20th ISPE International Conference on Concurrent Engineering
C. Bil et al. (Eds.)

doi:10.3233/978-1-61499-302-5-421

Prospective Evaluation of Assembly Work Content and Costs in Series Production

Ralf Kretschmer [a], Stefan Rulhoff [b] and Josip Stjepandić [b,1]
[a] *Miele & Cie. KG, Germany*
[b] *PROSTEP AG, Germany*

Abstract. Strategic decisions in early production planning phases have a high impact on various production aspects. Decision making is often based on vague expert knowledge due to lack of a reliable knowledge base. Implications of this problem are especially observable in the field of assembly planning, which integrates results from various planning disciplines. This paper introduces a new concept and the corresponding data model for application of Data Mining (DM) methods in the field of production assembly planning and product design. The approach contains the usage of existing planning data in order to extrapolate assembly processes. Especially linked product and process data allow the innovative usage of Data Mining methods. The concept presents assistance potentials for development of new products variants along the product emergence process (PEP). With this approach an early cost estimation of assembly processes in series production can be achieved using innovative Data Mining methods. Furthermore, design and planning processes can be supported effectively.

Keywords. Product Realization, Manufacturing, Digital Factory, Assembly, Process Planning, Data Mining

Introduction

Today globally operating companies face additional challenges due to the increasing variability of products and complexity of processes. Therefore there are growing demands on the flexibility in the production system on the economic dispatch of new products in an existing production line. In the modern product emergence process production planning gets increasingly important and has to be executed in parallel to the product development [1]. In this early phase of the product creation a first start for planning processes is a cost calculation for the industrialization of the product in existing production lines regarding basic conditions [2]. The economic feasibility of series production must be assured with vague information on the product and given general conditions, e.g. shift model [3]. This is a great challenge especially to the planning of the cost-intensive assembly of the product [4] [5].

In order to meet this challenge PROSTEP AG supports Miele Cie & KG, one of the leading manufacturers of domestic appliances, developing innovative methods in the research project ProMondi. Aim of this project is the accurate estimation of the expected assembly work content and the resulting costs in an early stage of the product development. The approach contains the usage of existing planning data in order to

[1] Corresponding Author, Mail: Josip.stjepandic@opendesc.com.

extrapolate assembly processes. Especially linked product and process data allow the innovative usage of data mining methods. New processes appropriate to assemble the given new product shall be designed based on this existing linked product and process data. Automatic analysis with a specific data mining model can be used to create a first draft of the assembly process and estimate the expected costs. Additional use cases can be addressed. Following production planning processes can be supported by automatic proposals of adequate assembly processes, which then can be customized. Moreover the design engineer can be supported at the selection of appropriate joining elements. With this approach an assembly knowledge based support of the designer in series production can be achieved using innovative data mining methods.

1. Use Case Miele

In order to address the challenges of data mining the integration of various planning tasks within the PEP, new concepts are necessary. Though, as a part of integrated product and process development there are different definitions for various phases and aspects of planning activities along the PEP. Regardless of the specific definition of these phases and aspects, however, based on the analysis it is certain that great amount of their containing information and knowledge are either utilized insufficiently and ineffectively or remain unused [6]. In this regard, the presented concept focuses on product design and production assembly planning. Subsequently, for the product designer and production planer, there are varieties of applications, which can assist the design or the planning process through information gathered by data mining.

Assembly process estimation: The focus is on the creation of an assembly process for a new product. Based on existing product and process data compilation of a first approximated assembly process for a new product could be developed. From this, the production planner can specify further details and thus determine a first estimation regarding assembly time. Based on the assembly time and associated calculation scheme, the planner can perform the first cost estimation in a very early planning phase.

1.1. Preparation and Requirements

The information in production planning and engineering processes can mutually enrich each other. Additionally intelligent interconnecting information from both areas creates added value. The newly obtained information supports the workflow throughout the PEP. Therefore, as part of this concept some requirements need to be met. Thus the pre-conditions attached to both systems as well as their respective processes have to be fulfilled [7].

