Procedia CIRP 90 (2020) 617–622

Contents lists available at ScienceDirect

Procedia CIRP

journal homepage: www.elsevier.com/locate/procir

27th CIRP Life Cycle Engineering (LCE) Conference

Big data analysis for the estimation of disassembly time and de-manufacturing activity

Claudio Favi[a,*], Marco Marconi[b], Marco Mandolini[c], Michele Germani[c]

[a] Department of Engineering and Architecture, Università di Parma, Parco Area delle Scienze 181/A, 43124, Parma, Italy
[b] Department of Economics, Engineering, Society and Business Organization, Università degli Studi della Tuscia, Largo dell'Università, 01100, Viterbo, Italy
[c] Department of Industrial Engineering and Mathematical Sciences, Università Politecnica delle Marche, via Brecce Bianche 12, 60131, Ancona, Italy

A R T I C L E I N F O

Keywords:
Design for disassembly
Big data analysis
Data mining
Disassembly time
De-manufacturing

A B S T R A C T

Design for disassembly is a key enabling strategy for the development of new business models based on the Industry 4.0 and circular economy paradigms. This paper attempts to define a method, based on Data Mining, for modelling disassembly data from large amount of records collected through the observation of de-manufacturing activities. The method allows to build a repository to characterize the disassembly time of joining elements (e.g. screws, nuts) considering different features and conditions. The approach was preliminary tested on a sample of 344 records for nuts disassembly retrieved by in-house tests. Disassembly time and corrective factors were assessed including the analysis of probability distribution function and standard deviation for each feature (i.e. disassembly tool).

1. Introduction

The concept of Industry 4.0 has gained growing importance for manufacturing companies in recent years. Industry 4.0 is the latest manufacturing paradigm blending the physical and digital worlds where a key feature is the creation of highly automated industries through human-machine interaction (BMBF-Internetredaktion, 2019). Industry 4.0 is enabled by nine technology breakthroughs in fields including: (i) robotics, (ii) cloud computing, (iii) big data analytics, (iv) additive manufacturing, (v) internet of things, (vi) augmented reality, (vii) cyber security, (viii) system integration, and (ix) advanced simulation. (Zhong et al., 2017) As a consequence, new players, business models and opportunities driven by these technologies are emerging. For example, circular economy (CE) is a new business model orientated to increase economic opportunities against resource shortages and waste. Industries and enterprises are supporting the CE perspective by adopting closed-loop life cycle models for the development of new products and services (Reike et al., 2018). Recycling, remanufacturing and reusing represent possible scenarios in dealing with this new paradigm. Product lifetime extension is another developing strategy aiming at the postponement or reversal of the obsolescence of a product through deliberate intervention (den Hollander et al., 2017). Product disassembly plays a critical role in CE. Disassembly is defined as a systematic method for separating a product into its constituent parts, components and subassemblies (Gungor and Gupta, 1998). Selective disassembly is adopted for both maintenance and EoL product components recovery (Vanegas et al., 2018). Product disassembly generally occurs in the last part of the product life cycle, but it originates in the preliminary phases of product design. Thus, design for disassembly (DfD) allows anticipating de-manufacturing issues minimizing disassembly time/costs and providing best disassembly sequences for each target component.

Disassembly time is one of the most important metrics to guide the implementation of DfD strategies and thus to adopt CE models (Mandolini et al., 2018). Disassembly time depends by several factors, such as component shape, joining elements, disassembly directions, disassembly tools and equipment (Güngör, 2006). It is influenced by the product workload (life cycle stress), working environment, chemical and physical degradation (ageing), deformation, cleanliness, material type, coating/painting processes, etc. (Duflou et al., 2008). The condition of the product and its constituent components could be uncertain when disassembly occurs, and this kind of information needs to be processed systematically in order to develop any realistic and credible disassembly plan (Zhu and Roy, 2015). Thus, disassembly time is not the reverse of assembly time and its estimation requires the analysis of real de-manufacturing operations performed in different contexts (e.g. dis-

* Corresponding author.
E-mail address: claudio.favi@unipr.it (C. Favi).

https://doi.org/10.1016/j.procir.2020.01.072

mantling centres and/or service centres) (Sodhi et al., 2004). This issue has been faced with the aim to collect all this information in a structured way (Favi et al., 2016).

