This is a post-peer-review, pre-copyedit version of an article published in IFIP Advances in Information and Communication Technology book series. The final authenticated version is available online at: https://doi.org/10.1007/978-3-031-16411-8_47

Fostering Circular Manufacturing through the integration of Genetic Algorithm and Process Mining

Federica Acerbi[1[0000-0002-0818-4620]], Adalberto Polenghi[1[0000-0002-3112-1775]], Walter Quadrini[1[0000-0003-0081-2255]], Marco Macchi[1[0000-0003-3078-6051]] and Marco Taisch[1[0000-0003-2016-3571]]

[1] Department of Management, Economics and Industrial Engineering, Politecnico di Milano, Piazza Leonardo da Vinci 32, 20133 Milan, Italy
federica.acerbi@polimi.it; adalberto.polenghi@polimi.it; walter.quadrini@polimi.it; marco-macchi@polimi.it; marco.taisch@polimi.it

Abstract. Recently, the increasing lack of raw materials is forcing the manufacturing sector in revising the internal operating and strategic activities to embrace Circular Economy (CE) principles thus, moving towards Circular Manufacturing (CM). CE principles are pursued during product design, product realisation, as well as product end-of-life. As an enabler of end-of-life CM strategies, disassembling represents the cornerstone to facilitate other ones to take place, as remanufacturing and recycling. Indeed, nowadays, empowering companies in the disassembling process by maintaining high their environmental sustainability performances is essential. Indeed, identifying the best disassembly sequence that is also energy-effective is an open challenge to guarantee a 360° application of CM strategies. Therefore, the objective of this contribution is to develop a framework able to automatically reconstruct the disassembly sequence while minimising the energy consumption. The solution is based on process mining technique, which aims at representing the original process, and genetic algorithm, which is instead in charge of identifying the solution with minimal energy consumption. Once the framework has been developed, its feasibility has been tested first at laboratory scale and then through a simulated case. The proposed framework represents a Proof of Concept that aims at promoting the pursue of CM strategies in the product end-of-life by facilitating the identification of the disassembly sequence which is also energy-effective.

Keywords: Circular Economy, Disassembly, Energy, Process Mining, Genetic Algorithm.

1 Introduction

The linear production model is no longer sustainable since it is rapidly and irreversibly damaging the environment requiring drastic corrective actions [1]. Indeed, hundreds of millions of wastes are generated yearly by developed countries together with an increased energy consumption trend causing carbon dioxide and other greenhouse gasses emissions augment [2]. Moreover, the shortage of resources requires the introduction of new strategies oriented towards the adoption of Circular Economy (CE) paradigm in manufacturing companies, named as Circular Manufacturing (CM) strategies [3]. In this context, although the design phase represents the main area in which applying the CE paradigm in an efficient and effective way, it is extremely important to ensure an easy and sustainable treatment of products at their end-of life (EoL) in case they wouldn’t be designed to be circular [4]. EoL-related strategies, such a recycling, remanufacturing and repair, are based on disassembly processes (DP) so to have unitary components and subassemblies available to obtain valuable parts or materials to increase economic efficiency [5].

Nevertheless, although the opportunities are many, there are still some limitations in the correct introduction and exploitation of DP, due to the uncertainty about the disassembly process itself, EoL product quality, demand and yield [4] and high skilled labour requirement to plan the entire DP [6].

In addition, to be effective, thus using few resources in terms of energy, cost and time [7], DP requires lots of data (such as component shape, size and weight, joining elements) [8] that must be available to the company in charge of disassembly.

The challenge this work aims at facing is the exploitation of advanced data analysis techniques to solve the disassembly sequence problem (DSP) in case of few data available, while keeping energy consumption minimisation as a target. The DSP repress the problem of selecting the best sequence of tasks so to minimise or maximise a certain objective function [9]. In the current work, the proposal is a framework based on process mining (PM) to automatically extract the assembly sequence to infer the disassembly one, and genetic algorithms (GA) to identify the DP whose energy consumption is the lowest amongst those DP that are feasible.

2 Background on DSP, process mining and genetic algorithm

As previously described, the DSP refers to the identification of the best disassembly sequence so to optimise a certain objective function. Hence, DSP challenges mainly refer to the identification of the feasible (set of) sequences and then the identification of the sequence whose performance optimise a certain objective. While GA are widely used to solve the DSP [10], they are usually used for both challenges. However, if the knowledge of the process is not given for granted, a workaround must be identified, and PM is identified as a well-fitted option. The overall framework is presented later in section 3, after a brief excursus on PM and GA that are the ground of this work.

