A market segmentation process based on usage context

Jiliang Wang, Bernard Yannou

A MARKET SEGMENTATION PROCESS BASED ON USAGE CONTEXT

Jiliang WANG, Bernard YANNOU

Ecole Centrale Paris, Laboratoire Genie Industriel, Grande Voie des Vignes, 92290, Châtenay-Malabry, France

ABSTRACT

Designing an appropriate product or service requires to know the targeted markets. Traditionally, market segmentation methods are based on socio-demographic data. In many situations, it is not sufficient enough to elicit relevant segments with clear orientations. We propose here to add usage context data so as to define relevant market segments before leading kansei studies in each segment. The case study of cutting a wood board with a jigsaw tool is taken in order to exemplify how usage context variables may influence segments.

Keywords: **usage model, usage context, market segmentation, cluster analysis**

1. INTRODUCTION

The analysis of customer's requirements and preferences is an important stage for product design. Due to limits of company's budget and marketing strategy, managers are always willing to choose profitable segments as target market segments for launching new products or redesigning existing products. Prediction of the profit for a given product in a given market segmentation is essential for the decision model of a company. In turn, it requires to model the demand to model the cost. So the precision of market segmentation, which determines the quality of demand prediction, becomes especially important. Market segmentation consists in defining homogeneous groups of customers and it is one of the building blocks of effective product or product family positioning. Products or services can no longer be designed and sold without considering customer needs and recognizing the heterogeneity of those needs. Various statistical techniques are available in order to identify market segments, among these, cluster analysis and latent class models. For the segmentation variables, traditional market segmentation methods are always based on marketing attributes and demographic attributes of customer, or correlations between them [1]. Nowadays furious market competition requires that companies design specific products for specific customer segments. Only slight differences from the competitors' solutions may lead to great successes if actual customers' requirements are captured. Traditional market segmentation modeling could not distinguish such latent attributes. Moreover, as customers become more conscious of their expectations on the products, markets often continue to be ever more fragmented into more technical niches. Exploiting the rich contextual information which exists in customers' product usage becomes more and more crucial for product design, especially in the situation that the customer is experienced and able to imagine the probable performances of the product regarding to their usage. As a result, traditional marketing segmentation methods should be extended to contain contextual attributes, in order to drive more efficient kansei analysis in each context segment.

In this paper, we begin by a general literature review of usage context models and market segmentation methods in research. Next, we present the basic principles behind our general usage model and show how it may be integrated into a method of market segmentation. Finally, we apply the usage context market segmentation method in a brief case study of cutting a wood board with a jigsaw tool in order to exemplify how usage context variables may influence segments. We conclude by reaffirming the utility of this integrated method and outlining the further research work that will be done in the future.

2. STAT-OF-THE-ART AND OBJECTIVE OF RESEARCH

Segmentation is a grouping task for which a large variety of methods are available. Segmentation methods can be classified in a-priori, when the type and number of segments are determined in advance by the researcher, and post-hoc, when the type and number of segments are determined on the basis of results of data analysis. Clustering methods are the most popular tools for descriptive segmentation both in research and industrial applications [2]. They are employed extensively in segmentations based on psychographics, benefits seeking or conjoint partworth data [3]. In general, researchers want to find individual segments that are internally as much homogeneous as possible. K-means cluster analysis is the most frequently used to find homogenous segments of customer, since it can accommodate the large sample size with market segmentation studies. A comprehensive review of K-means methods and of more efficient genetic combined K-mean methods can be found in [4].

