INFORMATION
SYSTEMS IN
MANAGEMENT

Information Systems in Management (2016) Vol. 5 (3) 336–346

IDENTIFYING PART COMMONALITIES IN A MANUFACTURING COMPANY DATABASE

JAKUB KWAPISZ [a], VIRGINIA INFANTE [b]

[a] *MITPortugal, Instituto Superior Tecnico, Universidade de Lisboa, Lisbon, Portugal*
[b] *LAETA, IDMEC, Instituto Superior Técnico, Universidade de Lisboa, Lisbon, Portugal*

Manufacturing companies that produce and assemble multiple products rely on databases containing thousands or even millions of parts. These databases are expensive to support, maintain and the inherent complexity does not allow end users to utilize fully such databases. Designers and engineers are often not able to find previously created parts, which they could potentially reuse, and they add one more part to the database. Engineered improvements without removal of the previous version of the component also cause the avoidable increase of elements in the database. Reuse of parts or planned development of common parts across products brings many benefits for manufacturers. Search algorithm utilized across part databases and varying projects allows identifying similar parts. The goal is to compare part names and attributes resulting in the assignment of a similarity score. Determining common and differentiating part attributes and characteristics between pairs of components allows nominating parts that can become shared in different products. The case study utilizes an industrial example to evaluate and assess the feasibility of the proposed method for identifying commonality opportunities. It turned out that it is possible to find many parts that can be potentially shared between different products.

Keywords: Commonality; Data mining; Industrial database analysis; Part reuse; Product platform, Product family

1. Introduction

Identifying existing parts that can be shared across different products allows implementation of strategies increasing competitive advantage. Manufacturing companies aim to fulfill particular customer needs by introducing product differentiation while taking advantage of mass production efficiency [10]. Development of a single product efficiently and effectively is not enough for high market penetration in many industries. Academia and industry have well recognized product families as a compelling method with which it is possible to gain the economy of scale to accommodate an increasing product variety [13]. The product family is a group of related goods that have similar physical characteristics and share the same production processes. An example of the product family is an Airbus A330 family that has six products the A330-200, A330-300, A330-200F, ACJ330, A330 MRTT, A330neo [17].

Product platforms serve as a base for different subassemblies that are shared by product families or a set of products. Platforms allow the consolidation of modules with standard geometries and interfaces in several combinations. Such approach leads to high flexibility and variety of products, at the same time it assures low amount of parts used [7, 12]. Companies that manufacture products for different markets such as customer goods, electronics, automotive, and previously mentioned airspace industry create product platforms. Typically key components that are used to constitute modules, or entire platform are engines, electric motors, batteries, controls, etc. It is important to have key internal parts common, to assure low manufacturing and development costs, and at the same time high exterior parts variety to serve different market segments [6]. Complex products that are made from the considerable amount of parts and are highly customized are challenging from platform design perspective. Companies that produce a range of such products may benefit from sharing not only key components but as well other additional parts. Sharing allows decreasing development time and cost as well as improves the economy of scale during manufacturing [5]. Commonality is a part of the product family and product platform strategy that is dedicated to intentional sharing of parts, processes, operations [14]. Implementation of such an approach requires a management support that will enforce usage of parts that are created as common and preventing from local optimizations. Common parts are designed in the way to meet the requirements of the most constrained product in which it will be used. In a less constrained product, such part could be potentially improved, but then overall benefits (economies of scale, learning curves, etc.) will be lost [4]. For many designers, it is often hard to understand the overall benefits of commonality. Designers tend to create parts the way they like and not reuse others designs or concepts [13]. In large manufacturing companies, every day hundreds of new parts are created. Part databases are becoming vast and complex to use. Such situation makes even harder for designers to find components that they could potentially reuse.

This paper is trying to tackle commonality problem from a different perspective than previously found in the literature. The first research question is to find out: "Are there potential commonality opportunities in bus manufacturer part database?". The research was conducted with the usage of a part database from one of the leading bus manufacturers. This study not only investigates a new commonality approach but as well analyze an industry in which commonality and platform design approach is rarely implemented. Buses are composed of thousands of parts. Parts are usually developed and manufactured for a particular product and, apart from standard parts, are infrequently shared between different vehicles produced by the company. Every week hundreds of new parts are created and placed in the part database. The first hypothesis is that there are a large number of parts in already existing part databases that can become common. Parts created for a particular application, potentially may have similar enough form fit and function as at least one other part. If that is the case, such parts can be shared between different products. Finding the way of selecting parts for commonality study among of thousands of parts that are already made and stored in the part database is a key problem this paper investigates.

