DTU Library

Automatic Identification of Similarities Across Products to Improve the Configuration Process in ETO Companies

Shafiee, Sara; Kristjansdottir, Katrin; Hvam, Lars

Link back to DTU Orbit

UDK 658.5:005

Automatic Identification of Similarities Across Products to Improve the Configuration Process in ETO Companies

Sara Shafiee
Industrial PhD Student, Department of Management Engineering, Technical University of Denmark, Denmark,
sashaf@dtu.dk

Katrin Kristjansdottir
PhD Student, Department of Management Engineering, Technical University of Denmark, Denmark, katkr@dtu.dk

Lars Hvam
Professor, Department of Management Engineering, Technical University of Denmark, Denmark, lahv@dtu.dk

Abstract

Engineer-To-Order (ETO) companies making complex products face the challenge of delivering highly customised products with high quality, affordable price and a short delivery time. To respond to these challenges, ETO companies strive to increase the commonality between different projects and to reuse product-related information. Therfore, ETO companies need to retrieve data about previously designed products and identify parts of the design that can be reused to improve the configuration process. This allows companies to reduce complexity in the product portfolio, decrease engineering hours and improve the accuracy of the product specifications. This article proposes a framework to identify and compare products' similarities. The framework (1) identifies the most important product variables available in the Product Configuration System (PCS), (2) retrieves data of previously designed products in an Enterprise Resource Planning (ERP) system, (3) identifies a method to compare products based on the main products variables and (4) sets up an IT system (database) with data of the previously designed products to integrate with the PCS. The proposed approach (the framework and the IT system) is tested in an ETO company to evaluate the application of the framework and the IT system. We retrieved the needed data from the ERP system at the case company and developed the IT system in Microsoft Excel, which is integrated with the PCS.

Key words: *Clustering, Framework, Integration, IT system, Product Configuration System (PCS), Similarities*

1. INTRODUCTION

A product configuration system (PCS) supports users to specify different variables[*] of a product by defining how predefined entities (physical or non-physical) and their properties (fixed or variable) can be combined [1]. PCS offers a good opportunity to enhance a company's resale and production processes starting from the improvement of the quotation process [2]. Several benefits can be gained from utilising a PCS, such as a shorter lead-time for generating quotations, fewer errors, an increased ability to meet customers' requirements with regards to the functionality and quality of the products, and increased customer satisfaction [3–6]. To realise the advantage that can be gained from utilising a PCS, the organisations and the support systems need to change in the order acquisition and fulfilment processes [7,8].

In Engineer-To-Order (ETO) companies producing complex and highly engineered products, a significant problem arises when calculating the prices in the presale and sale processes, especially when domain experts cannot determine accurate price curves, or when vendors fail to provide sufficient information to model within the PCS. Therefore, estimates are often used and mark-up factors are added. Alternatively, ETO companies use prices and other data based on previously made products as a base for the new design. However, this method affects the accuracy of calculations because previous projects are not easily accessible and significant work is required in manual comparison of new products with previous products to find the relevant information [1].

---

[*] A variable is a value that can change, depending on conditions or on information passed to the program.

Hvam et al. [1] presented a solution to the discussed problem, based on a real case. The authors described an ETO company that strives to reuse information from previously made products to calculate the price based on weight and capacity [1]. Price and weight curves are drawn up by inserting the capacity, price and weight based on information from three to five previously produced machines [1]. A curve is then drawn through the points to identify the prices and weights for machines that have not previously been produced, as shown in Figure 1.
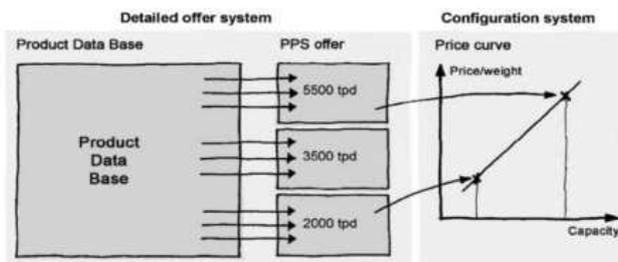
Figure 1. Price and weight curve for the main machines in FLSmidth [1]

However, with regard to highly complex products, the price curves may not be the most accurate because there are several dependent variables and a large number of neighbours on the curve. Another important drawback of the price curves is that the user is only provided by access to some of the previously made projects, and thus the most similar previous projects might be missed.

To identify the similarities of previously designed products and new products, an automated IT system[†] can be beneficial, which makes it possible to produce the customised products while using the least possible amount of time and resources.

Previous research has described how modules across different products [9, 10] can be used to compare different products. Kristianto et al. [11] claimed that platform-based designs can result in economies of scale by mass-producing the same modules and lowering design costs from not having to redesign similar products. Standardisation or system level configuration strategies can be applied in the ETO context [11]. Thus, if an existing product has standardised and decoupled interfaces, the design of the next product can borrow heavily from the modules of the previous product [12].

Thevenot and Simpson discussed a framework that uses commonality indices for redesigning the product families to align with cost reductions in the product development process. They argued that standardising and modularising the product structure incorporated into the PCS can make it easier to select the relevant variables or add them to the PCS [13]. Mäkipää et al. [14] presented the solution of design-configurators for ETO companies. However, they concluded that there are certain limitations of design-configurators, such as

handling calculations and adjusting the design accordingly [14].

