Marketing Mix Decision Making Using Scanner Data and Self-Organizing Maps

Krycha, Karl A. and Gerhard Kranner[1]

Keywords: Self-Organizing Maps, price elasticity, marketing mix planning, graphical representation

Abstract

During the last few years, an increasing number of retailers have introduced point of purchase scanning facilities allowing for the permanent collection of up-to-date sales information. Besides that development the introduction of Artificial Neural Network (ANN) modeling approaches into marketing science have put contemporary managers in a position making it easily possible to apply accurate marketing management instruments for marketing mix decisions making.

In our contribution we propose an application of Self-Organizing Maps (SOM) to the problem of managing the product line with respect to price decisions. A data set containing a wide range of product information is used to simultaneously establish purchase patterns as well as information on expected total sales. This is accomplished in the proposed ANN model by considering price elasticities for the products offered. The marketing decision process is supported by the visualization of the relevant marketing mix information for every customer profile detected in the sample.

1. Introduction

Originally, research in the area of ANN was oriented towards representing real biological systems. Due to the high complexity of these systems only heavily simplified artificial networks have been considered for computational purposes so far. Our special interest is in the performance of Artificial Neural Networks in the domain of marketing. Although marketing seems to offer a lot of application opportunities for neural nets, the first of these applications are appearing in the literature only lately. In their paper, Wierenga and Kluytmans presented examples for the prediction of television audiences for the BBC, the prediction of in-store and gasoline volume sales for *Specialty Consumer Outlets* and for choosing the most appropriate sales promotion tool (Wierenga and Kluytmans, 1994). The suitability of ANN for market segmentation can be found, e.g., in (Mazanec, 1993). A more detailed survey consisting of a sample of more than 90 papers published from 1990 to 1996 in the area of management science can be found in (Krycha and Wagner, 1998). The authors found that, based on the number of experiments documented in the investigated sample, the multilayer perceptron (MLP) has turned out to be the standard architecture used in 96% of all cases considered.

2. Problem Definition

2.1 Management of Product Lines

Typically, the management of product lines is confronted with two very closely related problems. On the one hand it has to *set up the marketing mix* for the product mix offered to its customers. This consists of determining the most favorable levels for all available marketing instruments. As pricing decisions turned out to be most important in this context, we will concentrate on decision making concerning price. Price lining has to decide how to specify particular pricing points within the offered product line. This process consists of determining the lowest-priced product and its price, determining the highest-priced product and its price, and determining the price differentials for all other products in all lines offered. The products in each line are likely to have

[1]Dr. Gerhard Kranner is managing director and senior consultant of eudaptics software gmbh, Vienna and Dr. Karl A. Krycha is assistant professor at the Institute of Management, University of Vienna.

Addresses for correspondence:

Dr. Karl A. Krycha, University of Vienna, Institute for Management Science, Brünner Straße 72, A-1210 Vienna, phone: +43.1.29.128-434, e-mail: karl.krycha@univie.ac.at.

Dr. Gerhard Kranner, eudaptics software gmbh, Helferstorferstraße 5/8, A-1010 Vienna, phone: +43.1.532.05.70, e-mail: office@eudaptics.com.

interrelated demand considerations, competitive characteristics, and cost interrelationships. As an objective in pricing, the marketer seeks to maximize profits for all lines but not necessarily of each individual item.

On the other hand, marketing management is fine-tuning its activities for various sets of consumers. This comprises the *structuring of heterogeneous markets (profiling)* in order to identify homogeneous clusters. More precisely, profiling is an essential step in *target marketing* which can be separated into the identification of market segments, the selection of one or more of them, and the subsequent development of marketing mixes tailored to each. The cross-sectional point where pricing decisions and profiling links together can be found in the consumers' perception of the product line with respect to its own price expectations. With too low a price at the bottom, the whole line may be considered cheap; with too high price at the top, buyers may perceive the line as above the mass market. Additionally, price differences must make sense to customers (cf. price signaling; Kinnear, Bernhardt and Krentler, 1995).