However, a new challenge is emerging related to the processing of big amount of data, which is one of the enabling technologies of Industry 4.0. Indeed, Industry 4.0 big data could come from many and diverse sources and include product and/or machine design data, product and process-quality data, records of manual operations carried out by staff, manufacturing execution systems, fault-detection and other system-monitoring deployments, logistics information including third-party logistics, customer information on product usage, feedback, and more (Mourtzis et al., 2016). Some of these data are structured (such as sensor signals), some are semi-structured (such as records of manual operations), and some others are completely unstructured (such as image files). Thus, a large and heterogeneous amount of information is currently available as output of new manufacturing systems based on the Industry 4.0 technologies. Latest trends in the big data domain is moving towards providing a level of abstraction to utilize popular data processing platforms (Gokalp et al., 2016). One of the most popular tools in this context is the data mining (DM), technique which allows to sort and cluster data to be re-injected in a design domain, using statistical algorithms. DM is often part of a larger data modelling process, called Knowledge Discovery from Data (KDD). Indeed, KDD refers to the overall process of discovering useful knowledge from data. It involves the evaluation and interpretation of patterns to make the decision of what can be qualified as knowledge (Fayyad et al., 1996). It also includes pre-processing, sampling, and projections of the data prior data mining. DM is defined as "the nontrivial process of identifying valid, novel, potentially useful, and ultimately understandable patterns in data" (Fayyad et al., 1996). DM allows "extracting valid, previously unknown, comprehensible information from large databases" (Friedman, 1998). In recent years, only few attempts have been done to translate manufacturing data in structured design information by the use of DM (Kretschmer et al., 2017; Bae and Jinhwa, 2011).

The application of DM techniques in DfD has not yet been investigated, and it requires an initial collection of de-manufacturing operations. This gap emerges by the analysis of the scientific literature and industrial needs. Thus, companies adopt cutting-edge platforms that leverage the value of manufacturing big data using analytic tools (i.e. data mining). Manufacturers today seek to achieve true business intelligence through collecting, analysing, and sharing data across all key functional domains (from manufacturing to design).

The paper proposes a method to collect all relevant data from different industry sources (dismantling/de-manufacturing centres and, service centres) and extract significant information, to be stored within structured repositories, that can be re-injected in the design domain to support DfD activities. The method is based on statistical tools able to assess standard disassembly time of joining elements (bolts, screws, nuts, rivets, etc.). The DM process includes the characterization of corrective factors related to the joining element features (e.g. type, dimension, tool used) and the overall condition of the product during disassembly (e.g. wear, deformation, rust). The use of DM as enabling technology in the context of product disassembly and DfD is the main novelty of this work.

The paper is structured as follows. After this introduction, the main section of the paper presents the DM approach for the assessment of standard disassembly times and characterization of features related to disassembly operations. The proposed method has been tested for evaluating the most relevant corrective factors to be considered for estimating the unscrewing time of multiple types of nuts. At last, the conclusion section summarizes benefits and drawbacks of the proposed approach.

