Identifying preferable product variants using similarity analysis

Günther Schuh, Michael Riesener*, Stefan Rudolf* 

*Laboratory for Machine Tools and Production Engineering, Department of Innovation Management, Steinbachstraße 19, 52074 Aachen, Germany  
* Corresponding author. Tel.: +49-241-80-28201. E-mail address: m.riesener@wzl.rwth-aachen.de

1. Introduction

In times of ever-increasing fragmentation of the markets, the ability to offer customized products at competitive prices is a crucial success factor for companies. Due to higher customer orientation, companies are forced to meet customer requirements despite small numbers and regardless of the resulting higher internal product complexity and process variance [1, 2, 3]. However, in many companies there is a lack of transparency concerning product variety and the costs and benefits related to it. Analyzing the orders in many companies from the field of mechanical engineering reveals the concept of cross-subsidization of product variants as shown in Figure 1.

In the past, many companies were able to deliver standard product variants to customers. Today there is a trend towards customer-individual product variants. Simultaneously the prices for these exotic product variants are often lower than they should be in comparison to the costs that are caused by this new variant, since dissimilar products are more expensive than standard variants due to higher adjustment costs for their development. As a consequence, exotic variants are cross-subsidized by standard variants, which causes a competitive disadvantage in the standard variant and a loss in many exotic variants (see Figure 1) [4]. Analyzing the sales process of many companies in this field conducted within the framework of several industrial projects leads to the conclusion that they try to satisfy customer requirements at any price instead of offering standard or slightly modified variants to the customer that would cause much less internal effort. However, companies from the field of mechanical engineering face the challenge that the customer intervention (Customer Decoupling Point) takes place earlier in the order fulfillment process than in many other branches. As a result, a significant portion of new features or components is determined directly by the customer. Especially during the ramp-up phase of new
products usually many changes to the product are made. Therefore, the control of the quotation process is necessary for the ramp-up management of products in the field of mechanical engineering.

At an early stage in the order fulfillment process mainly three divisions are involved: sales, product management and development. The overall aim should be to enable sales to offer preferable product variants to the customer, that cause minimum internal effort while it satisfies the customer requirements as well as possible. Today there is no systematic approach for the identification and visualization of these preferable product variants. Therefore, companies often to not control there new product variants especially during the ramp-up phase, which leads to an increasing product variety.

As shown in Figure 1, two main directions of impact can be realized by the use of the methodology of similarity-based product configuration:

1. By offering preferable “standard” variants to the customer, the typical position of product variants can be changed towards the standard area.
2. By analyzing the costs for product variants at an early stage in the quotation process, the costs are visualized and prices can be adapted accordingly.

The part of the methodology presented in this paper mainly addresses the first potential. The paper is organized as follows: After having presented the problem in Section 1, Section 2 gives a short overview of relevant aspects in the fields of preferable product variants and the identification and visualization of similarities. Section 3 deals with related work and the definition of the research gap. In Section 4 the methodology for the identification of preferable product variants is presented. The last section provides the conclusions of this paper.

2. Basic information and definitions

For a better understanding of this paper some definitions need to be clarified in the following section before presenting the methodology.

2.1. Preferable product variants

As described in the introduction, the number of product variants in companies has increased steadily since the early nineties. Many companies try to cope with this problem by structuring their product in modular product platforms, in order to configure product variants from a given set of components [5, 6, 7]. According to this basic idea preferable product variants can be defined, as long as the product structure is relatively modular.

Preferable product variants can be described as basic variants, that can be offered to the customer in a first step. Based on these basic variants product variations can be developed. Preferable product variants have been introduced to the automotive sector in the last years by the building of different packages. In the mechanical engineering sector, packages are not used to the same extent, since the customer’s influence is much higher. Products in mechanical engineering, which are focused in this paper, usually belong to the field of Make-to-Order. This means that the so called Customer Decoupling Point is at a very early stage in the order fulfillment process. Due to the basic idea of the paper (development of product variants based on existing orders) the object region is also extended to the field of Engineer-to-Order. In mechanical engineering the customer has to select his desired product design within a pre-defined solution space, but over and above that a large part of the product needs to be constructed according to the customer requirements.