2. Concepts related to commonality search

In previous academic work, it was assumed that it is easy to find or nominate commonality parts. As an example "These five elements can be whatever the designer feels would be good candidates to make common among a family of designs" [3]. In practice in large companies with different divisions and many employees it is hard to make such a decision. Moreover incorrect selection of common part leads to divergence and additional part development that is expensive [1]. Since existing literature presenting how commonality nominations can be made is limited, interviewing three different manufacturing companies allows finding out an industrial practice. Commonality nominations in some companies are made not by the designers but a few specialists or in some cases entire teams. The specialists preference often comes from expertise that they have and do not take to account all of the available data or know-how of other employees. Finding a way of searching for commonality opportunities with the usage of information's available in databases allows closing a gap in academic research as well as helping practitioners to implement widely commonality strategy in their companies.

Since there is no currently available part commonality search tool, a variety of solutions from different areas were investigated. The closest to similar part detection are duplicate detection algorithms. Duplicate document or file is the same as the other record [15]. Some methods use feature selection, similarity measure, or discriminant derivation to find duplicated documents [9]. For each sentence in the record, special terms are selected as features. Then the similarity degree between features of one document is compared with the other document. Such approach is not applicable for commonality part detection because large text documents rarely describe information about parts. There is a stream of research focusing about database entry duplication. In that case, two records are duplicates if they identify the same real world entity [8]. Such research is applicable mostly to the databases containing information about people (patients, clients, citizens, criminals, etc.) or places (restaurants, stores, businesses, etc.), but as well other database entries [2, 11]. In those situations, duplicates are found and removed from a database for clarity. There are few key differences, from the method and research objective perspective, between duplicate detection approach and the approach to be selected for part commonality nomination. First of all attributes about people and places have a limited number of options and differentiations between attributes entries can be predicted. Abbreviations and spelling mistakes of addresses or names are finite. For example, "Avenue" will be either "Ave." or "av.". Moreover, last name and first name of a person may be exchanged or misspell in few typical ways such as “John” and “Jon”. In the case of part names there are no standards and spelling and phrases used may vary significantly. To improve the speed of search databases containing information about people or places, the records may be pre-sorted alphabetically. After alphabetic sorting, pairs are compared only within a range to avoid time complexity of $O(n^2)$ comparison. In the situation of manufactured parts, alphabetic sorting and searching only in limited range would decrease detection rate of commonality opportunities. Detection of exact duplicates of a part would be easier, but it is not the point of the commonality search. Two database entries that are exactly the same increase part database complexity, but do not cause manufacturing or procurement problems. The part number is determining if the database entry is recognized as real part. In a situation when part numbers are different, and the other attributes of the pair of parts are the same such situation may lead to problems or wasted opportunities. Another difference between approaches is a similarity measure of part attributes not duplication. Part dimensions, dates, tolerances, and other part attribute values even if not identical they may indicate that parts are commonality opportunity. For example steel bar which length is 100mm may potentially become a common part, in the exchange of steel bar 98mm. This is just an example, and in particular situation additional evaluation need to be made before one part will be shared between products. Commonality opportunity search is comparable but different than typical data mining and data matching. Described in next chapter part similarity detection uses elements of available tools and approaches that are usually used in various applications.

3. Detection of part similarity for commonality nomination

This chapter describes a way of nominating parts that can potentially be shared between different products. The initial step is the extraction of useful data from a part database. A part is usually characterized by many attributes, whose similarities help to decide about the pair of parts similarity. It is crucial to select the part attributes from a database that allows the record match. Documentation experts, depending on comprehension of the part database, choose attributes. Fig. 1 presents all the major steps of part similarity method.