The needs of customer may be considered as Direct Needs and Latent Needs. Direct Needs are those that customers have no problem to express as something they are concerned about when sollicitated to talk about the product. Latent Needs are typically not directly expressed by customers, then some stimulations must be done for customers to talk about [5]. Usage contextual information contains and reflects user's unrevealed requirements during product application. Meanwhile, designer can comprehend the contextual information and help user probing these latent needs. Context is any information that can be used to characterize the situation of a person, place, or object that is considered relevant to the interaction between a user and an application, including the user and applications themselves [6]. While the formal study of usage context permeated the field of marketing years ago, it has only just begun to be applied directly to the design and engineering of new products. Green et al. have published three successive papers on the subject [7][8][9], with the goal of forming a comprehensive product design methodology that includes contextual factors. In these papers, important first steps in the field were taken, including the definition of key terms and concepts in design engineering: “usage context, as it relates specifically to products, is defined as the unique combination of application and environment in which a product is used.” Furthermore, usage context is framed as one part of a larger product design context, which also includes market and customer context [9]. During the course of the studies, customers were found to have distinct product preferences under different usage contexts. In the domain of hi-tech product design, context-aware systems - knowing the activity context and taking it into account for system behavior – are emerged. Context-aware system for mobile cartography has been shown in [10], which uses formalization to describe situations and contexts for finding typical context patterns. Additionally, evidence supported that contexts could be differentiated based upon functional attributes, indicating a link between engineering parameters and perceived usefulness, which occurs under the influence of different usage contexts.

So our objective in this paper is to investigate deeply the product contextual information and integrate it while segmenting market with cluster analysis method. Latent requirement pattern can be found from resulting segments. In each context segment, customers’ usage behavior and demographic attributes are quite homogeneous, which could serve as a basis for future kansei analysis or further experiments.

3. MODEL AND METHODS

3.1. Usage Context Model

We have proposed a Usage Context Model so as to get a more thorough marketing model based on sets of permitted usages for a product-service instead of the conventional perceived marketing attributes [11]. In this model, customers are understood as product employers and products as service providers. This method proposes to quantify individuals’ performances during product usage, depending on the usage context and on the personal skills of the individuals. It offers the advantage of linking with user experience to introduce the perceived quality of a product’s service, as well as to consider particular service delivery conditions.

In our usage context model, a product is defined by its design parameters X, such as physical dimensions of product. Variable set E represents any variables that describe the conditions under which the product is used to provide the service; this is the usage context. Cs are user-related parameters that affect performances. Then one states that individuals’ performances Y during product usage depend on the product itself X, on the context E and on the user’s skills Cs. Such performance equations could either be obtained by physics-based model or rating-based regression model.

$$Y = f\left(X, E, C_s\right)$$

A usage needed is defined as a set of expected service contexts Ei associated with a usage percentage Fi.

$$U_{needed} = \left\{\left(E_i, F_i\right)\right\}$$

The usage context E consists of the environment of usage and the object of usage. One example of building UCM for the situation of using jigsaw to cut wood boards or sticks is illustrated in the work [11], where performance values are established using a physics-based model.

3.2. Process of Usage Context Information Collection

The process of usage context information collection is very crucial for the whole usage context based market segmentation process. Because the kind of contextual information collected and its quantity play an important role in the following clustering process. Figure 1 reveals the principal stages of this process.

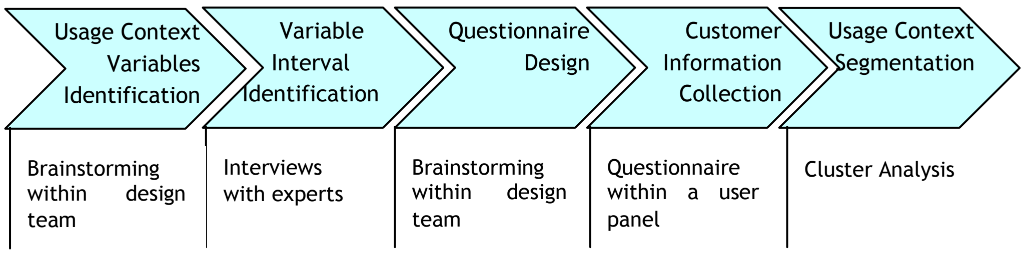

Figure 1: Process of Usage Context Information Collection

First of all a brainstorming is carried out among design team members to identify the important usage context variables. Brainstorming is well-known for generating a flood of new ideas, in which a group of open-minded designers from different spheres of life bring up any thoughts that are related to usage context variables of a given product. These variables can describe properly the conditions under which the product is used to provide the service.