This paper tries to give a first hint how to identify and visualize preferable product variants in the field of mechanical engineering.

2.2. Identification and visualization of similarities

As mentioned before, this paper aims at visualizing similarities between product variants and identifying preferable product variants. The analysis of a set of individual objects (for example products) concerning the similarity in certain criteria is called similarity analysis. Having identified the similarities, the classification groups these objects into segments, which should be as similar as possible (homogenous) in relation to the criteria and as dissimilar as possible (heterogeneous) among themselves. With respect to product design, “similar” represents a broad range of potential commonality levels in the whole spectrum between “common” (identical in fit, form, and function) and “unique” (different part numbers) [8].

For the calculation of dissimilarities or distances between two objects several distance metrics can be employed. In this paper the most common distance metric, the Euclidian distance, is used. The Euclidian distance corresponds to the geometric distance, which makes it easy to interpret. For the identification of segments different types of cluster analysis can be used: graph-theoretic methods, hierarchical methods (agglomerative and divisive), partitioning methods and optimization methods. Due to the practical relevance this paper uses hierarchical methods with a focus on agglomerative methods. For the visualization between objects different ways of visualization are possible, for example dendrograms or multidimensional scaling. Multidimensional scaling is a family of statistical techniques for analyzing the structure of (dis)similarity data. Multidimensional scaling represents the data as distances among points in a geometric
space of lower dimensionality. This visualization can help to see patterns in the data that are not obvious from the data matrices [9]. In this paper multidimensional scaling with two-dimensional representations is used in order to visualize the distances between several customers and products.

3. Related work

With focus on the object area of the order fulfillment process in mechanical engineering, different approaches in the field of product configuration management are examined concerning different criteria. Derived from the goal of this paper, the consideration of the market level, the customer requirements level, and the similarity analysis of previous product orders build the base for the examination of previous research.

There are several existing approaches for the matching between customer segments and customer requirements [10-15]. However, up to now there is no approach that integrates features and specifications on the two levels focusing on the derivation of preferable product variants. Similarity analysis is used in some approaches in order to identify similar groups of customers [16], customer requirements [17, 18] or products [19-23]. The mentioned approaches make use of different solutions. Some approaches use dendrograms or metrics [19], other approaches focus on cladistics [22].

By looking at the relevant literature it becomes obvious, that some approaches fulfill important aspects of the identification of preferable product variants by the use of similarity analysis, but no approach meets all requirements. Literature review does not provide a holistic approach, that considers both aspects of the identification and visualization of preferable product variants, the customer requirements of a new customer, and the former variants that have been designed and produced in the past. In order to minimize the creation of new, exotic product variants existing product variants in the company need to be considered, which is not focused in the approaches described before.

4. Methodology of the identification of preferable product variants

The following section is structured as follows. First, the methodology concept and framework is presented. Second, the identification of similarities between orders is explained. The third part deals with the visualization of similarities between orders. In the fourth part the determination of preferable product variants is described in detail. The section ends with a case study.

4.1. Methodology concept and framework

During the early stage of the order fulfillment process usually three divisions of a company are involved in the process: sales, product management and development. Therefore, the methodology for the coordination of the three divisions addresses all three view points on the process and combines them in one integrated approach. The overall approach has been described in detail by Schuh et al. [24]. This paper aims at addressing the first two levels in order to present the identification of preferable product variants on the customer requirements level.

The framework of this paper is presented in Figure 2. The level of detail increases from sales to development, since the process starts with a general idea of the customer requirements and ends with a defined bill of materials.

Fig. 2. Methodology framework.

From sales perspective, customers can be described and allocated to a specific customer segment. Product management usually focuses on the customer requirements that constitute a product segment. The view point of the development department is characterized by modules and components that build a product. In the next section the identification of similarities on the first two levels is described in detail.

