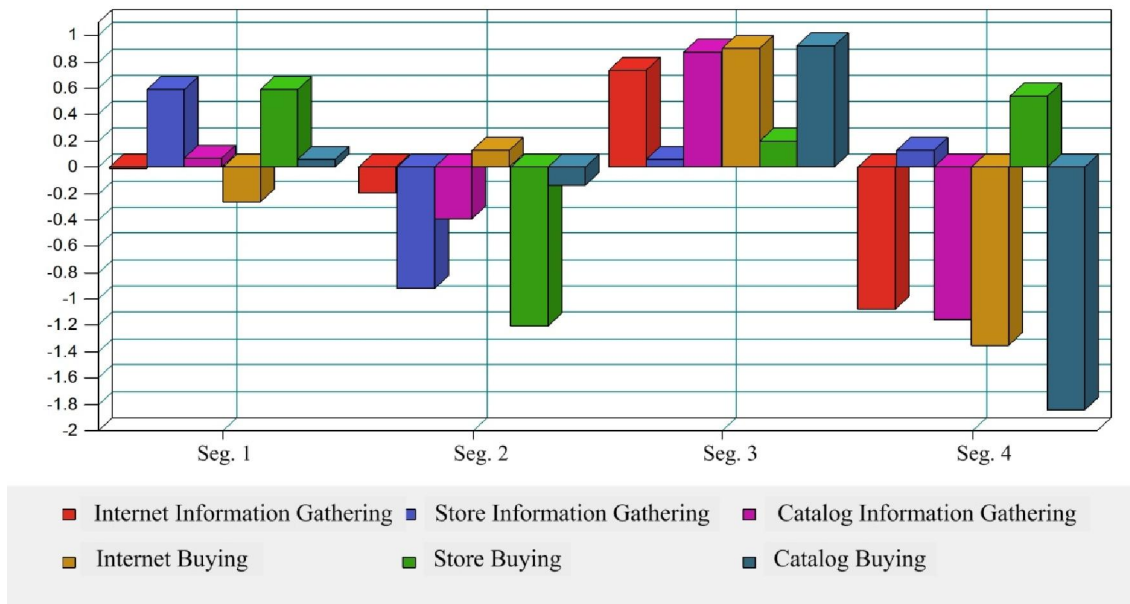

Fig. 3. Relative position of the segmentation variables

As we can observe from the results of Fig. 3, the participants who are located in the second group are not interested in either gathering information or purchasing goods or services through internet, shopping centers or catalog. In contrast, consumers who are located in the third group are highly interested in gathering information and purchasing goods or services through internet, catalog or shopping centers. For the implementation of regression analysis, we consider the information of the second group as reference, dependent variable, and the information of the first, third and fourth groups are considered as the independent variables. Table 2 summarizes of details of our findings.

Table 2
The summary of the results of our logistic regression

Variable		Segment 1	Wald	Sig.	Segment 3	Wald	Sig.	Segment 4	Wald	Sig.
Price consciousness		0.08	0.26	0.6	0.05	0.08	0.77	-0.13	0.3	0.58
Quality consciousness		0.66	12.94	0.00	-0.03	0.03	0.84	-0.31	1.6	0.2
Mavenism		0.02	0.01	0.91	0.16	0.65	0.42	-0.26	1.12	0.29
Innovativeness		-0.31	3.73	0.04	0.23	1.59	0.2	0.09	0.16	0.68
Shopping enjoyment		0.14	1	0.31	0.07	0.21	0.64	0.39	4.81	0.05
Risk aversion		0.08	0.2	0.65	-0.07	0.15	0.69	0.79	7.47	0.00
Time pressure		-0.28	4.95	0.02	0.23	2.75	0.09	-0.23	1.39	0.23
Motivation to confirm		-0.04	0.1	0.75	0.1	0.36	0.54	-0.6	7.14	0.00
Gender	Male	-0.79	7.47	0.00	-0.16	0.27	0.6	-0.23	0.3	0.58
	Female	.00	-	-	.00	-	-	.00	-	-
Age	18-25	-1.57	8.53	0.00	0.18	0.09	0.76	-1.84	5.26	0.02
	25-30	-0.29	0.48	0.48	0.7	1.99	0.15	-0.94	2.3	0.12
	30-35	0.17	0.2	0.65	1.63	13.41	0.00	-0.79	2.04	0.15
	35-40	0.28	0.57	0.44	0.96	4.23	0.04	-0.32	0.36	0.54
	>40	.00	-	-	.00	-	-	.000	-	-
Years of education	<12	1.71	1.78	0.18	1.24	0.74	0.38	-0.86	-0.2	0.65
	12	1.54	1.71	0.19	1.04	0.63	0.42	-1.09	0.38	0.53
	14	1	0.75	0.38	0.45	0.12	0.72	-1.21	0.5	0.47
	16	1.39	1.51	0.21	1.45	1.26	0.25	-0.99	0.35	0.55
	18	1.07	0.98	0.32	1.72	2.03	0.15	-1.2	0.55	0.45
	22	.00	-	-	.00	-	-	.00	-	-
Income	0.5-1	-0.93	3.69	0.05	-2.13	17.77	0	-0.53	0.57	0.44
	1-1.50	0.14	0.08	0.76	-1.5	8.31	0	0.75	1.2	0.27
	1.5-2	-0.43	0.88	0.34	-1.48	9.73	0	-0.38	0.36	0.54
	2-2.5	0.25	0.24	0.28	-1.14	4.38	0.03	0.23	0.1	0.74
	2.5-3	1.01	3.52	0.06	-0.2	0.13	0.71	-0.56	0.5	0.47
	>3	.00	-	-	.000	-	-	.000	-	-
Marital Status	Married	0.23	0.53	0.46	-0.16	0.22	0.63	-0.12	0.05	0.81
	Single	.000	-	-	.000	-	-	.000	-	-