2.2 Data Sources for Marketing Mix Decision Making

In this particular situation, the decision process can be significantly streamlined with the assistance of a data-based methodology that would help to more accurately predict the outcomes of possible management interactions. The decisions are of concern to wholesalers and retailers (e.g. supermarkets). The methodology used should not only help to find new groups of consumers but also to undertake a simultaneous grouping of customers, brands, and/or products. The development of new sources of buyer behavior data now makes it possible for the marketing man to introduce a wide range of new methods to this discipline.

A good example for these new opportunities can be found in the retailing industry of consumer durables. Compared to former time, the situation has improved significantly today: progress in bar code technology has made it possible for retail organizations to collect and store massive amounts of sales data. A record in such data typically consists of the transaction date and the items bought in the transaction. This kind of data is sometimes referred to as *basket data* (Agrawal and Srikant, 1994). Successful organizations view such databases as important pieces of their marketing infrastructure. These businesses are interested in understanding information contained in their data in order to use it most effectively for their business advantage. Therefore, an increasing need for data-driven marketing management decision models has arisen. To investigate that kind of data we used a SOM-based approach capable of detecting relevant consumer segments and its reactions to price adjustments based on buying patterns and reactions to price adjustments.

3. Self-Organizing Maps for Price Lining

3.1 Basic Properties of Self-Organizing Maps

The basic properties of a SOM-type Artificial Neural Network are illustrated in Fig.1. This type of unsupervised learning neural network defines a nonlinear projection from the input data space $\Re^n$ onto a regular two-dimensional array of nodes.

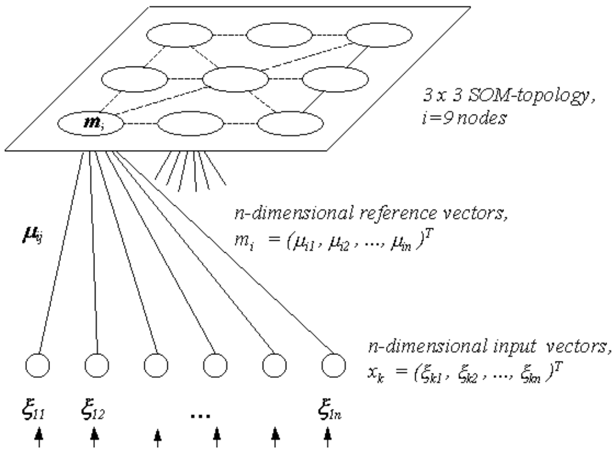

Fig.1. Basic properties of a SOM-type Artificial Neural Network

With every node i in the mapping, a parametric reference vector $m_i \in \Re^n$ is associated. The lattice type of the array can be defined to be rectangular or hexagonal. In the course of learning, an input vector $x \in \Re^n$ is presented to the network and compared to its codebook vectors m_i. The best match is defined as *response* and the input is mapped onto this location. The comparison itself may be undertaken in any metric, in practical applications the smallest of the Euclidean distances is made to define the best matching node, signified by the subscript c :

$$\|x - m_c\| = \min_i \|x - m_i\| \ . \tag{1}$$

During learning, those nodes that are topographically close in the array up to a certain distance will activate each other to learn from the same input (t indicates the progress step in the course of learning). The construction of the clusters mapped is done by stochastic approximation considering in each iteration step several neighboring nodes in the lattice (Bock, 1997). Useful values of the m_i can be found as convergence limits of the following learning process, whereby the initial values of $m_i(0)$ can be chosen randomly (Kohonen, 1997):

$$m_i(t+1) = m_i(t) + h_{ci}(t)\left[x(t) - m_i(t)\right]. \tag{2}$$

In this relationship, $h_{ci}(t)$ indicates the so-called neighborhood function, a smoothing kernel defined over the lattice points. Usually, $h_{ci}(t) = h(\|r_c - r_i\|, t)$, where $r_c \in \Re^2$ and $r_i \in \Re^2$ are the radius vectors of nodes c and i , respectively, in the array. With increasing $\|r_c - r_i\|$, $h_{ci} \to 0$. The average width and form of h_{ci} defines the stiffness of the elastic surface to be fitted to the data points. Widely applied neighborhood kernels are the bubble neighborhood or the Gaussian neighborhood; see (Kohonen, 1997) for more details.