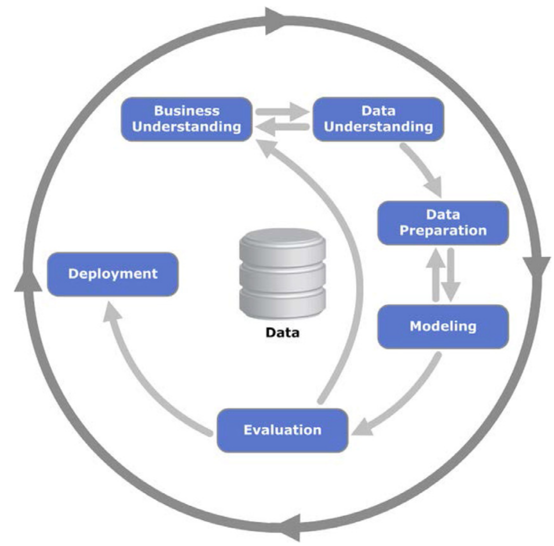

Fig. 1. CRISP-DM Process Diagram (Source: Wikipedia, based on (Shearer, 2000)).

2. A knowledge discovery in database approach for disassembly time calculation

This section explains the Data Mining process (called *Dis-DM*) thought for defining knowledge required for estimating disassembly times, to be used by the method presented in (Mandolini et al., 2018). According to such approach, the disassembly time required for removing a selected component (or a group of components) from a product depends by two aspects: (i) the sequence of operations expected by the disassembly path, and (ii) the disassembly time needed to perform each operation. A disassembly operation is defined as the set of actions for removing a liaison that connect two or more components. Disassembly time of a single operation is computed by multiplying the standard disassembly time of a liaison (evaluated considering ideal conditions, such as lack of wear, deformation, rust, etc.) with a vector of corrective factors (e.g. dimensions, material, weight, level of wear, corrosion, deformation). Corrective factors are necessary input to consider the specificity of a component/liaison (e.g. dimensions), as well as its conditions during disassembly (e.g. the more a component is rusty, the more the disassembly time is high). The DM process that will be presented hereafter (*Cross-industry standard process for data mining* – CRISP-DM (Shearer, 2000)) aims to identify, for each kind of liaison, the most relevant corrective factors with relative values for each condition (e.g. '*not deformed*', '*partially deformed*' and '*deformed*' for '*deformation*' corrective factor). Fig. 1 reports the CRISP-DM Process Diagram used within this work. This is a systematic approach based on the six steps expected by a standard DM process: (i) business understanding, (ii) data understanding, (iii) data preparation, (iv) modelling, (v) evaluation, and (vi) deployment.

2.1. Business understanding

When starting a new DM process, it is necessary to define the *business goals* (i.e. what are the business needs?). In this study (*Dis-DM*), the main goal consists in defining elementary information to be used in a systematic approach (explained in (Mandolini et al., 2018)) for analytically estimating time for selective disassembly. The *data mining goal* (i.e. what are the patterns to be defined?), instead, consists in defining the most relevant corrective factors, and relative values, for characterizing the liaisons conditions during disassembly. Such values are then used for estimating the product disassembly time.

2.2. Data understanding

This step aims at the definition of data source, and then to gather and understand the data. Hereafter, is reported the list of activities to be performed for *Dis-DM* process:

- *Define a wide set of corrective factors*: it is required to list all the parameters that can be measured or evaluated during disassembly operations. In this step, a large set of parameters needs to be considered, including the ones that will be neglected during data modelling phase. These parameters are classified in three levels: (i) general parameters that are applicable for all liaisons and components (e.g. working environment, physical degradation, deformation), (ii) liaison class-related parameters (e.g. wear for mechanical liaisons), and (iii) liaison type-related parameters (e.g. head type, length, diameter and unscrewing tool for a threaded liaisons). The set of general parameters should contain data concerning products and components characteristics, as well as the conditions about where and how the disassembly is carried out (e.g. senior/junior operator, manual/automatic disassembly).
- *Select products and components categories*: identify products and components for creating the initial data sets. This cluster contains the liaisons that require a definition of corrective factors (it must be in accordance with the data mining goals).
- *Select data sources*: select and train dismantlers and maintainers (operators in general), which are the sources of disassembly data. Training about the method to be used for collecting data is required in order to avoid possible sources of bias.
- *Define disassembly procedures*: define all the procedures that operators need to adopt during data collection (i.e. the template where classify and characterize liaisons and collect disassembly times). In addition, it is important to characterize the equipment used for disassembling the products, including teaching documents used by operators for establishing the value of each condition (e.g. what is the meaning of '*not deformed*', '*partially deformed*' and '*deformed*' for the parameter '*deformation*'?).
- *Collect disassembly time and liaisons parameters*: direct observation/video recording of operator's activity for collecting data about liaisons (e.g. duration of each disassembly task, need of special tools, difficulties of the disassembly or extraction operation).