Inakoshi et al. [15] proposed a framework to support the PCS, which frames the integration of a constraint satisfaction problem with case-based reasoning (CBR), where the framework is applied to an online PCS. In ETO companies, the integration of existing PCS technologies with recommended approaches is crucial for supporting end-users in their configuration processes [16, 17]. Felfernig et al. [16] discussed different recommendation systems, divided into Collaborative Filtering (CF), content-based filtering (CBF) and knowledge-based recommendations (KBR). The available recommendation technologies in e-commerce are potentially useful in helping *customers* to choose the products' variables. Comparing the new project with previous ones could also result in developing a recommendation system in the companies.

Existing literature do not respond to the need of a structured automatic solution for retrieving the data of previously designed products to reuse in the configuration process.

In this paper, we aim to use a PCS to make a connection between previously designed products and the new products being configured. When generating quotations in the PCS, it is valuable if we can compare the configured products with the previously designed products by comparing the main products variables. This means that, if there is a high percentage of similarity between the new product and a previous made product, the previous documentations and specifications designs can be reused for the new product. Thus, the costs and resources required to generate the product specification can be significantly reduced (i.e. costs in the sales, engineering and production phases).

To achieve this, we develop a framework, as a supporting structure for ETO companies. The framework aims to identify previously designed products that are most similar to the one that a customer is asking for in the configuration process. The framework considers different steps, which guide the company to fulfil this gap. Based on the proposed approach, a framework and an IT system can be generated, where clustering methods are coded to compare the similarities of the products variables.

The remainder of the paper is organised as follows. Section 2 elaborates on the research method. Section 3 details the framework development and discusses each of the proposed steps. Section 4 presents the results of the case study and Section 5 discusses the limitations and presents the conclusions of the research.

2. RESEARCH METHOD

This research was conducted in two phases. The first phase developed the framework for identifying the similarities from previously designed products and new products. The second phase validated the framework.

[†] An IT system is a group of components that interact to produce information [18].

2.1 Phase 1: Framework development for identifying similarities between products

The main purpose of the framework is to define the similarities between previously designed products and new products. To provide a foundation for the proposed framework, we evaluated existing literature focusing on identifying and retrieving the most important product variables, retrieving data of previously designed products and clustering methods to compare products based on the main variables. The literature provides the sequences of steps and methods by which to identify the product similarities. Next, we study the integration of PCS with another IT system in the previous literature [15, 16].

The framework is developed and improved in an iterative testing process, which is described in detail in the subsection 2.2. The next step assesses the framework validation by developing and testing an IT system to automate the process based on the framework.

2.2 Phase 2: Framework assessment through case application

After clarifying the available literature on clustering methods, retrieving the product data and finding the sequences of steps, we developed an IT system to use in a pilot project at a case company that produces highly engineered complex products. The project team formed at the case company included four researchers from the Technical University of Denmark and three experts from the company. The experts from the company included a specialist from the configuration team, a manager and an IT engineer in the IT department.

Based on the proposed framework, we specified the product variables in the PCS and ERP systems at the case company. We identified the product variables from the PCS and managed to collect, treat and structure data from the ERP (SAP) system using MS Excel. We decided to run the pilot project to avoid additional costs by integrating the PCS and ERP and by only coding the clustering constraints in MS Excel. In this step, we selected the clustering methods based on the literature, tested them in the case company and compared the results of the tests.

We prepared the IT system by storing the data from the previously designed products in MS Excel and coding the selected clustering method. However, the success criteria had to indicate what kind of data should be retrieved from previously designed products and how the clustering should be done for the purpose of comparison. Thus, the acceptance criteria for the IT system in the case company were determined as follows:

1. The MS Excel developed IT system should demonstrate its capability to store and retrieve the relevant product variables need to search for similar products.

2. The selected clustering method for comparing the similarities with previously designed products in the configuration process have to be programmed in MS Excel.

3. The IT system (MS Excel) should be integrated into the PCS.

3. FRAMEWORK DEVELOPMENT

Section 1 provided the theoretical bases for developing the framework by covering subjects as: identifying product variables, clustering the data for the comparison purpose, creating an IT system and integrating it with the PCS [9, 10, 12, 14–16, 18–22] The framework aims to fill a gap in the literature, which fails to discuss how the clustering methods can be used to identify similar previously designed products or develop an IT system that can be integrated to the PCS. The proposed framework consists of the following four steps:

1. *Identify the most important product variables available in the PCS*
   The first step of the framework involves defining clear objectives to guide the development and the implementation processes. This includes describing the nature and characteristics of the product and listing the main variables of the products that have to be included.

2. *Retrieve data of previously designed products in the ERP system*
   The second step involves retrieving the data from the identified product variables from the ERP system or any other available database storing the product information.

3. *Identify a method to compare products based on the main variables*
   The third step involves defining a method for clustering the main variables to find the similarities between the products.

4. *Set up the database with data of the previously designed products to integrate with the PCS.*
   The *last step* involves setting up an IT system using the following steps [23]:
   (a) Requirement analysis.
   (b) Conceptual database design.
   (c) Logical database design.

The following subsections explain the individual steps in more detail. In Section 4, the IT system is implemented in the case company and the framework is assessed. Section 4 provides a visual representation and elaboration of the individual steps in the case company.

3.1 Identify the most important product variables available in the PCS

Different techniques can be used to demonstrate, identify and communicate product structure and variables, such as Product Variant Master (PVM) [1] and Product Family Master Plan [24]. A company's product range is often large, with a vast number of variants. To obtain an overall view of the products, the product range is drawn up in a PVM (Figure 2).