1.1.1. Attributes and Data Sources

Data mining is a process of discovering valuable information from observational data sets, which is an interdisciplinary field bringing together techniques from databases, machine learning, optimization theory, statistics, pattern recognition, and visualization. Data mining has been widely used in various areas such as business, medicine, science, and engineering. Many books have been published to introduce data-mining concepts, implementation procedures and application cases [8] [9]. The overall goal of the data

mining process is to extract information from a data set and transform it into an understandable structure for further use.

Data mining methods can be used for data clustering and classification, however criteria for comparison of data sets have to be identified. To determine these criteria, within the scope of ProMondi project, a survey of users as well as an analysis of various tools of the DM was performed. The objective of this analysis was to identify attributes relevant for assembly processes that could be assigned to products and parts in CAD [10], PDM and production planning systems. In CAD systems attributes assigned to parts contain mainly geometric information including volume and weight. The PDM systems contain organizational information, such as creator, version and revision as well as the mentioned parts information form CAD [11]. In addition to the conventional systems for design and stacking product parts and assemblies, systems for process planning and time measurement were also taken into account. They sustain a comprehensive portfolio of information and therefore can be used to distinguish different product parts and assemblies. The results of this analyzing are capsulated as an object oriented data model, further described in section 3.1.

1.1.2. Data Collection and Availability

The necessary enrichment of product and process data on the fly for the presented concept requires additional efforts in the design. This additional expenditure also relates to the assembly connections and includes the acquisition of new information form the designer's know how. The designer usually defines assembly connections either implicitly through formed locked joints by the shaping of the parts or explicitly by connecting elements such as in screwed fasteners.

The designer of the assembly connection considers all these information in the design but cannot store them in the CAD model because the CAD tools for the most part are not able to define the necessary attributes.

To overcome this problem as part of the concept presented in this paper, the designer is provided with an additional tool in the CAD system. It can be used to create assembly connections and gives additional information and explicit design possibilities. These additional assembly informations are named below as "product assembly information". Thus, data will be collected in the source system, the CAD system in particular. Since the defined objects are not part of PDM systems an extension is necessary in order to implement connections as objects and to store them after the transfer in the PDM system persistently. In the further processing, the product information is linked to the planning processes. Unless the storage of product data are in the same system as for production planning, the information flow from the PDM system to the planning system as well as the Data Mining tool, for further analysis, has to be ensured.

1.1.3. Aggregation of the existing Data

In current planning systems often direct linking of processes to products is possible [11]. Thus an allocation of to be assembled product and the associated assembly processes is realized. In the assembly, however, parts are joined with other parts or product. These assembly connections have no digital equivalent object yet. However by means of an object such as the product assembly informations it is possible to store useful additional connection information, which relates directly to the respective assembly connection. As part of this concept, the combination of the products and

processes does not take place directly but through the product assembly information. The linking of product and process does not necessarily need to occur at the part level.

2. Solution Concept

The concept presented in this paper describes an assisting workflow to support the designer (Figure 1). As part of a new or modified design the designer creates new product data. In creating the assembly connections a software assistant supports the designer and enriches the CAD model with product assembly information for each connection. This product assembly information includes additional connection information including e.g. the torque screwed fasteners or the type and the form of a welded joint and information about other connection types. In the ongoing design process the designer can trigger an evaluation regarding the assembly connections in the model.

For this purpose, the characteristics of the CAD model are first prepared and analyzed with Data Mining. The analysis focuses on the product assembly information and their properties. The parts associated with the product assembly information and their geometric properties, are also included in the analysis as additional information set. Furthermore, an extended database is also provided and consists of product and process data of existing products, which are linked via the product assembly information. The characteristics of the product assembly information of the new product are compared with the properties of the product assembly information of the existing product in the extended database. Then the most similar product assembly information is determined from the existing products. This analysis can be restricted by a class of the connection types (screwed, weld, rivet) or deliberately left open to widen the solution space and to provide the designer with information about other assembly connections.