The intervals of the context variables are identified by interviews with experts: here design team members discuss the usage context information one single customer at a time. The interviewed customers are chosen to be “experts” on the product usage. They can estimate the importance of each context variable and its possible interval value. The interviews are held in the customer’s environment, where the customer uses the product. The design team records the customer responses.

In the questionnaire, the team develops a list of criteria it believes relevant to the customer’s concerns and organizes a customer panel for responses. The collected questionnaires are treated and translated to ordinal data for further cluster analysis process. A process example of jigsaw product case study is shown in the following section.

3.3. Cluster Analysis for Usage Context Segmentation

In order to ultimately find out profitable segments for product design, only the individual decision is insufficient. For instance, the customers with similar usage contexts for whom a single type of saw may be designed and marketed in case of cutting wood requirement. Interior homogeneous segments should be built first. Cluster analysis can be conducted to find similar segments of a population based on Environmental Context (E), User skill (Cs), Demographic attributes (S), or the complete set [E, Cs, S]. The following cluster types were examined from a hypothetical data set created for that aim:

- Clustering by Object (subset of E).
- Clustering by Environment (subset of E).
- Clustering by Skill and Demographics (Cs, S)
- Clustering by All attributes (E, Cs, S).

Usage similarity must take all usage environments and skill level attributes into account. Two methods are used for identifying similar usage segments: Latent Class Analysis, used when the usage attributes are categorical (e.g. type of wood) or ordinal (e.g. skill level on a scale of 1-3) measures. Latent class analysis enables a characterization of categorical latent (unobserved) variables from an analysis of the structure of the relationships among several categorical manifest (observed) variables. Maximum likelihood estimates (MLEs) of the conditional and latent class probabilities is calculated. [12]. Cluster Analysis when the usage variables are continuous (e.g. thickness of cut to be made). These methods are capable of identifying the optimal number of latent classes or clusters, and all the usage contexts which belong to each latent class or cluster.

4. EXPERIMENTATION

An example of “cutting wood sticks and boards” service is used. A given user may use a given saw tool. The experimental activities intend to reveal the usage context variables.

4.1. Saw Service Usage Context Information Identification and Collection

A brainstorming within the research group is carried out to identify the relevant saw usage context variables, so do several expert experiments to identify possible value interval.

Figure 2: Saw Service Experiments

The service context variables E may be:

- The location where the cutting operation must or is supposed to take place:

$$location \in \{\text{apartment,house}\}$$

- The type of wood to be cut: $Wood \in \{\text{hard,medium,soft}\}$
- The shape of object: $Form \in \{\text{stick,board}\}$
- The object thickness: $Thickness \in \{\text{thin,medium,thick}\}$
- The object size: $Size \in \{\text{small,medium,large}\}$
- The lighting condition: $Lighting \in \{\text{Yes/No}\}$
- The presence of a workbench: $Workbench \in \{\text{Yes/No}\}$

The customer related attributes S may influence preference of product, such as income, etc. Cs in saw case is user-related parameters that affect performances, but not preference, e.g. user's skill. While some user related parameters can belong to both types, e.g. gender.

A panel of users gives their opinions for their usage context information by answering a questionnaire which requires simply choosing among several categories about their usage of jigsaw. The sample questionnaire is shown in the Appendix A, which supposes that all the users are given a product of jigsaw BOSCH with specification X. An example of questionnaire for single usage situation can be found at: http://www.diaochapai.com/c3da8833-6005-4ba9-9104-add2e357c733; Internet version eases the collection of survey information.