In this paper, the PVM is used to break down the components of the product into a tree structure and identify the main product variables. The product structure, variables and rules in the PCS are illustrated using PVM to identify the different variables.
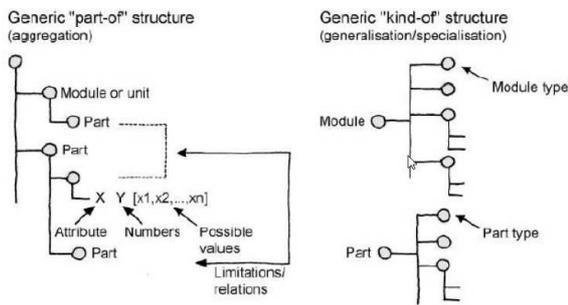

Figure 2. Principles of the PVM [1]

3.2 **Retrieve data of previously designed products in the ERP system**

The current generation of database systems is designed mainly to support business applications, and most of these systems offer discovery variables using tree inducers, neural nets and rule discovery algorithms [25]. One of the fundamental problems of information extraction from ERP systems is that the format of the available data sources are often incompatible, requiring extensive conversion efforts [26]. Knowledge discovery (KD) in databases represents the process of transforming available data into strategic information, which is characterised by issues related to the nature of the data and the desired features [27, 28]. Brachman et al. [29] broke the KD process into three steps:

1. Task discovery, data discovery, data cleansing and data segmentation;

2. Model selection, parameter selection, model specification and model fitting; and

3. Model evaluation, model refinement and output evaluation.

KD includes the derivation of useful information from a database, such as "which products are needed for the specific amount of engineering hours for installation?" [30]. In this article, the specific steps of KD are followed to retrieve the data from the ERP system.

Most companies use the traditional technique called "British classification" when naming different components according to the product variants. However, as products become more complicated, this technique becomes more impractical. When using this technique, as shown in Figure 3, a "surname" of five digits represents the general class of an item and a "Christian name" of three digits provides a particular item with an exact identity [31].
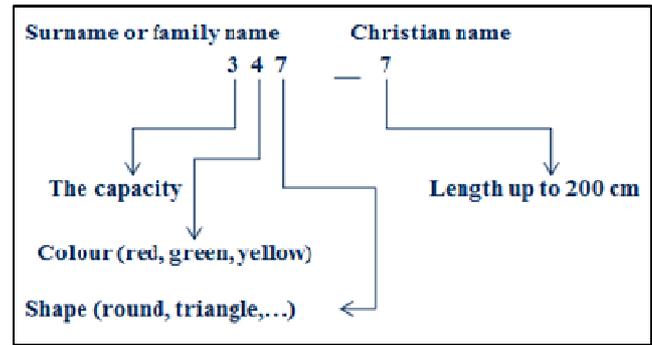

Figure 3. Expansion of a major class [31]

The British classification can be used to assess the products similarities with a high level of abstraction. Thus, we used this technique to decode the high level data from the ERP system.

3.3 **Identify a method to compare products based on the main variables**

Clustering techniques are required for identifying and clustering relevant products variables. Burbidge [31] described how to cluster the product components and code them by introducing the Group Technology (GT) method. Martinez et al. [32] provided an example of using the GT technique in a manufacturing plant to minimise unnecessary diversity by making designers aware of existing components.

The aim of clustering and coding is to provide an efficient method of retrieving information and improving the decision-making. Leukel et al. [33] discussed the design and components of product clustering systems in business to business (B2B) e-commerce and suggested a data model based on XML. Fairchild [34] discussed the application of clustering systems and their requirements. Simpson [35] used GT for adding, removing, or substituting one or more modules to a product platform that should improve the design of the product platform and the customisations. Fairchild et al. [34] suggested an automated clustering system for the specialisation of life cycle assessment. Ho [28] introduced a system, called OSHAM, generated in a hierarchical graphical browser, which competes with C4.5. Software Product Line Engineering (SPLE) was introduced to represent the combinations of features that distinguish the system variants using feature models [36].

A popular non-hierarchical clustering method is the k-means clustering algorithm, which is recognised for its efficiency [37]. This method aims to minimise the k-means algorithm considering the squared differences between the observational data vectors and the cluster centroids overall observations and k-clusters [37]. A method proposed by Anzanello and Fogliatto [38] is based on six steps: (1) Obtain experts' variables, (2) Model the variables, (3) Define bounds, (4), Select the variables, (5) Check whether the upper bound is selected, and (6) Identify the best variables and clusters. Euclidean distances are typically used to calculate the distance between observations because a Silhouette Graph can be generated for displaying the

performance of a clustering procedure [39]. The method provides, for each observation j, the SIj, which can vary from –1 to +1. The closer SIj is to one, the less is the distance within a cluster, meaning that it is properly assigned to the correct cluster [40]. SIj is estimated as follows:

$$SI(j) = \frac{b(j) - a(j)}{\max(a(j), b(j))}.$$

(1)

3.4 Set up the database with data of the previously designed products to integrate with the PCS

Ramakrishnan et al. [23] provided an overview of database design based on the following three steps:

1. Requirement analysis: Understand what data are to be stored in the database, what applications must be built on top of it and what operations are most frequent and subject to performance requirements.
2. Conceptual database design: The information gathered in the requirements analysis step is used to develop a high-level description of the
3. data along with the constraints to be stored in the database.
4. Logical database design: The Database Management System (DBMS) has to be chosen to implement the database design, and the conceptual database design must be converted into a database schema in the data model of the chosen DBMS.