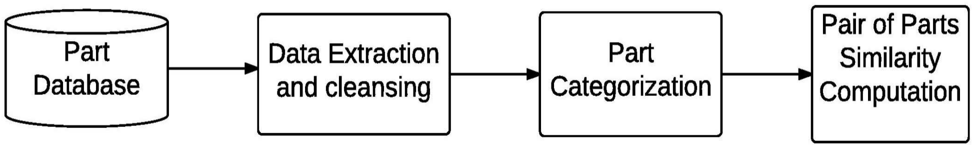

Figure 1. Scheme of a method identifying similar parts in part database

After information about parts is extracted from the database, a separate part database can be created and stored in a format that is uncomplicated to process. Moreover, data need to be cleansed from the entries and symbols that are not helpful from part similarity perspective. During part categorization algorithm find and display a list of words with number how many times they appeared in part names. Most frequently reappearing words are displayed on top. A user chooses with which type of parts he would like to proceed (the selected keyword represents that). After initial keyword is chosen, then this process is repeated in order to have two keywords part categorization. Next step is a pair of parts similarity computation. A part is usually characterized by many attributes that contribute differently to deciding whether a pair of parts is a commonality opportunity. Part attribute weight is assigned based on the importance of its contribution to the commonality decision-making. Ideally, company experts assign the weights. In the next step, the algorithm calculates and displays score that determines how the pair of parts is similar. The two main element of the total score assigned to a pair of parts under evaluation is Jaro–Winkler distance and part attributes similarity. Jaro-Winkler algorithm is found effective for fields like name or other short descriptions [16]. In the proposed algorithm, Jaro-Winkler is used on part name dataset to determine the similarity between a pair of parts part names. In the last step, the algorithm displays the list of pair of parts with the total similarity score. The user needs to decide if the pair of parts is a true positive or not. Commonality opportunities may be treated as a nomination for further in-depth study or in the situation when enough information is available a commonality decision may be immediately made.

4. Industrial Example

Described method was implemented in leading bus manufacturer part database. The first step was to check the current status of parts shared between the same types of vehicles. Figure 1 a) presents a three the same type of buses developed for different cities. A number of individual parts per bus is different, and the number of common parts is smaller than expected. Figure 1. b) shows types of parts that are used multiple times in a bus developed for a particular city. The thickness of the edges, drawn as curved links, indicates if a part is used many times or in small quantity.

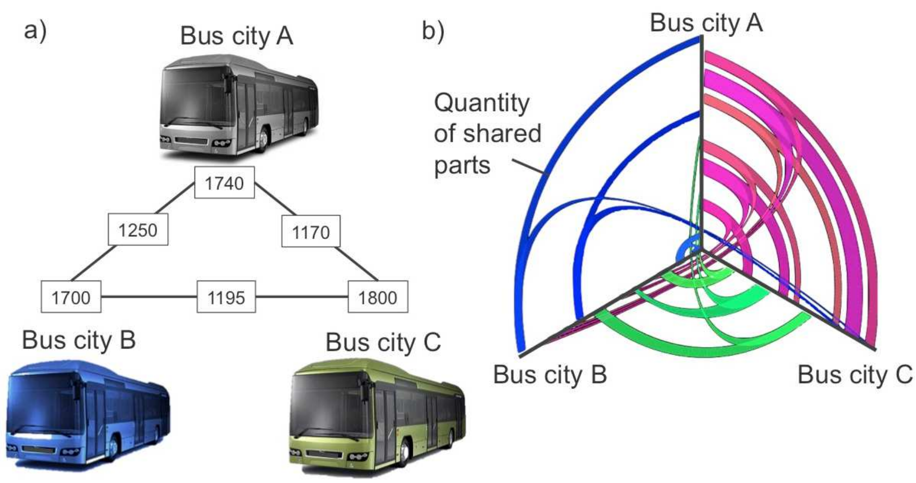

Figure 2. Current commonality status at investigated company. **a)** A number of individual parts per bus and number of shared components in the same bus type. **b)** Different quantities of the same part types used in buses (an example of top 6 parts)

In presented case the thickness of the edges was changing, it means that in the same type of buses different quantities of identical parts are used. Such situations allow making a claim that there are solutions that possibly fulfill the same function but have a different design. The total number of extracted parts from the database with part attributes was almost 380,000. Extracted information had to be cleansed to remove counterproductive part attributes and entries. After cleansing, every part had part number and 6 part attributes, to which company expert assigned weights of importance as described in previous chapter. Part attributes with weights are as follow: revision (0.1), unit (0.05), part name (0.65), supplier (0.1), drawing number, date (0.025), and designer (0.075). Company expert was asked to choose an interesting keyword from displayed list of the most popular word is part names.

The recurring word ‘sensor' was selected, which was used 1568 times. Algorithm analyzed part names that included word sensors and displayed list of the most reoccurring words in the group of sensors. The word ‘temperature' was used 180 times. That is why ‘temperature sensors' were chosen for pairwise comparison. To confirm the pair of parts nomination for commonality opportunities they were checked manually. Figure 3 presents the number of commonality opportunities found within a range of results.