2.1 Process Mining

Process Mining (PM) is a process management and data analysis technique which allows to discover, monitor and improve production processes through the inspection of data coming from the information systems present in manufacturing companies (Van Der Aalst, 2012). Benefits like reduced complexity at low effort required from the PM usage, had been evidenced in other fields of study. Indeed, the highlighted PM potentialities are limited time needed for the initial scoping of the data, the PM fast development of an actual process model, the PM capability of mapping complex processes into simple process models and last the PM possibility to analyse and evaluate deviations from the planned process [12].

2.2 Genetic Algorithm

Genetic algorithms are meta-heuristic algorithms based on principles of biological evolution through crossover, mutation, replication, cut-paste and break-joint mechanisms. Species face a process of continuous evolution, generation after generation, intending to achieve better adaptable offspring concerning environmental conditions in which they live. The chromosomes produce offspring chromosomes in such a way that offspring are different but keeping some of the original gene fragments. Indeed, the variation between the two generations is caused by the mutation mechanism that helps to increase chromosome diversity allowing to reach new possibilities [13]. Chromosome's adaptiveness is measured by a fitness value, which is evaluated through a fitness function: an absolute value that represents the goodness of the solution proposed by the chromosome itself [14]. In DSP, GA are widely used since they are able to identify the best disassembly sequence according to specific objectives the company is pursuing.

3 Framework development

The developed framework is based on PM and GA, which allows to determine the best energy-effective disassembly sequence with the aim of enabling EoL-related CM strategies. The framework has been developed according to three phases: i) analysis of the scientific literature to assess the selected algorithms and techniques and to mathematically define the DSP in terms of objective functions, variables and constraints; ii) framework design and development taking into account required data for each step; iii) framework testing both in a laboratory-scaled environment and through a simulated case.

The literature review was systematic through the following keywords ("Disassembly") AND ("Process Mining" OR "Data Mining" OR "Genetic Algorithm"), which, inserted in Scopus, resulted into 318 documents in the last 10 years. After title and abstract screening, and full reading, a final sample of 59 documents were selected as eligible. In addition to GA, also Artificial Bee Colony (6%) and Grey Wolf Optimizer (4%) have been used.

The optimization algorithm used for DSP usually considers as fitness function the cost (39%) and the completion time (34%). Energy-related objectives account for 24%.

The development of algorithms for DSP must consider various constraints. Worth to mention are sequential disassembly constraints (no parallel operations) and precedence constrains (some operations must be performed before other ones).

With a throughout view of the scientific state of art, it is possible to develop a framework, whose aim is to identify the energy-effective sequence and automatically identify the disassembly sequence through process mining. Indeed, this technique is used very few for DSP according to the literature. The framework is depicted in Fig. 1.

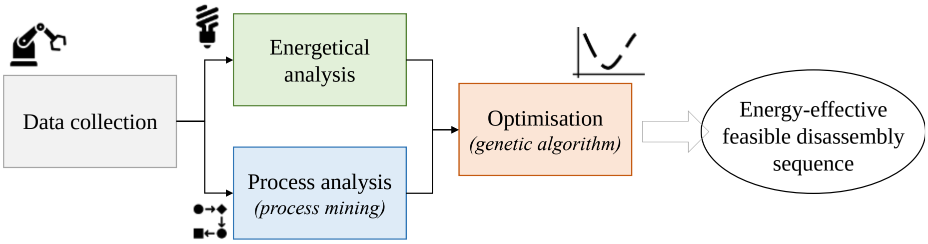

Fig. 1. Framework proposal for energy-effective DSP.

The steps of the framework are the following, going from left to right:

1. Data collection, which refers to different types of data, as operational one for energy analysis and log data for process representation.
2. Energetical analysis to identify hotspots.
3. Process analysis to identify the process sequence through process mining, to be later transformed into disassembly sequence.
4. Optimisation through genetic algorithm for the identification of the cost-effective feasible disassembly sequence.