In this paper, we used the database design instruction proposed by Ramakrishnan et al. [23]. First, we performed the requirement analysis, which is discussed and elaborated in step 1 of the proposed framework. Next, the conceptual database design is built based on the analysis from step 1 and the retrieved data in step 2. Finally, the logical design of the database is followed by choosing MS Excel, and the logics are built upon the selected clustering method.

4. CASE STUDY

The proposed framework was tested in an ETO company by developing the IT system. Figure 4 illustrates the process of fulfilling the framework steps to deliver the IT system to the case company over four months. The stakeholders of this pilot project are the sales engineers, sales managers and technical designers from the relevant department. The main potential benefits from using this IT system in the case study were discussed by the stakeholders and listed as the following project aims:

Recommendation system: The decision was made to design the system and its user interface to be replaceable by a recommendation system in the sales process.

Price estimation: It would be beneficial if the IT system could be used to analyse the relationship between costs and variables in the cluster analysis. Thus, the calculated estimated costs from the PCS could be

verified or corrected accurately after configuring the product by comparing them with the previously designed products.

Statistical analysis: It would be preferable if a more detailed overview of the product complexity, the most sold products and the products never sold was provided. This would help the company to reduce the complexity in product ranges based on market requests, clean up the product range and replace it with new product variables based on the knowledge from the market.

4.1 Step 1: Identify the most important product variables available in the PCS

The first step involves selecting the main product variables to be compared across new and previously made products. The PVM is used as the tool to identify the main product variables [1]. The tree structure of the PVS is then used to structure the entire product and to break the main overall product structure down into small enough issues to analyse. Using the PVM, we determined the main product variables of the chosen products.

4.2 Step 2: Retrieve data of previously designed products in ERP system

In the second step, all the main product variables and data were retrieved from the ERP system using KD [29]. The main customised variables were determined as the main variables of the selected products (e.g. weight and cost). Based on these customised product variables, one specific component with different variables was selected, and the IT department helped to retrieve the cost documents from the ERP system into MS Excel. The retrieved data were then divided into subparts (based on the specific variables from the PCS) and the project numbers were decoded to make the deliverables more generic.

4.3 Step 3: Identify a method to compare products based on the main variables

After testing multiple clustering methods, this paper uses k-means and Euclidean distance measurement methods. The first objective in this step was to select the most suitable set of clustering variables leading to an optimised product grouping. Therefore, the k-means procedure was run for every combination of the variables. Each one belonged to a different Excel sheets. In this case, there were four sheets for each cluster: x-y, x-z, y-z and x-y-z. We assess which sheet would lead to the optimal clustering, where the average Silhouette Index (SI) for all the analyses was stored. A higher SI means more accurate clustering. The next step was to calculate the distance between the previously designed and the new product based on the Euclidean distance. This distance was calculated for all combinations of the variables—three variables (x, y, z) and six possibilities (xyz, xy, xz, yz, x, y, z). A small distance between the new product and the previously designed product indicated a high similarity. The formula shown in Figure 5 is based on a Euclidean distance measurement. The final step of the

comparison platform is to list the products based on similarity. This was done by ranking the distance measurements. As shown in Figure 6, the distance rank a6 has the shortest distance to the new product, a4 has the second closest product and a7 is the third closest to the new product among the previously made products. The cluster was initially placed and based on the k-means algorithm, and a final position for the cluster's centroids was found. The algorithm continued with the second iteration, where the same procedure was

applied. As a result of the further iteration, the cluster centres moved according to their belongings, which resulted in an increase in the average SI. A higher SI means a more accurate clustering. The algorithm continued until the cluster centres stopped moving. Figure 7 illustrates the situation resulting from several iterations.
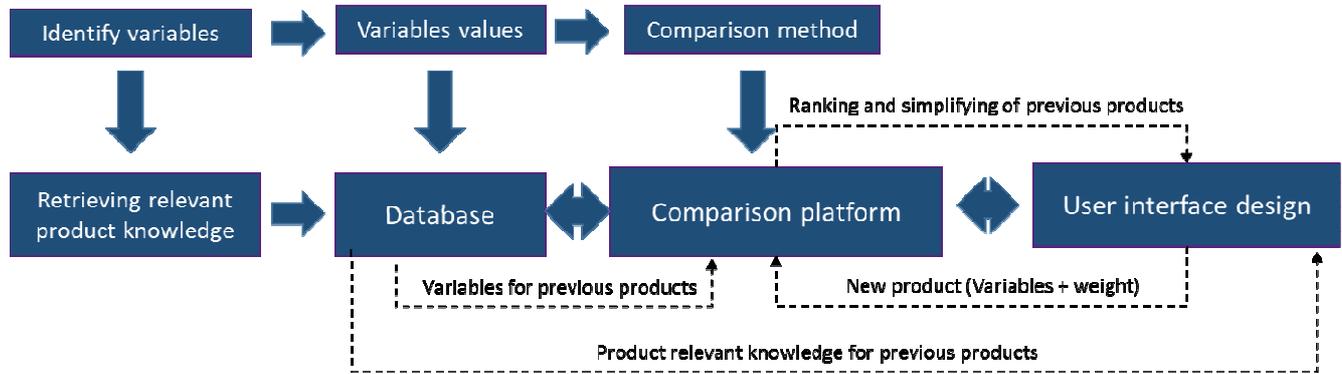