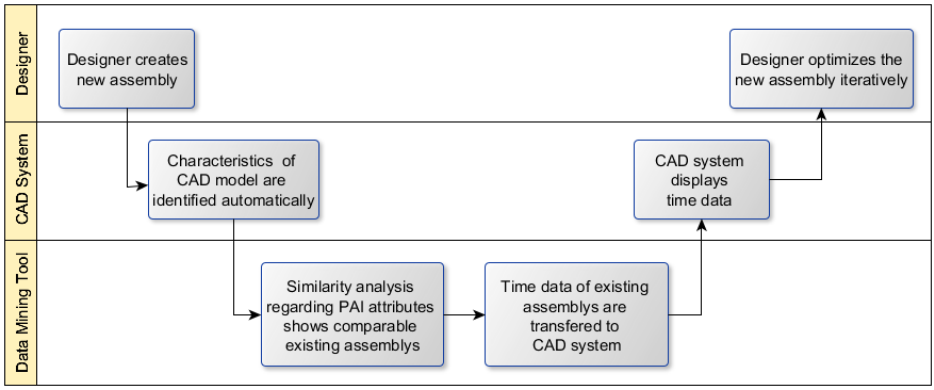

Figure 1. Design optimization with additional time data.

A limitation on the particular type of connection yields as a result of the closest realized assembly connection of the same kind. Depending on the properties of the parts, other mounting connections can also be found and proposed to the designer in a proposal list.

The presented application for the support of the design process uses the product assembly information identified in the analysis of the PDM database to determine the

respective associated and related sub processes. Each product assembly information represents an assembly connection. By multiple connections within the assemblies multiple sub-processes for the assembly can be determined. These processes contain the time data relevant for the new product design. Therefore the corresponding time information of the existing products and if requested an alternative proposal list is transferred in the CAD System and displayed. This assembly time information of the existing product represents a first approximated assembly time for the new product. So the designer is provided with this additional information regarding the assembly time and with an enterprise specific factor the corresponding cost of the current design solution. In the final step the designer is able to optimize the product iteratively on the base of anticipated assembly time and costs for each design.

2.1. Data model

Based on determined assembly characteristics a range of attributes is derived to classify the assembly of the parts. Figure 2 shows an overview of the generated data model for the data mining analysis.

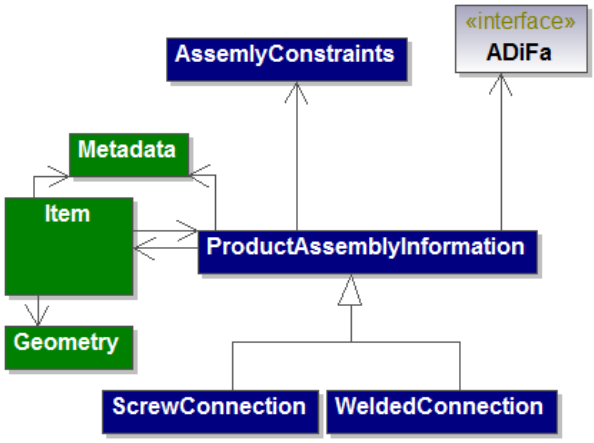

Figure 2. Data model overview.

The ProductAssemblyInformation (PAI) is the central element in this data scheme and represents the assembly of the product parts. References for time analysis, assembly requirements, designed parts or products as well as a wide range of meta data including the assembly department and other information are lodged. This element is supplemented with attributes of different connection types (see Figure 3). Further connection types can be added to the data model. To provide the required information for the time analysis a standardized data model is applied. In this regard, ADiFa project's "Application-specific data models", so called ADiFa Application Protocols, were used, which offer the integration of processes and data for different DM systems [12].

The second fundamental object in the data model is the Item. It contains references to existing sub-assembly units, geometrical characteristics as well as ProductAssemblyInformation. Each Item refers to the ProductAssemblyInformation, which also refers to further used Items. This construct is chosen to enable Data Mining

methods to determine exact similarities between new parts and/or products and other existing parts. Furthermore, it makes comparison parts and products the new and existing ones, in any order and combination interchangeably possible.