The semantic answers of the responders are translated into ordinal values. A matrix of each user's usage context and demographic pattern is build based on their answer sheet. Appendix B shows an example of simulated answer data of 30 users on their typical (single) usage information and their demographic information.

4.2. Cluster Analysis for Usage Context Information

First of all, a two-step clustering method is carried out to find a relative good final number of clusters. The main process of two-step clustering is to group cases into pre-clusters, and then standard hierarchical clustering is applied to the pre-clusters in the second step. This method is used when one or more of the variables are categorical (not interval or dichotomous). We have assumed that data are interval in level or are true dichotomies for hierarchical and k-means clustering, though two-step clustering can handle categorical data. So when at least one variable is categorical, two-step clustering must be used. One experimental hypothesis is that all observations are independent. Randomization of cases is recommended. Because K-means and two-stage cluster analysis usually generate different solutions, depending on the sequence of observations in the dataset.

We have try SPSS for our case with usage context variables. Since our variables are mainly categorical, only the process of two-step clustering is applicable. The process gives final classes based on preset criteria. If we let the process choose the final clusters, based on Akaike's information criterion (AIC), a measure of the goodness of fit of an estimated statistical model. Given a data set, several competing models may be ranked according to their AIC (see Figure 3), with the one having the lowest AIC being the best.

Auto-Clustering

Number of...	Akaike's Information Criterion (AIC)	AIC Change[a]	Ratio of AIC Changes[b]	Ratio of Distance Measures[c]
1	536.303			
2	474.811	-61.492	1.000	2.123
3	461.720	-13.091	.213	1.032
4	449.962	-11.757	.191	1.135
5	443.168	-6.795	.111	1.036
6	437.660	-5.507	.090	1.309
7	440.540	2.880	-.047	1.293
8	449.568	9.027	-.147	1.010
9	458.795	9.227	-.150	1.131
10	470.433	11.638	-.189	1.020
11	482.438	12.005	-.195	1.269
12	498.263	15.824	-.257	1.027
13	514.454	16.191	-.263	1.023
14	530.958	16.504	-.268	1.217
15	549.867	18.910	-.308	1.140

Figure 3: Auto-Clustering results of AIC

We can see from the resulting table of Figure 3 that 6 clusters satisfy the AIC criterion. So another constrained two step cluster analysis is carried out with SPSS to find the 6 optimal clusters. Table 1 shows how the 30 response cases are distributed in the 6 final clusters.

Table 1: Two Step Cluster Analysis Results

ID	1	7	16	29	6	12	15	21	28	2	4	10	14	19	23	25	9	17	24	27	5	11	18	22	26	30	3	8	13	20
6	1	1	1	1	2	2	2	2	2	3	3	3	3	3	3	3	4	4	4	4	5	5	5	5	5	5	6	6	6	6

The 6 typical usage context and demographical combined clusters are detailed in table 2.