4. Discussion and conclusion

According to Table 1, Table 2 and Fig. 3, which show the overall channel orientation of consumer segments and also psychographic and demographic variables that have a significant role in profiling consumers in each segment, following results about labeling and profiling consumer segments can be concluded:

Multichannel-seekers (Segment 1): This segment includes 37.63% of total consumers. This group of consumers has favorable attitudes toward multi-channels for information searching but prefer store channel for purchasing. In terms of demographic characteristics, aggregation of young consumers is very low in this segment and most of them are with an income level of between 25 million to 30 million Rials. In addition, most of multichannel-seekers are women. In terms of psychographic characteristics, Multichannel-seekers are quality-conscious consumers with low level of innovativeness and time pressure.

Uninvolved shoppers (Segment 2): This segment includes 27.38% of total consumers and includes consumers with generally neutral attitudes toward all channels for both searching information and purchasing phases.

Multichannel consumers (Segment 3): This segment, which includes 23.88% of total consumers, can be assumed as the multichannel segment. This group of consumers shows favorable attitudes toward all three channels for searching information and purchasing. In terms of demographic characteristics, consumers of this segment are within the middle age groups like 30-40, and most of them are with an income level of above 30 million Rials. In terms of the psychographic characteristics, this segment includes consumers with high level of time pressure.

Store-focused consumers (Segment 4): This consumer segment that is the smallest and can be known as the traditional part of the market, includes 11.13% of total consumers. In terms of demographic characteristics, most of the consumers in this segment are within the high age groups. In terms of psychographic characteristics, this segment includes consumers that are risk averse with high level of shopping enjoyment and motivation to confirm.

Acknowledgement

References

Ansari, A., Mela, C. F., & Neslin, S. A. (2008). Customer channel migration. *Journal of Marketing Research*, *45*(1), 60-76.

Bhatnagar, A., & Ghose, S. (2004). A latent class segmentation analysis of e-shoppers. *Journal of Business Research*, *57*(7), 758-767.

Chandon, P., Wansink, B., & Laurent, G. (2000). A benefit congruency framework of sales promotion effectiveness. *Journal of marketing*, *64*(4), 65-81.

Chen, J. S. (2003). Developing a travel segmentation methodology: A criterion-based approach. *Journal of Hospitality & Tourism Research*, *27*(3), 310-327.

Darian, J. C. (1987). In-home shopping: are there consumer segments?. *Journal of Retailing*, 63(2), 1987, 163-186.

Dolnicar, S., & Leisch, F. (2003). Winter tourist segments in Austria: Identifying stable vacation styles using bagged clustering techniques. *Journal of Travel Research*, *41*(3), 281-292.

Grønholdt, L., & Martensen, A. (2005). Analysing customer satisfaction data: a comparison of regression and artificial neural networks. *International Journal of Market Research*, *47*(2), 121-130.

Inman, J. J., Shankar, V., & Ferraro, R. (2004). The roles of channel-category associations and geodemographics in channel patronage. *Journal of Marketing*, *68*(2), 51-71.

Kumar, V., & Venkatesan, R. (2005). Who are the multichannel shoppers and how do they perform?: Correlates of multichannel shopping behavior. *Journal of Interactive Marketing*, *19*(2), 44-62.

Konuş, U., Verhoef, P. C., & Neslin, S. A. (2008). Multichannel shopper segments and their covariates. *Journal of Retailing*, *84*(4), 398-413.

Rangaswamy, A., & Van Bruggen, G. H. (2005). Opportunities and challenges in multichannel marketing: an introduction to the special issue. *Journal of Interactive Marketing*, *19*(2), 5-11.