In the first approach the attributes for screw connections are clustered and evaluated regarding the relevance for assembly operation. Figure 3 shows the identified attributes classified in the categories fasteners, installation / assembly situation, tools / equipment, installation regulations and additional assembly elements. These attributes are represented in the data model in different object classes here illustrated by color. The evaluation regarding the influence on the assembly time provides a first indication for the relevance in the data mining analysis. Which attributes really are significant for the similarity of assembly connection have to be determined in a data mining analysis with a large quantity of product data.

Category	Product Assembly Information (PAI)	influence on the assembly time
fasteners	screw head diameter	low
	thread type	medium
	number of thread transitions (used)	low
	screw diameter	low
	screw length (thread)	medium
	screw type	medium
	material	low
	magnetic screw	high
installation/assembly situation	chamfer on screw	high
	output of the screw	low
	additional elements	high
	lack of space	high
	visual disability	high
	risk of injury	low
	additional fix the add part	medium
	working both hands feasible	medium
	screwing in	medium
tools, equipment	threaded sleeves used	high
	glove used	low
	equipment used	low
	tool	high
installation regulations	additionally tighten	low
	tightening torque	low
	check torque	low
	assembly sequence	medium
	multi-stage screwdriving	medium
additional elements	flat washer	medium
	nut	medium
	LocTite	low
	grease	low

Figure 3. PAI attributes example screwing connection.

2.2. Data mapping and data mining

After aggregating and appending the data subsets from different sources and systems, it is necessary to remove redundant data sets [13]. Data removal for the presented concept is only based on syntactic similarities of attribute structures and data

sets. The next step is converting and porting data in the presented data model. Depending on data source the conversion is either fully automated or partially automated with further manual adjustment. Often value and scale of different attributes are often heterogeneous (Figure 4). In these cases a normalization of ratings prevents the undesired high or low impact of certain attributes on the results and evaluation process. In this regard a [0, 1] linear normalization has been used. Additionally, a further attribute prioritizing via weighting can be necessary to define the importance of each attribute for the evaluation. An automated learning of the weights via machine learning methods depends on the existing data sets and their quality possible. Otherwise they are determined based on expert knowledge or a combination of both methods. To prevent further expansion of scope and the complexity of existing problem expert knowledge was applied to determine the attribute weights. Furthermore, it is possible to have more than a single weight vector. This approach is useful, if there are various object types or parts, which have different prioritization for their attributes [14]. To identify the objects with most similar product assembly information for a new object the classification algorithm k-nearest neighbor (kNN) [15] with Euclidean distance as evaluation function is used. From the identified objects a list is generated and the most related one can be manually chosen, which passes its assembly process data to new object. To assure the reliability of the presented method and prevent overfitting problem a cross validation [16] is used.

Category	Product Assembly Information (PAI)	values	influence on the assembly time	construction		work planning	
				system	knowHow	system	knowHow
fasteners	screw head diameter	mm	low	x		x	
	thread type	fine thread / coarse thread	medium	x		x	
	number of thread transitions (used)	yes/no	low	x		x	
	screw diameter	mm	low	x		x	
	screw length (thread)	mm	medium	x		x	
	screw type	tapping screw, metric screw	medium	x		x	
	material	(metal sheet, wood,...)	low	x		x	
	magnetic screw	yes/no	high	x		x	
installation/assembly situation	chamfer on screw	yes/no	high	x			
	output of the screw	yes/no	low				
	additional elements	yes/no	high		x	x	
	lack of space	yes/no	high		x		x
	visual disability	yes/no	high		x		x
	risk of injury	yes/no	low		x		x
	additional fix the add part	yes/no	medium		x		x
	working both hands feasible	yes/no	medium		x		x
	screwing in	vertical,horizontal	medium		x		x
tools, equipment	threaded sleeves used	yes/no	high		x		x
	glove used	yes/no	low		x		x
	equipment used	yes/no	low		x		x
	tool	without, manually, automatically	high		x		x
installation regulations	additionally tighten	yes/no	low				
	tightening torque	Nm	low		x		
	check torque	yes/no	low		x		
	assembly sequence	part A,part B, part C, ...	medium		x		x
	multi-stage screwdriving	yes/no	medium		x		x
additional elements	flat washer	number(1,2,3...)	medium		x		x
	nut	number(1,2,3...)	medium	x		x	
	LocTite	amount(ml)	low		x		x
	grease	amount(ml)	low		x		x

Figure 4. Weighting of Product Assembly Informations (PAI).