Starting from the entire batch of data collected from the production systems, those relevant for the DSP problem must be collected. In the scope of this work, the gathered data are related to the consumed energy per machine and the data log recording values of the control variables, usually Boolean, and others like timestamp, the piece ID on which the activity is carried out and the name of the resource (machine) that performs the activity. Once available, the energy data are used to identify hotspots, which are energy demanding machines in production systems. The results are a matrix relating production system resources (machines) with their energy consumption. Furthermore, data logs are used for PM, to infer the disassembly process sequence without a direct knowledge. Indeed, the direct assembly sequence is discovered and reversed so generate a Petri Net that is then translated into a Precedence Matrix, both representing the dismantling precedence.

Finally, the GA takes as input the energy matrix and the Precedence Matrix and looks for a disassembly sequence that respects the dismantling constraints, i.e. feasible, and that also reduces the overall consumed energy, i.e. energy-effective.

4 Framework testing

The first test sessions have been performed in the laboratory environment of the Industry4.0Lab at Politecnico di Milano. This laboratory is a semi-automated assembly line devoted to the manufacturing of real-like smartphones, equipped with commercial PLCs and connected to a digital infrastructure compliant with the standards ISA 95 and IEC 62264. The line is composed of seven stations, realising one operation each, as reported in Fig. 2, together with the product parts. In a previous work by [15], the line has been demonstrated to be capable of some disassembling operations too. For this reason, it has been used as first testing environment for this research, although the authors are aware about key limitations currently present: the disassembly process does not include the transportations and the disassembly phases considered for this experiment are based on the simple reverse flow of the assembly phases. Specifically for this Proof of Concept (PoC), the last key limitation implies that the energy of disassembly operations is assumed through corrective coefficients.

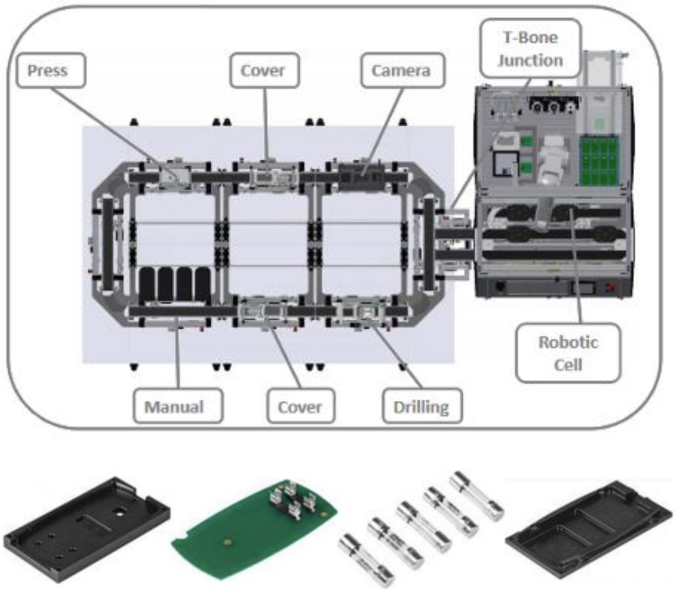

Fig. 2. Industry 4.0 Lab line (rendered).

In the PoC, after an explorative analysis, two variables of interest have been selected for acquisition (one Boolean signal describing the presence of a product in working position and the integer value of the active power consumed), then a chain of Python scripts implementing the methodology described in Section 3 has been developed to feed the GA with the Precedence Matrix and energy matrix. After the phases of generating the initial population, crossing-over, mutation, union of populations and selection of the best chromosome, the best and feasible disassembly sequence in terms of energy consumption have been selected, as displayed in Fig. 3**Error! Reference source not found.**.

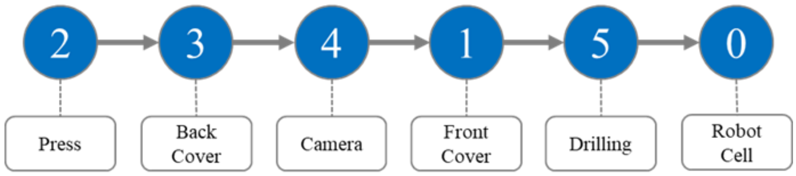

Fig. 3. Identified process sequence for the disassembly process.

Given the inherent limitations of the line (in terms of paths allowed), a second run of the procedure has been performed in a simulated environment, where more products were assumed, hence more paths were allowed (as displayed in the Petri Net of Fig. 4).