Figure 4. Structure and information flow of the IT system in the case company

Figure 5. Distance matrix

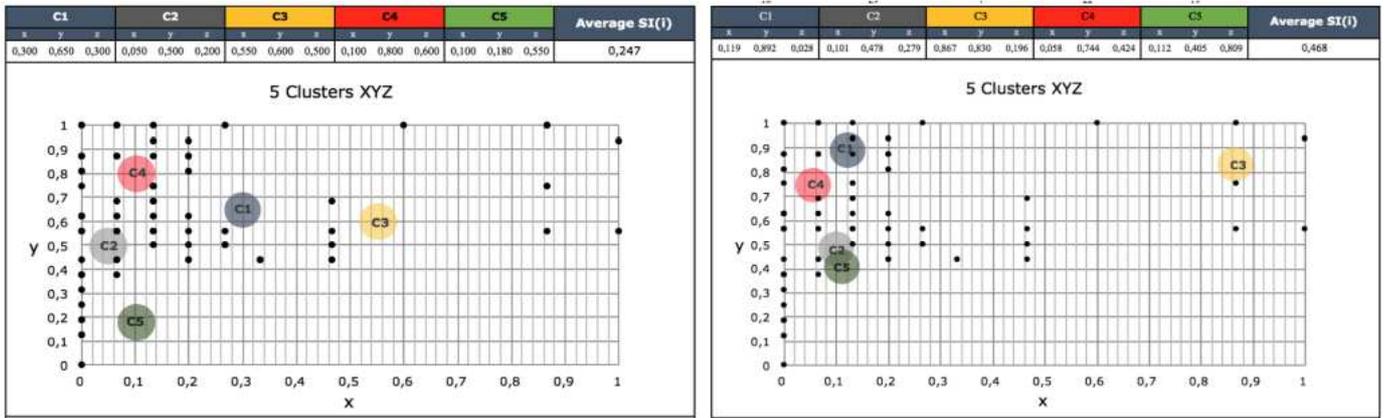**Figure 6.** Ranking the distances and projects

Figure 7. Movement of the clusters and average SI, at the beginning (left) and at the end of the k-means algorithm (right)

4.4 **Step 4: Set up the database with data of the previously designed products to integrate with the PCS**

The PCS at the case company is based on a commercial platform, where the integration with MS Excel forms part of the standard system. The aim of the user interface is to return similar previously made products when the user configures a new project. Based on this, the user can use product-relevant information from previous projects.

The IT system, which was developed based on the proposed framework, was tested in the case company with one of the current PCS. Figure 8 shows the simple user interface after the Excel sheet is generated from the PCS, where the main product variables are exported to MS Excel. Furthermore, MS Excel is integrated with and receives the relevant input from the PCS.The inputs were received from the PCS and added to the MS Excel. However, there is an input area in the Excel spreadsheet in case the PCS is not used.

The input part is covered by the three upper-left boxes in the user interface, which can be seen in Figure 9. The white fields are where the user can enter inputs. Therefore, the use of the Excel sheet is not only limited to the PCS.

Users can exclude products variables if they are not relevant. If a variable is taken out of the interface, it will be taken out of the distance calculation and other products will be recommended. The "elimination feature" is also integrated into the PCS. Figure 10 shows how the user can eliminate variables by clicking "YES" or "NO" and indicates how this impacts the output.

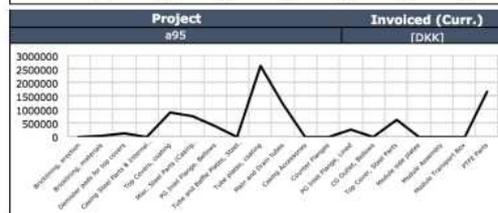To visualise the output, a bar graph was added to the user interface. Data for the graph are based on the relevant product information chosen as first priority in the input field. Thus, it is possible for the user to change the data subsequently. In addition, the graph was programmed so that it would fit the number of recommended products (Figure 11).

Figure 8. Final user interface of the IT system

Product Configuration

	Variables		
	x	y	z
Choose if relevant:	Yes	No	Yes
Input:	4	20	64
Value options:	(1,2..80)	(1...20)	(60..72)
Variable weight:	150	70	20
	100,0%	46,7%	13,3%

	Recommendation conditions		
	# projects	Max distance	Maximum age
Input:	7	1	10
Variable value options	(1,2..120)	(0..1)	(0,1..50)

	Relevant product information		
	1st	2nd	3rd
			sportatio cost
	Expected Cost		
	Real Cost		
	Lead time		
	Quality issue		
	Drawings		
	Working hours		

Figure 9. Choice of product-relevant information

Product Configuration

	Variables		
	x	y	z
Choose if relevant:	Yes	No	Yes
Input:	4		64
Value options:	(1,2..80)	Yes / No	(60..72)
Variable weight:	150		20
	75,0%	100,0%	10,0%

Recommended Projects

	Variables xz	Distance	x	y	z	Age [years]	Price Estimation	Drawings	Transportation cost [DKK]
1	a66	99,2%	4	22	65	14	1350	D70	110
2	a35	99,2%	4	28	65	14	1313	D36	104
3	a18	99,2%	4	28	65	8	1249	D19	254
4	a23	98,8%	4	23	62,5	2	1204	D24	246
5	a11	98,8%	4	29	62,5	12	1378	D12	104
6	a48	98,7%	3	28	65	3	1353	D52	249
7	a26	98,7%	3	26	65	6	1242	D27	165