2.3. Data preparation

This is generally the most labour-intensive and long step, since data sources need to be selected, cleaned and formatted into the desired form. Since data are defined by operators (data does not come from internet or other channels), this step is quite fast and straightforward. Data should be collected in a n x m matrix, where n is the quantity of tests collected, and m is the quantity of parameters describing a liaison with the addition of one column used for collecting disassembly time.

2.4. Modelling

In this step, statistical tools and algorithms are used for defining data patterns. For *Dis-DM* process, the relevant algorithms used are the following:

- *Correlation test*: it is the analysis which aims to identify the parameters (i.e. corrective factors) that mostly influence the disassembly time of a certain liaison. This phase is required since, when observing disassembly operations, dozens of parameters may be registered, even those ones poorly related to disassembly time. Since such parameters will be used by a software tool for estimating the disassembly time, it is suggested to limit their quantity (Mandolini et al., 2018), considering only the most relevant ones. For this task, the coefficient of determination (R^2, proportion of the variance in the dependant variable that is predictable from the independent variables) can be adopted. It determines which ones of the investigated parameters are significant corrective factors in the determination of disassembly time.
- *Clustering of parameters*: observed parameters are classified in two categories: (i) discrete (e.g. screw head type, screwing type), and (ii) continuous (e.g. screw length, screw diameter). For the first ones, clustering may not be required, since these parameters assume a discrete number of values. Continuous parameters must be clustered because correlation functions are not managed by the database to be filled in (Mandolini et al., 2018)[20]. *K-means* algorithm can be used for this goal. It consists in partitioning n observations into k clusters so that observation belongs to the cluster with the nearest mean. The "*Elbow*" method is one of the methods used for defining the optimal number of clusters for each observed parameter. This method considers the relation between the total within-cluster sum of squares (WSS) and the quantity of clusters. The optimal quantity of clusters is defined as that number for which, adding another cluster, WSS improvement will be less than 10%. The clustering process is performed on corrective factors for isolating the contribution of the analysed parameter on the disassembly time.
- *Analysis of Variance (ANOVA)*. This is a statistical method used to check whether the means of multiple groups are significantly different from each other. ANOVA checks the impact of one or more factors by comparing the means of different samples. The output of this analysis, for each liaison (e.g. screw) and corrective factor, established with the correlation test (e.g. head type), is a list of disassembly times for each condition (e.g. ScrewHeadType.Hexagonal = 6.9 [s], ScrewHeadType.Cylindrical = 7.1 [s]).

2.5. Evaluation

Patterns extracted in the previous stage are evaluated considering business objectives set at first. Since this is a preliminary study, test is mainly focused on evaluating the *distribution fitting* of values established for each corrective factor. This test allows evaluating errors (i.e. standard deviation) determined by associating a certain value to a specific condition of a corrective factor.

2.6. Deployment

This step consists in presenting the knowledge, gained through the DM process, to the involved stakeholders and formalizing such knowledge within software tools. For this specific *Dis-DM* process, corrective factors computed in the previous stages, with relative values, are stored within the database of the software tool described in (Favi et al., 2019).