2.3. Aggregation of Information and utilization of the concept

The implementation of the presented approach was challenging due to high requirement for interconnection and the overall quality of the existing data in different Data Mining systems. In particular the number of realized and existing assembly connections and thus necessary instances of a product assembly information as well as the quality of the data are important. As proof of concept, feasibility of the presented

concept is verified with artificial test data. But in order to evaluate the quality of the results, it is necessary to rerun the analyses with production data. Furthermore, the selection of the properties and attributes for the analysis in particular also has to be determined based on production data to ensure the reliability of generated results. In this scope a special focus is on the characteristics of the parts and of the connection itself. In conformity with the presented objectives and concept a utilization of the methodology is described as follows.

Suggesting assembly connections: The designer creates a new module with already known and new assembly connections in the CAD system. He designs assembly connections and complements these connection properties in the context of the new module. Via the automated Data Mining process, he is provided with various information about the assembly connections. Moreover, for each assembly connection a list of alternative or ever realized connections can be created. Depending on the product properties the five most similar product assembly informations are made available to the designer as a prepared proposal list, which is generated through cluster analysis of existing product data. If the analysis is dispensed with the filtering of associated connections with the product assembly information, the designer can also be provided with other not associated connections as alternatives.

Estimation of assembly process and information: The production planner drafts an initial assembly process for a new assembly at an early stage of product development. Analogous to the use case of the designer, for known assembly connections that are implemented in the new product as well as in the old product data, the right product assembly informations and thus the assembly processes are found. For new unknown connections the most similar product assembly informations and related assembly processes from the database are determined and duplicated. Each of the founded product assembly information represents a single connection and the linked process represents precisely the assembly work content for this connection. The sum of the individual connections for the new product is its first assembly process. Thereby an initial draft of an assembly process of the new module can be generated. The founded individual connections, the individual process, as well as the overall process can be used in different ways to assist the designer and the production planner. The planner and designer also get a first estimation for the expected assembly time and cost in the automated process. In addition, the production planner can increase the quality of the process by manual intervention. On the one hand he adapts the product assembly informations, which are created by the designer, before the Data Mining analysis. On the other hand he can complete the product assembly informations in the attributes with practical knowledge. Thus he has an impact on the input of the Data Mining analysis and increases the quality of the result thereby. Furthermore, the designer has a first draft for the assembly process at one's disposal and a first estimated assembly time in the current CAD system. By a company-specific factor, the designer receives also information about the cost of the connection in the assembly. By verifying this information, the designer can evaluate and compare the alternatives for different connections.

3. Conclusions and Outlook

Through utilization of Data Mining tools the quality of planning results and planning processes can be increased, while simultaneously time and cost reduction can be

realized. In this regard, the presented approach contributes an important added value to production design and planning through usage of knowledge in the existing systems. Thus reduction of planning time, increasing availability of information in product design as well as making the cooperation between the designer and product planning teams easier are the consequences. The technical feasibility of the proposed solution has been shown by a prototypical implementation of the concept in CAD and PDM systems. However, to produce reliable outcomes the product data have to fulfil high requirements in regard to connection elements. Concurrently the necessary data model and some tool sets are provided to make the data integration easier. In the future further development of tool sets and methods could help to reduce the high initial effort for adjustment of the data even more. Besides the evaluation of the results based on product data it is important to investigate the behavior and results of the methodology for new and innovative assembly technologies. Furthermore, for analyzing more complex data sets as well as obtaining better results, it is important to develop and refine the concept and to apply further Data Mining methods.