Product Configuration

	Variables		
	x	y	z
Choose if relevant:	Yes	No	No
Input:	4	20	
Value options:	(1,2..80)	(1...20)	Yes / No
Variable weight:	150	200	10,0%
	75,0%	100,0%	10,0%

Recommended Projects

	Variables x	Distance	x	y	z	Age [years]	Price Estimation	Drawings	Transportation cost [DKK]
1	a71	100,0%	4	27	60	8	1268	D75	245
2	a66	100,0%	4	22	65	14	1350	D70	110
3	a46	100,0%	4	21	68,5	9	1357	D48	214
4	a35	100,0%	4	28	65	14	1313	D36	104
5	a32	100,0%	4	21	72	1	1393	D33	252
6	a25	100,0%	4	24	62	10	1395	D26	197
7	a23	100,0%	4	23	62,5	2	1204	D24	246

Figure 10. Exclusion of variables

Information	[Unit]
PO Price (Curr.)	[DKK]

Information	[Unit]
PO Price (Curr.)	[DKK]

Figure 11. Graphs for recommended products

5. DISCUSSION AND CONCLUSION

As products become increasingly complex, it becomes more difficult to generate precise product specifications from the PCS, especially for complex products. Integration of existing configuration technologies with recommended approaches is crucial to support end users in the configuration processes [16, 17]. Researchers have proposed various support measures to help to integrate PCSs with other IT systems [15], and existing literature provides examples o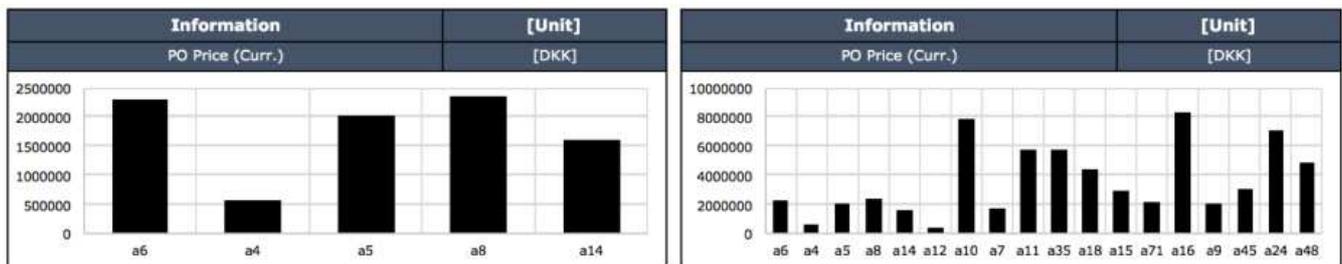f clustering methods [37–40]. However, there is no automatic solution for retrieving and reusing product information in the configuration process. This solution proposed in this paper thus builds on the available literature on clustering and integration. Based on the literature and experiences working with PCSs, the users of PCSs check and compare some of the old projects they are capable of remembering before sending out new order proposals. In this way, they might be able to find similar products and thus reduce the necessary time and resources, improve the quality, increase the accuracy of their calculations and eliminate the engineering processes or even offer customers the same product at a lower price. However, even when there are similar

products, it can be difficult to find them in the ERP system, and this process of finding similar products can become even more challenging once the proposal phase is accepted and the engineering phase has started. The engineers sometimes waste time repeating the same processes without realising that another similar project was completed earlier, and they could simply reuse the data.

In this paper, we propose a framework for creating an IT system of previously completed products and compare against new projects. This approach allows efficient comparisons to be made while using the available methods and tools. An IT system was coded in a separate MS Excel file as the pilot project using the minimum resources at the case company. The IT system showed the ability to cluster and compare the product data and thus proved the feasibility of the concept. Moreover, we tested the proposed approach in a case ETO company to determine whether the framework and IT solution are practical in a real-life situation.

As discussed in Section 2, we need to determine some criteria at the case company for accomplishing the project. The criteria and deliverables fulfilled during the case company project are as follows: (1) We retrieved and stored the relevant product variables for the product in MS Excel, (2) we coded the selected clustering method for comparing the similarities from previously designed products in the configuration process in MS Excel, and (3) we integrated the Excel database into the PCS used at the case company.

The users of the system at the case company saved time and resources by using this IT system. Previously, they faced a number of problems estimating costs and engineering and workshop hours, which led them to check the previous projects manually.

The IT system, that was developed based on the proposed framework in this paper, helped the users of the PCS to manage the high number of previously designed products and the high level of customisation. The users of IT system did not have to overcome any challenges related to training or system changes because the engineers were familiar with the setup of Excel and it had a friendly user interface. They also mentioned that this clustering method and IT system not only saved around 50% of their time when making sales quotations but also reduced errors and increased the accuracy of their proposals. This paper is limited to a single-case study containing limited data. Limited numbers of clustering methods were tested. The coded IT system might not be efficient when the number of variables increases. This IT system needs to be continually maintained because it has to be aligned with the ERP system; otherwise, it will become outdated and forgotten after a number of projects have been sold. Therefore, in the future it might be more beneficial to integrate the PCS directly to the ERP system. As mentioned, the framework and IT system are eveloped in an iterative process in an ETO case company. However, the case study type allows the research group to face the complicated types of products and repeat the in-depth testing of the developed framework

and IT system. Meanwhile, the study of one case company allowed the team to have hands-on practice and make IT developments to assess the research in a real situation long-term.

