

# Online Scheduling: A Survey

Rita Guerreiro, André S. Santos

Interdisciplinary Studies Research Center (ISRC), Institute  
of Engineering, Polytechnic Institute of Porto  
Porto, Portugal  
1180543@isep.ipp.pt; abg@isep.ipp.pt

Anabela Tereso  
ALGORITMI Research Centre/LASI  
University of Minho  
Guimarães, Portugal  
anabelat@dps.uminho.pt

*Abstract* — In this article a deep search of the literature of online scheduling is conducted. This paper intends to assess the developments and solutions found for online scheduling problems. Online scheduling is a very important topic since most of the real scheduling problems have dynamic characteristics. First, it was developed a literature review about scheduling problems, dividing them in stochastic and deterministic problems as well as in online and offline problems. Then, a bibliometric analysis was performed. Finally, some case studies in the field of online scheduling were analyzed. Online Scheduling is mostly explored in industry and health areas. In some articles explored there is a rescheduling, and the sequence of task may change due to the arrival of new tasks. In other cases, the new tasks are introduced in blocks of time that do not affect the previous schedule. This last technique is limited, since, with the arrival of new tasks, the schedule is not re-evaluated. Therefore, it is thought that, in future work, within the scope of online scheduling, when new tasks or other significant changes enter the system, the system should be evaluated, allowing the necessary changes to be made to the existing schedule. The Industry 4.0 and the evolution of Internet of Things (IoT), Deep Learning and Machine Learning favours a continuous and real-time flow of information, which allows the implementation of real-time online scheduling. This is a branch that should be explored in future works.

*Keywords* - online scheduling; heuristics; internet of things, machine learning.

## I. INTRODUCTION

In this paper, the literature review of the field of online scheduling has been explored in order to discover recent developments and solutions for scheduling problems with dynamic components. Most of the scheduling literature considers static and deterministic systems. However, most real-world scheduling problems are distributed, dynamic and stochastic in nature. In this way, it is important to explore the online field of the scheduling problems in order to find better ways to optimize these problems. After this introduction, the second chapter will present a literature review on scheduling problems and also a bibliometric analysis on online scheduling problems. Then, in the third chapter, some case studies in the online scheduling field are presented. Finally, in the fourth chapter some conclusions are drawn.

## II. LITERATURE REVIEW

### A. Scheduling Problems

In both productive industries and in services, there are scarce resources, which are used in the development of the company's activities. Often managers are faced with uncertainties about the best way to allocate resources to tasks (or vice versa) and about which will be the schedule of tasks that optimizes a given performance measure. As described by [1], "Scheduling is a field in which there are some intriguing problems and some interesting answers". Scheduling problems began to arouse the interest of researchers around 1958. Scheduling is a decision-making process that distributes available resources between operations, during a given period. Scheduling can include allocation, sequencing and scheduling decisions. Through correct scheduling it is possible to optimise one, or more, performance measures. Scheduling can be applied in many different areas, but it is usually easier to apply scheduling in an industrial environment than, for example, in the service area [2]. In [2] it is used the example of an airport, where it is necessary to reserve boarding gates for specific aircraft or the reservation of rooms or other infrastructures for service companies, making the scheduling process more complex.

### B. Stochastic vs. Deterministic Scheduling

Scheduling can be divided into deterministic and stochastic: deterministic scheduling simplifies real problems that usually incorporate randomness, by assuming that all the parameters of the problem are known and can be described through deterministic values; stochastic scheduling are problems where the parameters of the problem are not all known with certainty and are typically approximated through statistical distributions that model that uncertainty. To characterize a stochastic scheduling problem, one needs information about the distribution of processing times, release dates, etc., and it is also necessary to know if the processing times of the tasks are independent or correlated and which classes to consider [2]. It is possible to resort to fuzzy models to incorporate uncertainty and subjectivity into the problem. Fuzzy scheduling serves mainly for scheduling with flexible constraints or scheduling with imprecise or missing information and one of its main advantages is the focus on the significant scheduling decisions [3]–[5].

### C. Online vs. Offline scheduling

Scheduling problems can also be classified as either dynamic (online) or static (offline), depending on the nature of arrival of the tasks. In a static scheduling problem, all the tasks, as well as their processing times or posting dates are available at the start. In an online scheduling problem, it is not known how many tasks exist nor what their processing times are. The decision maker only has the information about the existing tasks when they are posted. The algorithm must make decisions with limited information, and it is not known how many more tasks will be launched, nor when they will be launched [2], [6], [7]. In online scheduling the process is continuous, tasks arrive intermittently with no set arrival times. Online scheduling is very important since in the real-world information is often limited when a decision needs to be made. Stochastic scheduling also has to make decisions with limited information, however these two concepts are different. In stochastic scheduling, there is at least some information regarding the distribution of processing times and the launch and finish dates of tasks; in the case of online scheduling this information is not available [2]. There is a stochastic component in an online scheduling model, however a stochastic model may not be online. Generally, an online model is always stochastic, since it deals with randomness in the number of tasks of the problem, but a stochastic model may not be online if it deals only with randomness in the parameters of a given number of tasks, known a priori.