Table 2: The 6 typical usage context and demographical combined clusters

ID	Hardness	Size	Thickness	Form	Location	Lighting	Workbench	Gender	Incomes	Skill
- Skilled user, in apartment, cutting small hard thin stick, with lighting and workbench										
1	1	1	1	2	1	1	1	1	3	3
7	1	1	1	2	1	1	1	1	1	2
16	1	1	1	2	2	1	2	2	1	3
29	1	3	1	2	1	1	1	2	1	3
- Medium skill female user, cutting large thick board, with lighting and workbench										
6	2	3	3	1	1	1	1	2	1	2
12	2	3	3	1	2	1	1	2	1	2
15	2	3	3	1	1	1	1	2	2	2
21	1	2	3	1	1	1	1	2	2	3
28	1	1	2	1	2	1	1	2	2	2
- Beginner male user, cutting thick board, with lighting and workbench										
2	1	3	3	1	1	1	1	1	1	1
4	1	2	3	1	2	2	1	1	1	3
10	2	3	3	1	1	1	1	1	3	1
14	2	3	3	1	1	1	1	1	3	1
19	1	2	3	1	2	1	1	1	1	1
23	2	2	3	1	1	1	1	1	2	1
25	1	3	3	1	1	1	2	2	1	1
- Male medium income user, cutting thin stick in apartment, without lighting, with workbench										
9	1	2	1	2	1	2	1	1	2	3
17	3	3	1	2	1	2	1	1	2	1
24	3	3	1	2	2	1	1	1	2	3
27	1	2	2	2	1	2	2	1	2	1
- Medium skill male user, cutting thick board, without lighting and workbench										
5	2	3	3	1	1	2	2	1	2	1
11	2	2	3	1	1	2	2	1	3	2
18	1	3	2	1	1	2	2	1	2	2
22	1	1	3	1	2	2	2	1	2	2
26	2	1	3	1	2	2	1	1	3	2
30	1	3	3	1	1	2	2	1	1	2
- Medium skill female user, cutting medium thickness stick, without workbench										
3	2	2	2	2	1	1	2	2	2	2
8	3	3	2	2	2	1	2	2	1	1
13	3	3	2	2	2	2	2	2	2	2
20	1	2	2	2	1	2	2	2	1	2

As one may notice, the possible combinatory of attributes is huge, but the number of clusters is finally acceptable. Our method can transform a large number of usage and user information into a manageable number of usage patterns. Extreme and abnormal usage context situations with very few users are merged with similar usage segments. Finally, six typical usage and user patterns are derived: with certain attributes extremely similar, while other attributes can be little variant. The identical attributes will dominate as the number of survey increases. Take the example of last usage context segment: - Medium skill female user, cutting medium thickness stick, without workbench; but their hardness, size, location, lighting, income can take a spectral value, which leads to several merges when one considers design products for this niche.

As a result, market segmentation should be based on usage context and demographic attributes. The practical application of such a method is shown in the case, for example, assuring that members of a cluster are similar on several measures and would therefore all be more likely to prefer the same saw design. For example, if we exclude demographics from the clusters, we may design a saw that meets the object and environment context well for a given cluster, but may only appeal to a certain income segment within the cluster.

5. CONCLUSION AND PERSPECTIVES

In this paper, we integrate usage context information in traditional market segmentation methods. Cluster analysis is performed to different type of segment decision variables. As shown in the results, segmentation method which integrates usage model performs complete market segments, which are internally homogeneous. Different usage context patterns can be found from the results. Within each of these usage context patterns, a kansei analysis process may be carried out according to better investigate perceptual and emotional users' expectations.

REFERENCES

1. Green P.E., A guide to market segmentation, MasterCard International, 2000
2. Francesca Bassi, Latent class factor model for market segmentation, Stat. Meth. & Appl., Vol.16, p279-287, 2007.
3. Green P.E., Abba M. Krieger, Segmenting markets with conjoint analysis, Journal of Marketing, Vol. 55, p20-31, 1991
4. Kuo R.J. et al., Integration of self-organizing feature maps neural network and genetic K-mean algorithm for market segmentation, Expert Systems with Applications, Vol. 30, p313-324, 2006.
5. Otto K., Wood K., Product Desing: Techniques in Reverse Engineering and New Product Development, Prentice Hall, 2001
6. Anind K. Dey, Understanding and Using Context, Personal and Ubiquitous Computing, Vol.5, P.4-7, 2001
7. Green, M. G., R. P. K. Palani and K. L. Wood, Product Usage Context: Improving Customer Needs Gathering and Design Target Setting, ASME Design Engineering Technical Conferences, Salt Lake City, 2004.
8. Green, M. G., J. Tan, J. S. Linsey, C. C. Seepersad and K. L. Wood, Effects of Product Usage Context on Consumer Product Preferences, ASME / DTM, Long Beach, CA, 2005.
9. Green, M. G., J. S. Linsey, C. C. Seepersad and K. L. Wood, Frontier Design: A Product Usage Context Method, ASME IDETC/DTM 99608, Philadelphia, PA, 2006.
10. Reichenbacher T, Adaptive Methods For Mobile Cartograph, Proceedings of the 21st International Cartographic Conference (ICC), Durban, South Africa, 10 – 16 August 2003
11. Yannou B., Chen W., Wang J., Hoyle C., Drayer M., Rianantsoa N., Alizon F., Mathieu J.-P., Usage coverage model for choice modeling: principles, Proc. IDETC/DAC:ASME, August 30 - September 02, San Diego, CA, 2009.
12. Allan L. Mccuthcheon Latent Class Analysis, a SAGE University Paper, 1987