Further research should be conducted to enable generalisability of this approach and to test the proposed approach in more and different case companies with different products. Future research can focus on clustering and integrating the IT systems with the ERP system to update the knowledge automatically. The goal is to use the ERP as the main database and automatically retrieve the stored and updated data from the ERP system.

6. REFERENCES

[1] Hvam, L., Mortensen, N. H. and Riis, J. (2008), *Product Customization*, Berlin Heidelberg, Springer.

[2] Felfernig, A., Bagley, C., Tiihonen, J., Wortley, L. and Hotz, L. (2014), "*Benefits of Configuration Systems*", Felfernig, A., Hotz, L., Bagley, C. and Tiihonen, J. (Eds.). *Knowledge-Based Configuration* from *Research to Business Cases*. Knowledge-Based Configuration. Morgan Kaufman.

[3] Haug, A., Hvam, L. and Mortensen, N. H. (2011), "*The Impact of Product Configurators on lead times in Engineering-Oriented Companies*", Artificial Intelligence for Engineering Design, Analysis and Manufacturing, Vol. 25, No. 2, pp. 197–206.

[4] Hvam, L., Haug, A., Mortensen, N. H. and Thuesen, C. (2013), "*Observed Benefits from Product Configuration Systems*", International Journal of Industrial Engineering: Theory, Applications and Practice, Vol. 20, No. 5-6, pp. 1-6.

[5] Trentin, A., Perin, E. and Forza, C. (2012). "*Product Configurator Impact on Product Quality*", International Journal of Production Economics, Vol. 135, No. 2, pp. 850–859.

[6] Zhang, L. L. (2014), "*Product Configuration: A Review of the State-of-the-Art and Future Research*", International Journal of Production Research, Vol. 52, No. 21, pp. 6381–6398.

[7] McGuinness, D. L. and Wright, J. R. (1998), "*Conceptual Modelling for Configuration: A Description Logic-Based Approach*", Artificial Intelligence for Engineering Design, Analysis and Manufacturing, Vol. 12, No. 4, pp. 333–344.

[8] Shafiee, S., Hvam, L., Haug, A., Dam, M. and Kristjansdottir, K. (2017), "*The Documentation of Product Configuration Systems: A Framework and an IT Solution*", Advanced Engineering Informatics, Vol. 32, pp. 163–175.

[9] Zamirowski, E. and Otto, K. (1999), "*Identifying product portfolio architecture modularity using function and variety heuristics*" in proceedings of DETC99, pp. 1–12.

[10] Ericsson, A. and Erixon, G. (1999), *Controlling Design Variants: Modular Product Platforms*, Society of Manufacturing Engineers, Michigan.

[11] Kristianto, Y., Helo, P. and Jiao, R. J. (2015), "*A System Level Product Configurator for Engineer-to-Order Supply Chains*", Computers in Industry, Vol. 72, pp. 82–91.

[12] Ulrich, K. (1994), "*Fundamentals of Product Modularity*", Author, A. (Ed.), Management of Design, Springer, pp. 219–231.

[13] Thevenot, H. J. and Simpson, T. W. (2006), "*Commonality Indices for Product Family Design: A Detailed Comparison*", Journal of Engineering Design, Vol. 17, No. 2, pp. 99–119.

[14] Mäkipää, M., Paunu, P. and Ingalsuo, T. (2012), "*Utilization of Design Configurators in Order Engineering*", International Journal of Industrial Engineering and Management, Vol. 3, No. 4, pp. 223–231.

[15] Inakoshi, H., Okamoto, S., Ohta, Y. and Yugami, N. (2001), "*Effective Decision Support for Product Configuration by Using CBR*" in proceedings of the fourth international conference on Case-Based Reasoning (ICCBR), p. 8.

[16] Felfernig, A., Hotz, L., Tiihonen, J. and Bagley, C. (2014), "*Configuration-Related Topics*", Felfernig, A., Hotz, L., Bagley, C. and Tiihonen, J. (Eds.). *Knowledge-Based Configuration* from *Research to Business Cases*. Knowledge-Based Configuration. Morgan Kaufman.

[17] Felfernig, A. and Burke, R. (2008), "*Constraint-based recommender systems: technologies and research issues*" in proceedings of the 10th international conference on Electronic Commerce ICEC '08, pp. 1–10.

[18] Forza, C. and Salvador, F. (2007), *Product Information Management for Mass Customization: Connecting Customer, Front-Office and Back-Office for Fast and Efficient Customization,* Palgrave Macmillan, New York.

[19] Goel, A. K. and Craw, S. (2005), "*Design, Innovation and Case-Based Reasoning*", Knowledge Engineering Review, Vol. 20, No. 3, pp. 271–276.

[20] Ruet, M. and Geneste, L. (2002), "*Search and Adaptation in a Fuzzy Object Oriented Case Base*" in proceedings of the 6th European conference on CBR (ECCBR'02), pp. 723–730.

[21] Dahmus, J. B., Gonzalez-Zugasti, J. P. and Otto, K. N. (2001), "*Modular Product Architecture*", Design Studies, Vol. 22, No. 5, pp. 409–424.

[22] Holmqvist, T. K. P. and Persson, M. L. (2003), "*Analysis and Improvement of Product Modularization Methods: Their Ability to Deal with Complex Products*", Systems Engineering, Vol. 6, No. 3, pp. 195–209.

[23] Lafayette, W. (1995), "*Database Management Systems*", Computing, Vol. 27, No. 4.