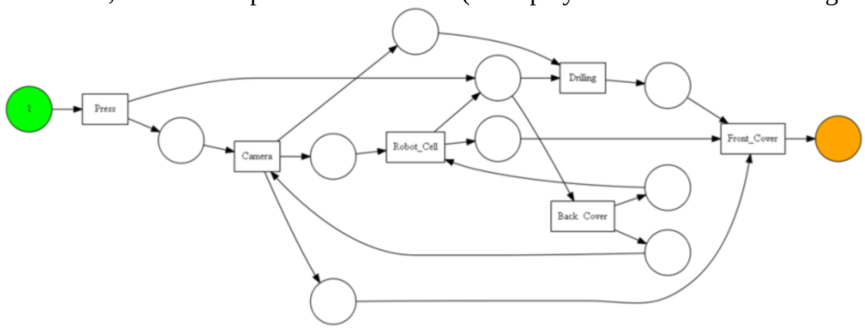

Fig. 4. Petri net of simulated scenario.

The same procedure has been hence run (Fig. 5 depicts the progresses of the GA) and the sequence of Fig. 6 has been found as optimal (unreachable for the real line of Industry4.0Lab, but coherent with the simulated scenario).

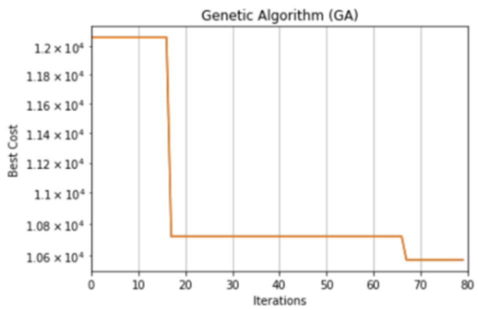

Fig. 5. GA fitness function value by iteration for the simulated scenario.

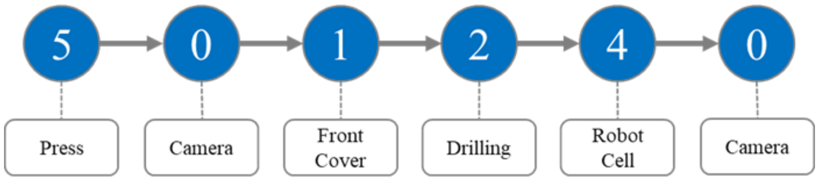

Fig. 6. Identified disassembly sequence for the simulated scenario.

At this moment, this optimal solution suffers of afore mentioned limitations. Especially, through process mining, only the assembly sequence could be discovered, while the disassembly operations should be put manually. This represents the main limitation, target of an ongoing research work, where ontological models may inform about the disassembly operations given the assembly ones.

5 Discussion and conclusions

The present contribution aims at developing and testing a framework to support an energy-effective DSP, as one of the main strategies to pursue circularity in manufacturing companies. The design and development of the framework is based on an ensemble of several techniques and algorithms. Firstly, PM is used to extract the knowledge of the current assembly sequence and reverse it to identify a first disassembly sequence. Then, energy matrix and Precedence Matrix are inputs to the GA that allows to find the disassembling sequence that is feasible and energy-effective at the same time. To test the feasibility of the framework, two applications have been performed. Starting with a successful but simplified case at laboratory scale, the sample of data has been amplified in a simulated scenario to test the robustness. The framework shows its capability to tackle DSP with a complete data-driven approach, given that it does rely only on data gathered from the production systems.

From a theoretical point of view, the proposed framework enables to overcome the limitations identified in the extant literature about the lack of models enabling the automated creation of a precedence tool enabling the identification of the optimal disassembly sequence in a case where no prior knowledge about the dismantling process exists. From a practical point of view the model enables to empower manufacturing companies in putting in practice an optimised disassembly sequence process by receiving it in an automated way the optimal sequence of tasks to be performed to limit the energy consumptions.

Nonetheless, at the current state of advancement, the main limitations to pinpoint are: i) simple disassembly sequence based on the reverse flow of the assembly one, ii) impossibility to practically test the disassembly sequence, iii) one-type product only (more products only simulated, hence more paths), iv) hypothetical energy consumption of the disassembly phases through corrective factors and v) relatively controlled environment in which the framework has been tested. Future research will overcome these limitations, especially related to the possibility to introduce coherent disassembly operations; at the moment ongoing research is focused on the introduction of ontological models that may infer, starting from the process mining information, the disassembly operations required, and the GA could then arrange them in the most energy-effective way. Furthermore, the solution must be extended so to test it in more complex environments, both simulated for numerical experimentation and real to test industrial viability. Additionally, in future research, the authors will apply the framework in an empirical environment, where specific disassembly phases are required and may differ according to the product type.