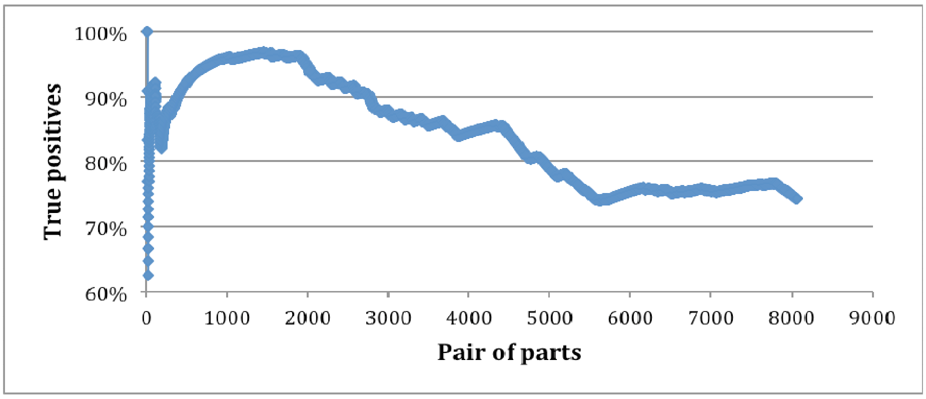

Figure 3. Commonality opportunity detection rate – percentage of true positives vs. a number of pair of parts compared

A total number of 4942 pairs of parts is above 80% threshold. True positives are the parts nominated by the algorithm and manually confirmed by the company representative.

Nominated pairs of parts are the commonality opportunities in the subgroup of temperature sensors. Since every temperature sensor was compared with all the other temperature sensors, it is possible that one part can be used in an exchange of more than one other part. From commonality perspective, the most beneficial is the situation when one part replaces as many as possible other parts without sacrificing their performance.

5. Discussion

The purpose of this part of the paper is to state the results interpretations, present opinions and make suggestions for next steps, which may lead to future research. Its primary function is to address the problem posed in the introduction and, how the achieved solution corresponds to the existing knowledge. Also, this section identifies potential limitations and weaknesses of the results and how they may affect the findings.

This paper tackles part commonality from another perspective than previously found in the literature. The research answered a question asking if there are potential commonality opportunities in bus manufacturer part database. The created algorithm found multiple parts, in a part database from a leading bus manufacturer, that are commonality opportunities. Parts are developed and manufactured for a particular product now have a greater chance to be potentially shared between different vehicles produced by the same company. The stated hypothesis turned out to be true; there are thousands of parts in already existing part databases that can become common. The time required to find and select parts for commonality study among of thousands of parts that are already made and stored in the part database is significantly reduced with the developed algorithm.

No previous research presents data that could directly be compared with the results shown in this paper. The closest stream of research is focusing on database entry duplication. Such research is applicable mostly to the databases containing information about people or places. There are few key differences from the method and research objective between this approach and the approach used to nominate parts for commonality study. Attributes describing people and places have a limited number of options. The differences between part attributes entered by designers are harder to predict. Detection of exact duplicates is not the point of the commonality search. Exactly the same database entries increase part database complexity, but do not cause manufacturing or procurement problems. Presented algorithm for commonality search is based on Jaro-Winkler distances that allow comparing part names. The originality of this work is a combination of string comparison with the weight assigned to part attributes. Another differentiation factor and gain for the body of knowledge is an industrial example. It proves that there is an opportunity and need for creation of such algorithm in the industry. In industry there is a greater implementation potential when solution is simple and straightforward. Company do not have to invest time and resources to learn how the algorithm works and ho to use it. Results presented from one case study are often treated as not enough statistically significant. Future work could also focus on additional companies from similar industry or with a similar data structure.

This paper presents an initial stage of the algorithm development and highlights its industrial application. Created algorithm could be improved to work faster, and a number of false positives could be reduced. It could be done by greater analysis of part names and possible inputs. On the other the same algorithm would applied in a different industry may have decreased performance due to such customization. Another limitation is that after commonality nominations are found there still need to be a vast amount of engineering work to make the common part. It is because common parts are designed in the way to meet the requirements of the most constrained product in which it will be used. That is why in some cases what is theoretically similar part it would have to be adapted to be utilized in few products.

The last problem is connected with the workflow that can be found in many companies. The designers often do not understand the overall benefits of commonality. Even with the usage of presented algorithm, the designers may not be willing to reuse existing designs or concepts. Such problem would have to be solved by an appropriate management approach.

6. Conclusions

Algorithm was proven to be a useful method to nominate parts for in-depth commonality study. This method allows searching a part database to find interesting groups of parts and present the results of the individual part comparison.

Implementing the method in an industrial example allows proving that there are possible commonality opportunities in bus manufacturer part database. Conducting research with the collaboration of the leading bus manufacturers allowed analyzing an industry in which commonality and platform design approach is rarely implemented. Algorithm allowed finding a large number of parts in already existing part database that can become common. Those parts that were created for a particular application, but have similar enough form fit and function as at least one other part, can be potentially shared between different products. Jaro-Winkler distance turned out to be a suitable method of comparing part names. Other types of string metric could be potentially used for measuring the difference between two part names. Only key part attributes that describe a part should be used to identify commonality opportunities. Weights assigned to part attributes are crucial for correct commonality opportunity search process. In this particular case, highest weight was assigned to the part names as the essential attribute in this part database.