[24] Mortensen, N. H. (2000), "*Design Modelling in a Designer's Workbench*", Ph.D. dissertation, Technical University of Denmark, Lyungby, Denmark.

[25] Imielinski, T. and Mannila, H. (1996), "*A Database Perspective on Knowledge Discovery*", Communications of the ACM, Vol. 39, No. 11, pp. 58–64.

[26] Bendoly, E. (2003), "*Theory and Support for Process Frameworks of Knowledge Discovery and Data Mining from ERP Systems*", Information & Management, Vol. 40, No. 7, pp. 639–647.

[27] Davies, R. (1989), "*The Creation of New Knowledge by Information Retrieval and Classification*", Journal of Documentation, Vol. 45, No. 4, pp. 273–302.

[28] Ho, T. B. (1997), "*Discovering and Using Knowledge from Unsupervised Data*", Decision Support Systems, Vol. 21, No. 1, pp. 29–42.

[29] Brachman, R. and Anand, T. (1996), "*The process of knowledge discovery in databases*" in proceedings of advances in knowledge discovery and data mining, American Association for Artificial Intelligence.

[30] Owrang O., M. M. and Grupe, F. H. (1996). "*Using Domain Knowledge to Guide Database Knowledge Discovery*", Expert Systems with Applications, Vol. 10, No. 2, pp. 173–180.

[31] Burbidge, J. L. (1975), *The Introduction of Group Technology*, Halsted Press, Heinemann, London.

[32] Martinez, M. T., Favrel, J. and Ghodous, P. (2000), "*Product Family Manufacturing Plan Generation and Classification*", Concurrent Engineering Research and Applications, Vol. 8, No. 1, pp. 12–23.

[33] Leukel, J., Schmitz, V. and Dorloff, F.-D. (2002), "*A modeling approach for product classification systems*" in proceedings of the 13th international workshop on Database and Expert Systems Applications, pp. 868–874.

[34] Fairchild, A. M. and De Vuyst, B. (2002), "*Coding standards benefiting product and service information in E-Commerce*" in proceedings of the annual Hawaii international conference on System Sciences, pp. 3201–3208.

[35] Simpson, T. W. (2005), "*Product Platform Design and Customization: Status and Promise*", Ai Edam, Vol. 18, No. 1, pp. 3–20.

[36] Lopez-Herrejon, R. E., Linsbauer, L., Galindo, J. A., Parejo, J. A., Benavides, D., Segura, S. and Egyed, A. (2015), "*An Assessment of Search-Based Techniques for Reverse Engineering Feature Models*", Journal of Systems and Software, Vol. 103, pp. 353–369.

[37] Taboada, H. A. and Coit, D. W. (2008), "*Multi-Objective Scheduling Problems: Determination of Pruned Pareto Sets*", IIE Transactions, Vol. 40, No. 5, pp. 552–564.

[38] Anzanello, M. J. and Fogliatto, F. S. (2011), "*Selecting the Best Clustering Variables for Grouping Mass-Customized Products Involving Workers Learning*", International Journal of Production Economics, Vol. 130, No. 2, pp. 268–276.

[39] Rand, W. M. (1971), "*Objective Criteria for the Evaluation of Clustering Methods*", Journal of the American Statistical Association, Vol. 66, No. 336, pp. 846–850.

[40] Kaufman, L. and Rousseeuw, P. J. (2009), *Finding Groups in Data: an Introduction to Cluster Analysis* (Vol. 344), John Wiley & Sons.

Unapređenje procesa konfiguracije kroz automatsku identifikaciju sličnosti proizvoda u kompanijama koje projektuju po narudžbini

Sara Shafiee, Katrin Kristjansdottir, Lars Hvam

Apstrakt

Kompanije koje projektuju i proizvode složene proizvode po narudžbini (engineering-to-order) se susreću sa izazovom isporuke kastomizovanih poizvoda visokog kvaliteta, u kratkom roku i po prihvatljivoj ceni. Kako bi odgovorile na ove izazove, kompanije koje projektuju po narudžbini teže da povećaju sličnost između različitih projekata i na taj način u više slučajeva iskoriste informacije vezane za proizvode. U tom smislu, kompanije koje projektuju po narudžbini treba da prikupe podatke o prethodno projektovanim proizvodima i identifikuju one delove koji su mogu ponovo koristiti kako bi se unapredio proces konfiguracije. To omogućava kompanijama da smanje složenost portfolia proizvoda i vremena za inženjering, kao i da unaprede preciznost specifikacije proizvoda. Ovaj rad predlaže okvir za identifikovanje i poređenje sličnosti proizvoda. Ovaj okvir (1) identifikuje najvažnije elemente proizvoda dostupne u sistemu konfiguracije proizvoda (SKP), (2) prikuplja podatke o prethodno projektovanim proizvodima u ERP sistemu, (3) identifikuje metod da uporedi proizvode na osnovu najvažnijih elemenata proizvoda i (4) postavlja IT sistem (bazu podataka) sa podacima o prethodno projektovanim proizvodima kako bi ih integrisao sa SKP-om. Predloženi pristup (okvir i IT sistem) je testiran u kompaniji koja projektuje po porudžbini kako bi se ocenila primena okvira i IT sistema. Prikupljeni su neophodni podaci iz ERP sistema kompanije i razvijen je IT sistem u excel-u, koji je integrisan sa SKP-om.

Ključne reči: *grupisanje, okvir, integracija, IT sistem, sistem konfiguracije proizvoda (SKP), sličnosti*