3. Case study and results

This section presents how *Dis-DM* process was used for defining corrective factors for a specific typology of threaded liaisons: a nut. Since the business and data mining goals are the same ones defined in the previous section and the deployment process is presented in (Favi et al., 2019), this section is focused on the key steps of the *Dis-DM* process: (a) data understanding and preparation, (b) data modelling, and (c) evaluation.

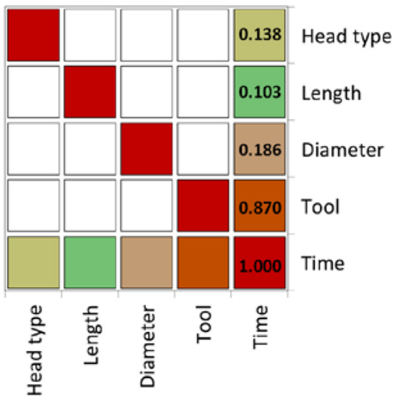

Fig. 2. Correlation matrix (red: high correlation, green: low correlation, white: no correlation).

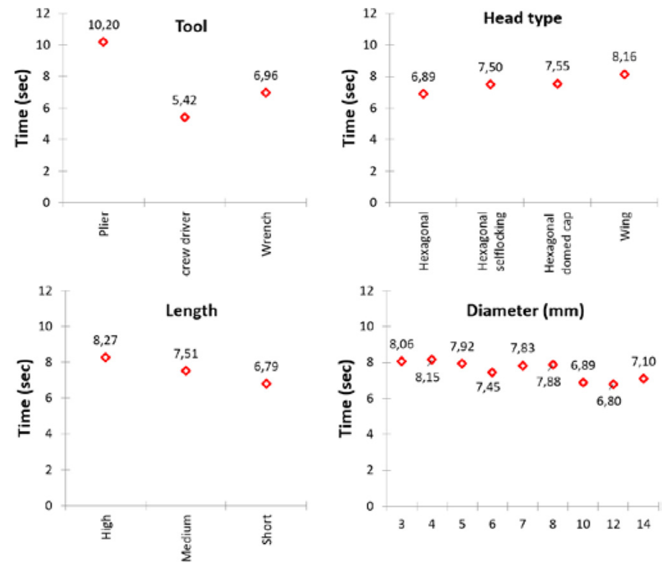

Fig. 3. ANOVA analysis.

3.1. Data understanding and preparation

This case study analyses corrective factors connected to the disassembly time of nuts. Several parameters were identified during the disassembly operations referring to two main groups: (i) product-related features (e.g. nut diameter, nut height, nut head type, etc.), and life cycle-related features (e.g. how long the product was used, deformation, wear, etc.). In this case study, the authors focus only on the first group which includes: nut diameter, nut height, nut head type and unscrewing tool.

Disassembly times collection has been done using a test bench realized and installed within a lab, which is simulating the facility of a de-manufacturing company. It consists of two metal plates connected by a pattern of screws and nuts, both with the same material (steel) and different diameters. The following nut types were considered: (i) four different heads (hexagonal, hexagonal domed cap, wing and, hexagonal self-locking), (ii) three height (short, medium and high) and, (iii) nine diameters (3, 4, 5, 6, 7, 8, 10, 12 and 14 [mm]). In addition, three disassembly tools were used in the tests: (i) a screwdriver, (ii) a wrench, and (iii) a plier. Screwing height is usually bigger than nut height, so in this case the screwing height was set as *1.5 * nut height*, while tightening torque is approximately 1, 2, 4, 8, 14, 20, 40, 50 and 70 [Nm], respectively to the previous list of diameters.

An operator, trained before beginning the test, carried out disassembly operations. Disassembly time was measured as the time required for removing a nut from the relative screw. A stopwatch with a 1/100 precision was used for the measurement. An electronic spreadsheet template was used by a supervisor for registering the disassembly times related to each test (combination of head types, head diameters, head heights and unscrewing tools), for a total of 344 records.