3.2 Training and Visualization of Results

A typical set of scanner data comprises several different variables directly usable for marketing-mix decision making. E.g., (Klapper, 1998) reports an american single source data set on ketchup data containing weekly information on price (cents per ounce), line advertising (line ad), major advertising (major ad), end-aisle display (end-aisle), front-aisle display (front-aisle), in-aisle display (in-aisle), and store coupon. For our purposes, a subset of about 2 980 observations of an Austrian wholesaler scanner data set is used. It contains information on the demand structure for three different color-TV sets, $b = A, B, C$. The collection took place for 24 months in Austria (t = 1, ..., 24) and consists of marketing mix information (i.e., product sold b, price $p_{b,t}$, quantity ordered $q_{b,t}$, and date of ordering t) as well as apriori customer profiling information (i.e., customer classification code and granted rebates).

The database comprises entries on a day-by-day basis; each observation accounts for one single purchase incidence. After appropriate aggregation for each product b specific demand patterns emerged. In Fig.2 the total sales of product A, B, and C are displayed.

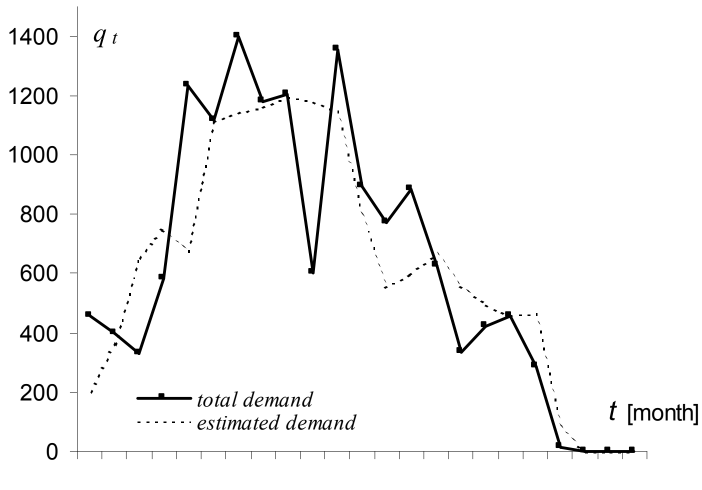

Fig.2. Aggregated demand for products A, B, and C

As can be seen in this representation, demand follows a life cycle pattern. This specific behavior is also reflected in the individual demand figures for each product not shown here. In order to capture the effects of marketing mix onto the demand variations, we compared several different modeling approaches. A market response function comprising a time dependent life cycle as well as a price component turned out to fit best in this particular situation:

$$q_{b,t} = a_b + \lambda_b r_b^{\,t} q_{b,t-1} - b_b p_{b,t}, \quad \forall b. \tag{3}$$

In this equation we account for carryover by λ_b and for obsolescence of the product by r_b (Simon, 1992, p. 260). For each product considered the price elasticity $\eta_{b,t}$ is defined to be:

$$\eta_{b,t} = \frac{dq_{b,t}}{dp_{b,t}} \frac{p_b}{q_{b,t}} = \frac{-b_b p_{b,t}}{a_b + \lambda_b r_b^{\,t} q_{b,t-1} - b_b p_{b,t}}, \quad \forall b. \tag{4}$$

The parameters for (3) are estimated independently for each brand by nonlinear least square optimization. The results of this step as well as additional information on the data set are given in Table 1. In particular, product *B* is marketed throughout all periods considered whereas product *C* is introduced to the market after 6 months in order to substitute product *A*.

	periods on the market	*purchase incidences*	*quantity sold*	a_b	λ_b	r_b	B_b	*SSE*)*
product A	2 – 7	468	1588	-1100.32	-0.05	0.98	-0.12	30615
product B	3 – 20	1279	8313	-1640.56	-20.04	0.42	-0.24	566469
product C	7 – 21	1582	5086	27059.64	0.03	1.21	2.57	470835
total		*2980*	*14987*					*1067919*