### D. Bibliometric Analysis in Online Scheduling

To carry out the bibliometric analysis, the Dimensions software [8] was used to collect data concerning the

bibliography related to the topic of online scheduling. Subsequently, these data were treated using the bibliometric library of the R studio software [9]. Effectively, it was possible to observe that as of the 2000s there was an accelerated growth of research within the scope of online scheduling, as can be seen by analyzing Fig. 1.

As can be seen in Fig. 2, the area of study with more publications is the information and computing sciences, followed by the area of engineering. It is also important to mention that clinical sciences and health sciences have a significant amount of publications in this area, mainly through research in the field of nurse scheduling, allocating them to patients.

In Fig. 3 it is possible to observe the keywords most used by the authors of the articles about online scheduling. In fact, the most used word is "algorithms" followed by "humans". These keywords are of great relevance, since algorithms are often used to solve scheduling problems and, generally, these schedules aim at allocating people (humans) to different tasks. The expression "reproducibility of results" should also be highlighted. Due to the complexity and unique characteristics of each problem, authors usually opt to study solutions for a specific problem, and the solution found, many times, does not apply to any other problem. For a fast and effective study of scheduling problems it is necessary to find solutions, having in mind their reproducibility, allowing to adapt the developed algorithms to different problems. The word "heuristic" is a key word for many authors since many of the solutions addressed in the context of online scheduling use heuristics. Within the scope of online scheduling, the expressions "Internet of Things (IoT)", "Artificial Intelligence (AI)", "Machine Learning" and

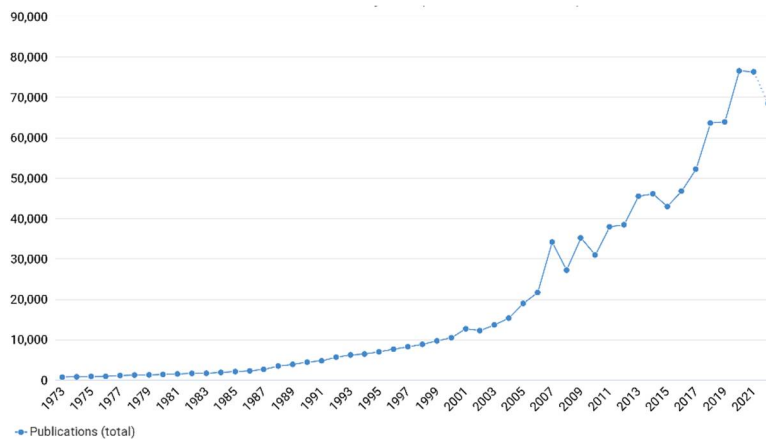


Figure 2. Publications per year in the online scheduling field

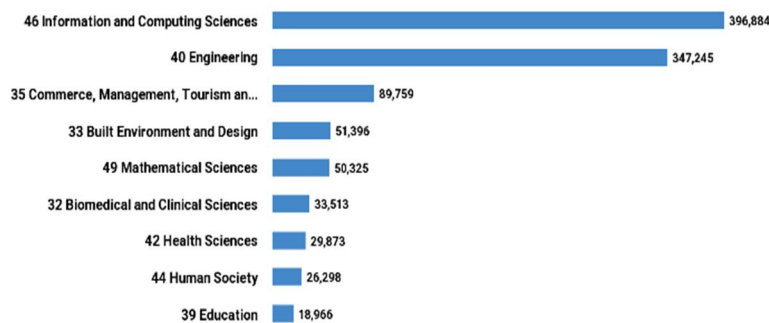


Figure 1 - Number of publications in each research category



for the current week, and SAA allows to evaluate each and retain the best one. The objective of this project is to sequentially build, week by week, the schedule for a set of nurses over a planning horizon of several weeks. In this dynamic process, the calculated scheduling for a given week is fixed before the demand and preferences for the next week are revealed. The primal-dual algorithm for online optimization aims at dynamically constructing pairs of primal and dual solutions. At each stage, primal decisions are made, and the dual solution is updated so that it remains feasible. The current dual solution leads the algorithm to better primal decisions, using these dual values as multipliers in a Lagrangian relaxation. The goal is to obtain a pair of feasible primal and dual solutions that satisfy the complementary relaxation property at the end of the process. Effectively, the solution developed by the authors works very well in a wide range of cases (from 30 to 120 nurses in a four or eight-week horizon). The results show that, although this procedure does not calculate the best scheduling, it finds feasible solutions for almost all examples of the competition, while providing high quality schedules.