3.2. Modelling

Once finished data collection, a table of 344 rows (1 for each test) and 5 columns (4 factors and 1 disassembly time) was used for data modelling. Correlation test was performed for evaluating which of the four factors was the most important one and whether some of these should be discarded or further processed. Correlation matrix (a symmetric matrix illustrated in Fig. 2) highlights a strong relation between the disassembly time and the tool used for unscrewing nuts (0.870). Hence, such parameter is certainly considered as one of the most important corrective factors. No correlation exists amongst head type, length, diameter and tool because such parameters were used for defining the set of 344 observations.

By analysing the correlation matrix and the related scatter plot, it is possible to observe that diameter, length and head type are poorly correlated with the disassembly time (correlation index lower than 0.2). First, diameters must be clustered (no high correlations can be retrieved with such data), then, diameter, head type and length should be further processed through *Analysis of Variance*. The results of diameter clustering are reported in Table 1 (corrective factor condition column). The ANOVA analysis (Fig. 3) was then computed for evaluating, one at a time, the relationship between the observed parameters and the disassembly time. From such analysis, it is possible to observe, for instance, that a wing nut is more difficult to be unscrewed than a hexagonal one (there are no specific tools for the wing nut head). Furthermore, the disassembly time for self-locking nuts and domed cap are slightly higher than in case of hexagonal nuts.

After that, the observed parameter "*diameter*" is clustered using a *k-means* algorithm. This analysis was performed starting from the output of the ANOVA analysis for such parameter (Fig. 3). Result consists of three clusters (the fourth cluster determines only a 4% improvement of the WSS), with the following centroids: *cluster 1* (diameter = 4 [mm]; time = 8.0 [s]), *cluster 2* (diameter = 7 [mm]; time = 7.7 [s]) and *cluster 3* (diameter = 12 [mm]; time = 6.9 [s]).

The last computing step consists in defining the corrective factors for each observed parameter and condition. For each parameter (e.g. tool) and related condition (e.g. plier, screw driver and wrench), ANOVA analysis was used for calculating the average disassembly time (10.20 [s], 5.42 [s] and 6.96 [s] for plier, screw driver and wrench, respectively). Then, for each parameter, the reference disassembly time is established (i.e. the minimum time for the selected parameter, which defines the optimal condition for this one). At last, the corrective factor for a specific condition is computed by dividing average time of this condition with reference time of each parameter (Table 1).

3.3. Evaluation

For evaluating the quality of corrective factors modelled in the previous phase, a distribution fitting analysis was performed. Through this analysis, it was possible to observe how much the estimated disassembly time, computed considering corrective factors established in the previous step, diverges from observed data. The distribution fitting analysis was carried out considering a corrective factor at a time. For a certain parameter, multiple distribution

Table 1
Corrective factors for "nut" liaison. (*) standard conditions.

Corrective factor type	Corrective factor condition	Average disassembly time [s]	Corrective factors [s]
Unscrewing tool	Plier	10.20	1.88
	Screw driver*	**5.42**	**1.00**
	Wrench	6.96	1.28
	Reference	*5.42*	
Head type	**Hexagonal***	**6.89**	**1.00**
	Hexagonal self-locking	7.50	1.09
	Hexagonal domed cap	7.55	1.10
	Wing	8.16	1.18
	Reference	*6.89*	
Length	High	8.27	1.22
	Medium	7.51	1.11
	Short*	**6.79**	**1.00**
	Reference	*6.79*	
Diameter (mm)	4	8.04	1.16
	7	7.72	1.11
	12*	**6.93**	**1.00**
	Reference	*6.93*	