**) R^2=0.74*

Table 1. Estimation results for product line

According to (4), we estimated $\eta_{b,t}$ and used this aposteriori customer profiling information to characterize price sensitivity for every product in the line for $t = 1, ..., T$. To train the map, we used the Viscovery SOMine-package from Eudaptics that is based on an enhanced version of the Kohonen algorithm. Through the implementation of new techniques, such as SOM scaling, batch SOM, and quick match, the speed in creating maps is notably increased compared to the original SOM algorithm. The system further supports full pre- and postprocessing, cluster search, association/recall, prediction, statistics, filtering, and animated system state monitoring. The result of our SOM analysis is presented in a color display which allows for analyzing nonlinear relationships and read off all relevant dependencies. The results are visualized on a highly sophisticated graphical display (Fig.3).

The U-matrix (Kohonen, 1997, p. 126) representation of the *cluster landscape* (Fig.3a) represents the relative distances between neighboring codebook vectors by shades in a gray scale. If the average distance of neighboring m_i is small, a light shade is used and vice versa. The resulting cluster landscape formed over the SOM then clearly visualizes the classification result. Each cluster can be taken to characterize an individual customer profile k, $k = 1, ..., K$. The representation of the n individual *component planes* (Fig.3b) of the trained map (Kohonen, 1997, p. 226) is especially useful to get a detailed insight into the exact structure of the solution established. It uses pseudo color scales to indicate the values of the input vectors at the respective position on the map. (Colors have been transformed to gray scales in this paper.) As an example, in Fig.3b a representation of the component plane for $\eta_{b,t}$, $-1 \leq \eta_{b,t} \leq 1$ $\eta_{b,t}$ is given where dark gray shades indicate elasticity close to 0 and light shades elasticities close to 0.61. In the representation of the component plane for the date of ordering t (not shown in this paper) we identified the period of four weeks before Christmas. It can be seen that the map is ordering these codebook vectors labeled “X-mas” close together along a longitudinal area on the right-hand side. The situation of product A, B, and C has been identified in a similar way and is indicated by uppercase letters. For some of the retailing customers we also indicated their position in the display by indicating its names (COSMOS and MediaMarkt).

In our application of SOM to retailing data several clusters have emerged. As an example, we have indicated only one profile in the center of the map labeled *cluster 1.* Further definition of profiles is straightforward.