Acknowledgements

The research project “Prospective Determination of Assembly Work Content in Digital Manufacturing (ProMondi)” is supported by the German Federal Ministry of Education and Research (BMBF) within the Framework Concept ”Research for Tomorrow’s Production” (funding number 02PJ1110) and managed by the Project Management Agency Karlsruhe (PTKA). The authors are responsible for the contents of this publication.

References

[1] U. Bracht; T. Masurat, The Digital Factory between vision and reality, *Computers in Industry 56, pp. 325-333*, 2005.

[2] H. Bley; C. Franke, Integration of Product Design and Assembly Planning in the Digital Factory, *Annals of the CIRP, Vol. **53/1**, pp. 25-30*, 2004.

[3] G. Boothroyd: *Assembly Automation and Product Design*, Second Edition, Taylor & Francis Group, Boca Raton, 2005

[4] B. Lotter, H.-P. Wiendahl, *Montage in der industriellen Produktion, Ein Handbuch für die Praxis*, 2. Auflage, Springer-Verlag Berlin-Heidelberg, 2013

[5] B. Rekiek, A. Delchambre: *Assembly Line Design - The Balancing of Mixed-Model Hybrid Assembly Lines with Genetic Algorithms*, Springer-Verlag London, 2006

[6] O. Erohin; P. Kuhlang; J. Schallow; J. Deuse: Intelligent Utilisation of Digital Databases for Assembly Time Determination in Early Phases of Product Emergence, *45th CIRP Conference on Manufacturing Systems 2012, Vol. 3, pp. 424-429,* 2012.

[7] J. Schallow; K. Magenheimer; J. Deuse; G. Reinhart: Application Protocols for Standardising of Processes and Data in Digital Manufacturing, in: *ElMaraghy, H. A. (Hrsg.): Enabling Manufacturing Competitiveness and Economic Sustainability - Proceedings of 4th CIRP Conference on Changeable, Agile, Reconfigurable and Virtual Production (CARV2011), 2.-5. October 2011*, Montreal, Canada, Springer, Berlin, Heidelberg, New York, pp. 648-653, 2011.

[8] J. Han; M. Kamber; J. Pei: *Data Mining: Concepts and Techniques*, third edition, Morgan Kaufmann Publishers, Waltham, 2012.

[9] Y. Yin; I. Kaku; J. Tang; J. M. Zhu: *Data Mining: Concepts, Methods and Applications in Management and Engineering Design*, Springer-Verlag, London, 2011.

[10] J. Hartung; J. Schallow; S. Rulhoff: Moderne Produktionsplanung - Integration in der Produktentstehung, *ProduktDaten Journal 19 1, pp. 20-21*, 2012

[11] M. Eigner; R. Stelzer: *Product Lifecycle Management - Ein Leitfaden für Product Development und Life Cycle Management*, Springer-Verlag, Berlin, Heidelberg, 2009.
[12] D. Petzelt; J. Schallow; J. Deuse; S. Rulhoff: Anwendungsspezifische Datenmodelle in der Digitalen Fabrik, in: *ProduktDaten Journal 16 1, pp. 45-48*, 2009.
[13] L. Ohno-Machado; H. S. Fraser HS; A. Øhrn: Improving Machine Learning Performance by Removing Redundant Cases in Medical Data Sets, *AMIA Fall Symposium, pp. 523-527*, 1998.
[14] D. Zhang; P. L. Yu; P. Z. Wang: State-dependent weights in multicriteria value functions, *Journal of Optimization Theory and Applications, Vol.74, No.1, pp. 1-21*, 1992
[15] S. Dhanabal; S. Chandramathi: Review of various k-Nearest Neighbor Query Processing Techniques, *International Journal of Computer Applications Vol. 31, No.7*, 2011
[16] R. Kohavi: A study of cross-validation and bootstrap for accuracy estimation and model selection, in: *14th international joint conference on Artificial intelligence, Vol. 2 (IJCAI'95)*, Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, pp. 1137-1143, 1995