Acknowledgements

This work was supported by the Fundação para a Ciência e a Tecnologia (FCT) under grant SFRH/BD/51595/2011. This work was also supported by FCT, through IDMEC, under LAETA, project UID/EMS/50022/2013.

REFERENCES

[1] Boas R, Cameron B, Crawley EF. Divergence and Lifecycle Offsets in Product Families with Commonality. Systems Engineering. 2013.

[2] Christen P. Data Matching: Concepts and Techniques for Record Linkage, Entity Resolution, and Duplicate Detection. Springer Science & Business Media; 2012.

[3] Corl MJ. Methodology for Optimizing Commonality Decisions in Multiple Classes of Ships. 2007 [cited 2015 Jun 28];Available from: http://deepblue.lib.umich.edu/handle/2027.42/55684

[4] Corl MJ, Parsons MG, Kokkolaras M. Optimal Commonality Decisions in Multiple Ship Classes [Internet]. In: Simpson TW, Jiao J (Roger), Siddique Z, Hölttä-Otto K, editors. Advances in Product Family and Product Platform Design. Springer New York; 2014 [cited 2015 Jan 7]. p. 625–645.Available from: http://link.springer.com/chapter/10.1007/978-1-4614-7937-6_25

[5] Fisher M, Ramdas K, Ulrich K. Component Sharing in the Management of Product Variety: A Study of Automotive Braking Systems. Manage. Sci. 1999 Mar;45(3):297–315.

[6] Hanafy M, ElMaraghy H. A modular product multi-platform configuration model. International Journal of Computer Integrated Manufacturing. 2015 Sep 2;28(9):999–1014.

[7] Hoppmann J, Rebentisch E, Dombrowski U, Thimo Zahn. A Framework for Organizing Lean Product Development. Engineering Management Journal. 2011 Mar; 23(1): 3–15.

[8] Kelkar BA, Manwade PKB, Patil PGA. Near Duplicate Detection In Relational Database [Internet]. ESRSA Publications; 2013 [cited 2015 Jul 14]. Available from: http://www.ijert.org/view-pdf/2879/near-duplicate-detection-in-relational-database

[9] Lin Y-S, Liao T-Y, Lee S-J. Detecting near-duplicate documents using sentence-level features and supervised learning. Expert Systems with Applications. 2013 Apr; 40 (5): 1467–1476.

[10] Liu N, Choi T-M, Yuen CM, Ng F. Optimal Pricing, Modularity, and Return Policy Under Mass Customization. IEEE Transactions on Systems, Man and Cybernetics, Part A: Systems and Humans. 2012 May;42(3):604–614.

[11] McCoy AB, Wright A, Kahn MG, Shapiro JS, Bernstam EV, Sittig DF. Matching identifiers in electronic health records: implications for duplicate records and patient safety. BMJ Qual Saf. 2013 Mar 1; 22 (3): 219–224.

[12] Meyer MH, Lehnerd AP. The power of product platforms: building value and cost leadership. Free Press; 1997.

[13] Nobelius D, Sundgren N. Managerial issues in parts sharing among product development projects: a case study. Journal of Engineering and Technology Management. 2002; 19 (1): 59–73.

[14] Shi J (Junmin), Zhao Y. Component commonality under no-holdback allocation rules. Operations Research Letters. 2014 Sep; 42 (6–7): 409–413.

[15] Wang G, Chen S, Lin M, Liu X. SBBS: A sliding blocking algorithm with backtracking sub-blocks for duplicate data detection. Expert Systems with Applications. 2014 Apr; 41 (5): 2415–2423.

[16] Yan S, Lee D, Kan M-Y, Giles LC. Adaptive Sorted Neighborhood Methods for Efficient Record Linkage [Internet]. In: Proceedings of the 7th ACM/IEEE-CS Joint Conference on Digital Libraries. New York, NY, USA: ACM; 2007 [cited 2015 Oct 28]. p. 185–194.Available from: http://doi.acm.org/10.1145/1255175.1255213

[17] A330 Family: A330-200, A330-300 - A330 photos, pictures, A330 videos, A330 3D view | Airbus | Airbus, a leading aircraft manufacturer [Internet]. airbus. [cited 2015 Oct 27];Available from: http://www.airbus.com/aircraftfamilies/passengeraircraft/a330family/