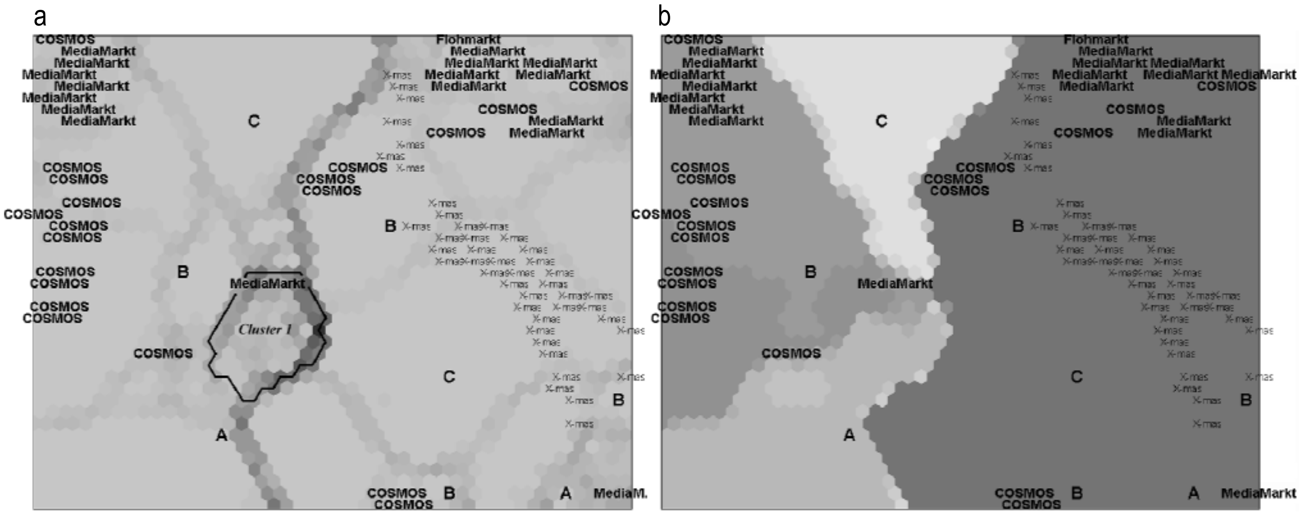

Fig.3. Trained map and component plane for price elasticities

For this customer profile we also derived the relevant marketing information for assessment of changes in price on the market response. (Cf. Table 2; for anonymity reasons we had to leave out some information and shifted the actual time frame.) As a result, the expected sensitivity to price changes in cluster 1 is given by the average price elasticity of 0.49; the average order comprises about 4 TV sets. Managerial relevance is added to the results by computing total sales for the whole product line resulting in 1,508,352 ATS for cluster 1 during the given time frame of 507 days (≈ 1½ years).

Component	Mean	Standard deviation	Minimum	Maximum	Sum
date of ordering [days]	34353	106	34232	34739	–
quantity ordered [#]	3.8	5.3	1	46	–
price [ATS]	–	639	8,462	10,709	–
sales [ATS]	37,713	–	–	–	1,508,352
elasticity []	0.49	0.18	0.0	0.6	–

Table 2. Marketing Mix information derived for cluster 1

Conclusions

Studies indicate that the amount of data in a given organization doubles every five years. This fact is a big opportunity to apply more accurate methodology to specific problems in marketing management. In our contribution, we presented an application of adaptive methods to marketing mix decision making using wholesaler scanner data. The use of locally approximated price perceptions allows to link together the previously separated tasks of consumer profiling and market structuring. As a marketing decision instrument, we demonstrated the use of a sophisticated SOM-based graphical display to support price lining of color TV-sets. Future investigations could extend the presented approach by taking into account consumer heterogeneity (e.g., by estimating (3) for every customer or groups of customers). Another possible avenue of our research is the representation of interproduct relationships by considering cross price elasticities or simultaneously estimating elasticities and training the map.

Bibliography

[1] Agrawal, R. and R. Srikant (1994): *Fast Algorithms for Mining Association Rules*, Proceedings of the 20th VLDB Conference, Santiago, Chile.

[2] Bock, H.H. (1997*): Simultaneous visualization and clustering methods as an alternative to Kohonen maps*, in G. Della Riccia, R. Kruse and H.-J. Lenz (eds.): Learning, networks and statistics, CIMS Courses and Lectures no. 382, Vienna et al.

[3] Ehrenberg, A.S.C. (1988): *Repeat-Buying. Facts, Theory And Applications*, 2nd edition, Oxford University Press: New York.

[4] Kinnear, T.C., K.L. Bernhardt, and K.A. Krentler (1995): *Principles of Marketing*, 4th edition, New York.

[5] Klapper, D. (1998): *Die Analyse von Wettbewerbsbeziehungen mit Scannerdaten*, Physika: Heidelberg.

[6] Kohonen, T. (1997): *Self-Organizing Maps*. 2nd edition, Springer: Heidelberg.

[7] Krycha, K.A. (1992): *A Comparison of Standard Methods with a Neural Network Approach to Forecast Univariate Time Series*, Research Memorandum No. 310, Institute for Advanced Studies, Vienna.

[8] Krycha, K.A. (1998): *Data—driven approaches for Market Segmentation and Consumer Profiling*, Proceedings of the 1998 AMA Marketing Exchange Colloquium.

[9] Krycha, K.A. and U. Wagner (1998): *Applications of Artificial Neural Networks in Management Science - a review*, to appear in Journal of Retailing and Consumer Services, Vol. 5, Issue 4.

[10] Mazanec, J. (1993): *Apriori and Aposteriori Segmentation: Heading for Unification with Neural Network Modeling*, Proceedings of the 22nd EMAC Conference, European Marketing Academy.

[11] Simon, H. (1992): *Preismanagement. Analyse – Strategie – Umsetzung*, Gabler: Wiesbaden

[12] Wierenga, B. and J. Kluytmans (1994): *Neural nets versus marketing models in time series analysis: A simulation study*, Proceedings of the 23th Annual Conference of the European Marketing Academy.