References

1. J. Korhonen, A. Honkasalo, and J. Seppälä, "Circular Economy: The Concept and its Limitations," *Ecological Economics*, vol. 143, pp. 37–46, Jan. 2018, doi: 10.1016/j.ecolecon.2017.06.041.
2. OECD, *Global Material Resources Outlook to 2060*. OECD, 2019. doi: 10.1787/9789264307452-en.
3. F. Acerbi and M. Taisch, "A literature review on circular economy adoption in the manufacturing sector," *Journal of Cleaner Production*, p. 123086, Jul. 2020, doi: 10.1016/j.jclepro.2020.123086.
4. I. Slama, O. Ben-Ammar, A. Dolgui, and F. Masmoudi, "Genetic algorithm and Monte Carlo simulation for a stochastic capacitated disassembly lot-sizing problem under random lead times," *Computers & Industrial Engineering*, vol. 159, p. 107468, Sep. 2021, doi: 10.1016/j.cie.2021.107468.
5. K. Wang, X. Li, L. Gao, P. Li, and S. M. Gupta, "A genetic simulated annealing algorithm for parallel partial disassembly line balancing problem," *Applied Soft Computing*, vol. 107, Aug. 2021, doi: 10.1016/j.asoc.2021.107404.
6. Institute of Electrical and Electronics Engineers, *An Economical Approach for Disassembly Sequence Planning*. 2019.
7. M. Marconi, M. Germani, M. Mandolini, and C. Favi, "Applying data mining technique to disassembly sequence planning: a method to assess effective disassembly time of industrial products," *International Journal of Production Research*, vol. 57, no. 2, pp. 599–623, Jan. 2019, doi: 10.1080/00207543.2018.1472404.
8. X. C. Zhuang, Y. Wu, J. Gu, and C. X. Li, "Disassembly process management based on BOM," *Liaoning Gongcheng Jishu Daxue Xuebao (Ziran Kexue Ban)/Journal of Liaoning Technical University (Natural Science Edition)*, vol. 25, no. 2, pp. 268–271, Apr. 2006.
9. H. E. Tseng, Y. M. Huang, C. C. Chang, and S. C. Lee, "Disassembly sequence planning using a Flatworm algorithm," *Journal of Manufacturing Systems*, vol. 57, pp. 416–428, Oct. 2020, doi: 10.1016/j.jmsy.2020.10.014.
10. H. E. Tseng, C. C. Chang, S. C. Lee, and Y. M. Huang, "A Block-based genetic algorithm for disassembly sequence planning," *Expert Systems with Applications*, vol. 96, pp. 492–505, Apr. 2018, doi: 10.1016/j.eswa.2017.11.004.
11. W. van der Aalst, "Process mining: Overview and opportunities," *ACM Transactions on Management Information Systems*, vol. 3, no. 2. Jul. 2012. doi: 10.1145/2229156.2229157.
12. C. Ortmeier, N. Henningsen, A. Langer, A. Reiswich, A. Karl, and C. Herrmann, "Framework for the integration of Process Mining into Life Cycle Assessment," in *Procedia CIRP*, 2021, vol. 98, pp. 163–168. doi: 10.1016/j.procir.2021.01.024.

13. H. E. Tseng, Y. M. Huang, C. C. Chang, and S. C. Lee, “Disassembly sequence planning using a Flatworm algorithm,” *Journal of Manufacturing Systems*, vol. 57, pp. 416–428, Oct. 2020, doi: 10.1016/j.jmsy.2020.10.014.
14. Y. Ren, C. Zhang, F. Zhao, H. Xiao, and G. Tian, “An asynchronous parallel disassembly planning based on genetic algorithm,” *European Journal of Operational Research*, vol. 269, no. 2, pp. 647–660, Sep. 2018, doi: 10.1016/j.ejor.2018.01.055.
15. R. Rocca, P. Rosa, C. Sassanelli, L. Fumagalli, and S. Terzi, “Integrating virtual reality and digital twin in circular economy practices: A laboratory application case,” *Sustainability (Switzerland)*, vol. 12, no. 6, Mar. 2020, doi: 10.3390/su12062286.