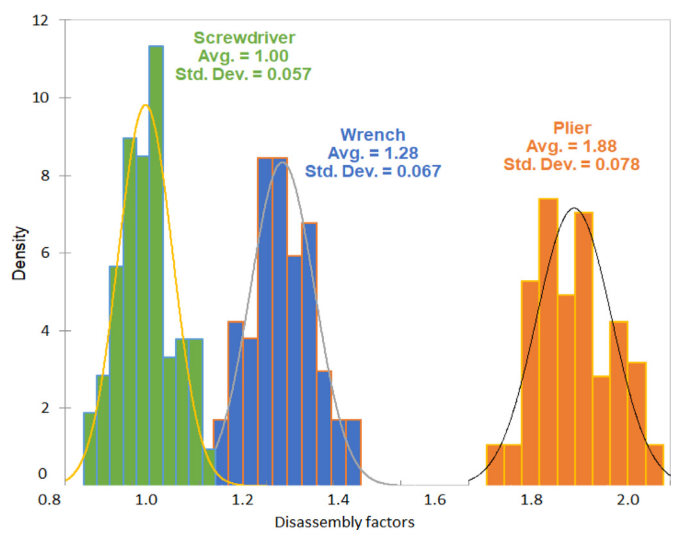

Fig. 4. Distribution fitting of the corrective factors used for considering the unscrewing tool.

fittings were performed (e.g. three distributions were computed for the unscrewing tool). Each distribution was performed considering the average factors retrieved from experimental data. For example, Fig. 4 presents the distribution fitting related to the unscrewing tool. It is possible to observe that the standard deviation (σ) for the three conditions are 0.057 (5.7%), 0.067 (5.2%) and 0.078 (4.1%) for screw driver, wrench and plier, respectively. Considering that observations can be represented with normal distributions (Fig. 4), the uncertainty in estimating disassembly time is proportional to the standard deviation (the lower the deviation the better the reliability). From such analysis, it is possible to conclude that considering one parameter at a time, the disassembly time of a certain liaison and a certain condition can be calculated considering the disassembly time in standard conditions and the corrective factors (one for each condition), used for correlating the actual conditions of a liaison. This is confirmed by the low standard deviation. For example, unscrewing a nut using a plier is 88% longer than using a screw driver ($\pm$8.2% in 2σ).

4. Conclusion

The paper presents a method (called *Dis-DM* and based on CRISP-DM) to sort and cluster big data related to disassembly time and operations from different industry sources. The work illustrates how adapt the six steps expected by the CRISP-DM approach for defining basic values to be used for product disassemblability evaluation, in accordance to the method and software tool presented in (Mandolini et al., 2018) and (Favi et al., 2019). The method aims to establish, for a certain set of liaisons, corrective factors, with relative values, to be considered for estimating disassembly time of a product, considering its conditions during disassembly. This is a systematic procedure where the most relevant statistical algorithms (ANOVA, k-means, distribution fitting, etc.) are exploited for discovering disassembly knowledge from wide data sources, build according methods and tools expected by the Industry 4.0 paradigm.

The proposed method was tested using a test bench equipped in a lab. Multiple nuts, with different diameter, head type, length, were disassembled using multiple unscrewing tools. Measured disassembly times were collected and managed according to the *Dis-DM* approach, with the goal of defining corrective factors to be used by a software tool for selective disassembly time evaluation.

The case study is a preliminary evaluation of the method, since a larger set of data and heterogeneous sources must be considered for stressing the proposed approach, while increasing algorithms to be used for data modelling. Indeed, authors expect to involve dismantlers and maintainers, with the aim to collect a large amount of real data from industrial sources. Further research should be also orientated in defining how to equip current dismantling centres with IoT sensors, so that disassembly times can be retrieved automatically, without deriving them from video recordings and direct observation.

CRediT authorship contribution statement

Claudio Favi: Conceptualization, Writing - review & editing. **Marco Marconi:** Methodology, Validation. **Marco Mandolini:** Formal analysis, Investigation, Data curation, Writing - original draft. **Michele Germani:** Supervision.