6. APPENDIX A: SAMPLE SURVEY QUESTIONAIRE TO IDENTIFY E, CS, S

Assume that you are given a jigsaw BOSCH as shown in the picture. What is your most usual usage context when you cut a piece of wood? Pleas answer the following questions:

1. Please tell us a little bit about yourself:

o Are you:

□ ,. male □ ,. female

o What is your skill level in terms of saw usage?

□ ,. Beginner □ ,. Intermediate □ ,. Experienced

o What is your monthly income?

□ ,. Low (<1000€) □ ,. Medium (1000€ ~3000€) □ ,. High (>3000€)

2. Please tell us about your typical saw usage situation:

o What kind of wood hardness do you mostly cut with a jigsaw?

□ ,. Soft wood □ ,. Intermediate □ ,. Hard wood

o What kind of size do you usually cutting with a jigsaw?

□ ,. Small size □ ,. Intermediate □ ,. Large size

o What thickness of wood do you mostly cut with a jigsaw?

□ ,. Thin □ ,. Intermediate □ ,. Thick

o What type form of wood do you mostly cut with a jigsaw?

□ ,. Board □ ,. Stick

o Where do you live in?

□ ,. Apartment □ ,. House

o Is there lighting system when you are doing your saw job?

□ ,. Yes □ ,.No

o Do you have a workbench while you cutting wood?

□ ,. Yes □ ,.No

7. APPENDIX B: QUESTIONAIRE RESPONSES DATA

ID	Hardness	Size	Thickness	Form	Location	Lighting	Workbench	Gender	Incomes	Skill
1	1	1	1	2	1	1	1	1	3	3
2	1	3	3	1	1	1	1	1	1	1
3	2	2	2	2	1	1	2	2	2	2
4	1	2	3	1	2	2	1	1	1	3
5	2	3	3	1	1	2	2	1	2	1
6	2	3	3	1	1	1	1	2	1	2
7	1	1	1	2	1	1	1	1	1	2
8	3	3	2	2	2	1	2	2	1	1
9	1	2	1	2	1	2	1	1	2	3
10	2	3	3	1	1	1	1	1	3	1
11	2	2	3	1	1	2	2	1	3	2
12	2	3	3	1	2	1	1	2	1	2
13	3	3	2	2	2	2	2	2	2	2
14	2	3	3	1	1	1	1	1	3	1
15	2	3	3	1	1	1	1	2	2	2
16	1	1	1	2	2	1	2	2	1	3
17	3	3	1	2	1	2	1	1	2	1
18	1	3	2	1	1	2	2	1	2	2
19	1	2	3	1	2	1	1	1	1	1
20	1	2	2	2	1	2	2	2	1	2
21	1	2	3	1	1	1	1	2	2	3
22	1	1	3	1	2	2	2	1	2	2
23	2	2	3	1	1	1	1	1	2	1
24	3	3	1	2	2	1	1	1	2	3
25	1	3	3	1	1	1	2	2	1	1
26	2	1	3	1	2	2	1	1	3	2
27	1	2	2	2	1	2	2	1	2	1
28	1	1	2	1	2	1	1	2	2	2
29	1	3	1	2	1	1	1	2	1	3
30	1	3	3	1	1	2	2	1	1	2