4.2. Identification of similarities between customers and products

Within the defined framework customers and products can be described using features and specifications. While the proceeding is equal in every use case, the used features can differ from company to company. On the market level customers are defined by general features such as price sensitivity, required quality, need for security or need for innovation. During the description the customer indicates his degree of fulfillment for the different features using the Likert scale. The specification is based on a rating scale, which consists of five steps. Every customer can be described by an n-dimensional vector, where n is the number of features (see Figure 3). In this case customer 1 rates his need for security as relatively low, while customer 2 rates it as high. Since this rating is done for every customer describing feature during an interview at the very early stage of the order process, the distance to other former customers can be identified quickly.

Fig. 3. Description of a customer by features.
The distance between all pairs of orders is calculated using the Euclidian distance metric. In order to represent the importance of the different describing features, a specific weight for each feature is used. A larger value for the weighting factor represents a higher importance of a specific feature for the company. Having calculated the distance between all pairs of customers, these distances can be displayed in a distance matrix. Huge distances between two customers represent a pair of customers, which is very dissimilar, while a small distance represents two similar customers.

\[ d_{ij} = \sqrt{\sum_{k=1}^{l} w_k (n_{k,i} - n_{k,j})^2} \]

- \( d_{ij} \): Euclidian distance between customer profile \( i \) and customer profile \( j \)
- \( n_{k,i} \): Specification for feature \( k \) of customer \( i \)
- \( n_{k,j} \): Specification for feature \( k \) of customer \( j \)
- \( w_k \): Weight of feature \( k \)
- \( l \): Number of features

4.3. Visualization of similarities between customers and products

Figure 5 shows two generic distance matrices on the market and customer requirements level. By using multidimensional scaling every distance matrix can be converted into a two-dimensional diagram that represents the similarities between two objects (i.e. customers and products). Customers or products that are allocated next to each other are very similar. According to the visualization (left part of Figure 5), there are four different customer segments in this example, wherein each of the mentioned customers belongs to a different segment.

In the right part of Figure 5 the customer requirements level is displayed. All orders are classified with multidimensional scaling based on the distance matrix.

![Distances between customer profiles](image)

Fig. 4. Calculation of distances between customer profiles.

On the customer requirement level products can be described by features and specifications addressing more detailed product features. On this level the product features need to be described from the customers point of view. Typical examples from the machine tool industry for features on this level of detail are the size of the machine, number of axes, type of spindle, color, or the type of the coolant system. Specifications for the feature color would be white, blue, or green, for example. Every order can be described by an m-dimensional vector, where \( m \) is the number of all feature specifications that exactly define the product on a customer requirements level. The similarity between orders is calculated based on the comparison between the customer requirements vectors (1 indicates an equal feature specification between two orders, 0 indicates non-equal feature specifications respectively).

The distance between all pairs of orders is calculated using the Euclidian distance metric again. According to the proceeding on the customer level, for each feature a specific weight is used and the distances of the orders pairs are displayed in a distance matrix. The weighting factor represents the importance of different customer requirements for the considered group of customers.

![Distances between customer profiles](image)

Fig. 5. Visualization of similarities between customer profiles and feature profiles by multidimensional scaling.

In this example there are apparently two groups of similar customer requirement vectors with each group consisting of several orders (one small group with product 1 and one large group with products 2, 3 and n). It is obvious that different customer segments do have different customer requirements. The distance between the customers and products 1 and 2 is also shown in Figure 5. Due to different feature weights between the market and the customer requirements level and due to different choices of the customers on the requirements level the distances \( d_{2,1} \) and \( d_{2,1} \) are not equal.

4.4. Determination of preferable product variants

The visualizations on the two levels as presented in section 4.3 build the base for the determination of preferable product variants. This step needs to be divided into two phases. At first preferable product variants need to be defined based on the product portfolio and previous sales experience. In the second phase during the configuration process one preferable
product variant needs to be derived based on orders of similar customers in the past. The definition of preferable product variants can be supported by the multidimensional scaling diagram shown before. Based on the distribution and piece counts of the previous configurations different segments can be derived. In each segment the most probable feature profile is determined (see Figure 6). Since the most probable feature profile is not necessarily a “possible configuration”, the “nearest” possible configuration is defined as the preferable product variant for this segment.