References

BMBF-Internetredaktion (21 January 2016). "Zukunftsprojekt Industrie 4.0 - BMBF". Bmbf.de. Retrieved 14 September 2019.
Zhong, R.Y., Xu, X., Klotz, E., Newman, S.T., 2017. Intelligent manufacturing in the context of industry 4.0: a review. Engineering 3 (5), 616–630.
Reike, D., Vermeulen, W.J.V., Witjes, S., 2018. The circular economy: new or refurbished as ce 3.0? - exploring controversies in the conceptualization of the circular economy through a focus on history and resource value retention options. Resour. Conserv. Recycl. 135, 246–264.
den Hollander, M.C., Bakker, C.A., Hultin, E.J., 2017. Product Design in a Circular Economy: development of a Typology of Key Concepts and Terms. J. Ind. Ecol. 21 (3), 517–525.

Gungor, A., Gupta, S.M., 1998. Disassembly sequence planning for products with defective parts in product recovery. Comput. Ind. Eng. 35 (1–2), 161–164.
Vanegas, P., Peeters, J.R., Cattrysse, D., Tecchio, P., Ardente, F., Mathieux, F., Dewulf, W., Duflou, J.R., 2018. Ease of disassembly of products to support circular economy strategies. Resour. Conserv. Recycl. 135, 323–334.
Mandolini, M., Favi, C., Germani, M., Marconi, M., 2018. Time-based disassembly method: how to assess the best disassembly sequence and time of target components in complex products. Int. J. Adv. Manuf. Technol. 95 (14), 409–430.
Güngör, A., 2006. Evaluation of connection types in design for disassembly (DFD) using analytic network process. Comput. Ind. Eng. 50 (1–2), 35–54.
Duflou, J.R., Seliger, G., Kara, S., Umeda, Y., Ometto, A., Willems, B., 2008. Efficiency and feasibility of product disassembly: a case-based study. CIRP Ann. 57 (2), 583–600.
Zhu, B., Roy, U., 2015. Ontology-based disassembly information system for enhancing disassembly planning and design. Int. J. Adv. Manuf. Technol. 78 (9–12), 1595–1608.
Sodhi, R., Sonnenberg, M., Das, S., 2004. Evaluating the unfastening effort in design for disassembly and serviceability. J. Eng. Des. 15 (1), 69–90.
Favi, C., Germani, M., Mandolini, M., Marconi, M., 2016. Includes Knowledge of Dismantling Centers in the Early Design Phase: a Knowledge-based Design for Disassembly Approach. Proc. CIRP 48, 401–406.
Mourtzis, D., Vlachou, E., Milas, N., 2016. Industrial Big Data as a Result of IoT Adoption in Manufacturing. Procedia CIRP 55, 290–295.
Gokalp, M.O., Kayabay, K., Akyol, M.A., Eren, P.E., Koçyiğit, A., 2016. Big data for industry 4.0: a conceptual framework. In: International Conference on Computational Science and Computational Intelligence, pp. 431–434.
Fayyad, U.M., Piatetsky-Shapiro, G., Smyth, P., Uthurusamy, R., 1996. Advances in Knowledge Discovery and Data Mining. AAAI Press/MIT Press, Cambridge, MA.
Friedman, J.H., 1998. Data mining and statistics: what's the connection? Comput. Sci. Statis..
Kretschmer, R., Pfouga, A., Rulhoff, S., Stjepandić, J., 2017. Knowledge-based design for assembly in agile manufacturing by using data mining methods. Adv. Eng. Inf. 33, 285–299.
Bae, J.K., Jinhwa, K., 2011. Product development with data mining techniques: a case on design of digital camera. Expert Syst. Appl. 38 (8), 9274–9280.
Shearer, C., 2000. The CRISP-DM model: the new blueprint for data mining. J. Data Warehousing 5, 13–22.
Favi, C., Marconi, M., Germani, M., Mandolini, M., 2019. A design for disassembly tool oriented to mechatronic product de-manufacturing and recycling. Adv. Eng. Inf. 39, 62–79.