During the quotation process the “optimal configuration” of a new product variant needs to be offered to the customer. In the best case, the preferable product variant works as an anchor for the further negotiations and the customer buys a product variant which is relatively similar to the preferable, standard variant. This second phase consists of two major steps: classification of the customer and derivation of the most probable feature configuration.

At first the customer is classified according to the features defined before. After having build up the customer vector \( v_n \), the distances to all former customers are calculated and the customer is embedded in the multidimensional scaling of the customers. In Figure 7 the new customer is indicated with \( n \).

Based upon this knowledge the most probable feature profile of the customer requirements of former orders is calculated. If the new customer rates the importance of the key customer feature “price” as very high, the probability distribution from all orders with a very high importance for “price” is shown for all customer requirements (see Figure 8). This results in the most probable feature configuration on the customer requirements level (see Figure 7). Since this profile might not be an existing or possible combination of features, one needs to derive the most similar preferable variant. In order to identify this configuration the most probable combination is embedded in the diagram by multidimensional scaling. Having embedded the new combination, the nearest preferable variant can be easily recognized (see Figure 7).

This preferable variant is offered to the customer. If the customer is not satisfied with the offered product variant adjustments need to be done. By adjusting the customer requirements, the product configuration “moves away” from the preferable variant that corresponds with one specific combination of modules and components on the product level. As a consequence, the costs for the realization of this product variant would increase and should be charged to the customer.

4.5. Case study

The methodology has been applied to a manufacturer of roller coasters with more than 300 orders in the last five years. On the market level the customers have been described by four features (price sensitivity, customer base, safety, size). On the customer requirements level twenty features could be identified (e.g. number of seats, seat material, upholstery color, seat heater, safety bar). As a result of the similarity analysis, different segments and preferable product variants could be identified (see Figure 9).
This practical example reveals several advantages of this methodology. First, the variance of different customers and customer requirements is combined in one method which can be used by the three divisions sales, product management and development. Second, the early classification of a customer and the offer of a preferred variant works as an anchor for the following configuration process. This leads to a reduction of internal complexity, since many customers remain with the recommended option due to the lower adjustment costs.

5. Conclusion

In the early phase of quotations there is currently no systematic approach that would result in a product variant that is similar to both, the internal standard variants and the customer requirements. In this paper, a methodology has been presented that allows companies to determine preferable product variants based on similarities between product variants. Similarities are identified with the help of features and specifications covering the market level and the customer requirements level. The third level, the product level, has been omitted in order to focus on the first part of the whole methodology, which has been presented by Schuh et al. [24]. Similarities between product variants are shown on each level by multidimensional scaling, that allows a simple visual evaluation. The definition of preferable product variants and the use during the quotation process by sales, product management and development reduces the number of exotic product variants. In combination with the second part of the methodology, that deals with the correct pricing of these product variants based on dissimilarities, the two main directions of impact presented in the introduction can be pursued. The results presented are aimed both at researchers and practitioners in the industry. With respect to the research community this submission is an important driver for the analytical identification of preferable product variants in a widespread product portfolio. With respect to the practitioners, the main implications can be seen in first application examples within mechanical engineering companies having a high product variety. Further studies in different companies will be conducted in the future. The proposed method has a few limitations especially concerning the evaluation of customer features and customer requirements. Therefore, future research will focus on the quantitative analysis of the benefit of different customer requirements. Furthermore, strategies for the configuration of products during the quotation process need to be derived from the similarities and corresponding diagrams.

Acknowledgements

The author’s sincere thanks go to the companies who supported the development of the approach.

References

[22] AIGeiddawy T, EIMaraghy H. Sustainability and Modularity Analysis of Varying Products and Families Using Cladistics. 3rd International Conference on Changeable, Agile, Reconfigurable and Virtual Production (CARV 2009), Munich, Germany.