#### C. Online Scheduling: Parallel Machine System

There are also many online scheduling publications and investigations in the industry. In [12] is presented the online scheduling of parallel heat treatment furnaces in a manufacturing company. The problem is an extension of the parallel machine scheduling problem. For the search of a solution for the realization of online production scheduling, the authors opt for a constructive, multiple start search approach. The Greedy Randomized Adaptive Search Procedure (GRASP) is used. GRASP will build a solution step by step, through a construction heuristic, determining the next job to be scheduled. The solutions are independent of each other, which means that each GRASP solution is in fact an autonomous search, built from scratch. The model in [12] is connected to the company's ERP and has three modes of operation, manual, periodic or automatic. The tasks arrive to the department dynamically and are introduced in the company's ERP, containing the necessary data. Then, manually, any employee, with access to the ERP, can run the GRASP. Alternatively, one can set a period for the algorithm to run, or it can be set to run automatically whenever a new task is entered into the system. To demonstrate the effectiveness of the proposed approach, the results of the algorithm are compared to other proven techniques. The results obtained demonstrate that efficient solutions can be found from the model proposed by the authors. According to the pilot study, it was observed that by using the model, the total downtime of the ovens was decreased by approximately 3,080 minutes in a month. In short, using the proposed algorithm, the expected annual gain is between 51,374.40 euros and 1,416,492.00 euros.

#### D. Real-time Adaptive Scheduling

In [13], a real-time adaptive scheduling method for logistics and production resources was formulated to solve advanced scheduling problems with fuzzy processing times and random job arrival. This method was used to solve the real problem of collaborative scheduling of logistic and production resources in

a high perturbation environment. Firstly, the real-time capacity model for Production-Logistics resources was built. Then, the triggering mechanism, based on the relationship between the disturbance level and the information utility, was developed. Finally, a real-time task-oriented scheduling approach was presented. The objective function aims to minimize the average task delay, the job completion time, and the total energy consumption of the resources. When a job enters the system, the cloud passes that information to the shop floor. That job is divided into multiple tasks. The information from the tasks being processed is sent to the different machines, which perform a capacity analysis, making a logistics plan that is sent to the Automated Guided Vehicles (AGV). These, in turn, analyze their service capacity, in real-time, and return the results to the respective machines. According to these results, each machine chooses the best AGV, forming groups of resources (allocation of each machine to an AGV), and the service capacity of each group is calculated. This information will be used in the scheduling method, which chooses the best group for a certain task, passing this information to the machine of the chosen group that, in turn, transmits the information to the AGV. The process is repeated for all tasks that follow, and the scheduling is always being updated, in real-time. The results show that this approach can effectively deal with the real-time production in dynamic environment.

#### E. Online Scheduling: Single Machine System

In [14], the authors developed an online scheduling tool for a single-machine system. In this system, tasks are launched dynamically and have stochastic execution times. The developed software alternates between 2 priority rules that meet the objectives, without the need for user interaction. The priority rules define the order (sequence) in which the tasks are executed. It is used a hyper-heuristic that alternates between the Earliest Due Date (EDD) and Shortest Processing Time (SPT). While the maximum delay does not reach the value defined by the user as maximum, the SPT heuristic is used to minimize the flow time, when it is detected that the maximum delay value has been reached, the EDD heuristic is used to minimize it. This algorithm is based on Kano's satisfaction model. The maximum delay represents a mandatory attribute, so if it is not met, it results in extreme customer dissatisfaction. As for the average flow time, it represents a proportional attribute, since it corresponds to a degree of satisfaction proportional to the degree of performance of the attribute, i.e., the shorter the average flow time, the higher the customer satisfaction.

In order to test the proposed framework, 100 tasks were created with stochastic attributes. Overall, the proposed model showed good results, with the maximum delay always being less than zero and the average flow time was minimized whenever possible.

#### F. Other Case Studies in the Online Scheduling Field

In addition to the papers presented above, there are many more that play a key role in the growth of research in the field of online scheduling in the most diversified areas, such as

transportation, space, military, supply-chain and crisis management [15]–[17]. Some of them are presented in Table 1.

TABLE 1. Articles In Online Scheduling

Authors	Summary
[18]	To solve the online and stochastic vehicle routing problem, the authors propose a multiple scenario approach (MSA). The key idea behind MSA is to generate new routing plans for scenarios that include existing as well as possible future requests. Experimental results indicate that MSA yields improvements over approaches that do not use stochastic information.
[19]	In this paper, the problem of scheduling oncology treatments in an online and stochastic way was presented. The goal is to provide a reasonable waiting time for the first treatment while maximizing resource utilization. To represent the uncertainty related to patient arrival and treatment duration, the authors developed an approach that combines stochastic optimization and online optimization. The results show that this works well in real cases and outperforms the current strategy.
[20]	In this study, the authors try to find a solution to the online scheduling problem of multiple control tasks with known execution times. The tasks can be dynamically incorporated into the scheduling through a modification of the optimization problem. In the control experiments, there was the reduction of the total cost by about 30%.
[21]	In this study, the rescheduling problem in job shop environments is investigated, considering that new tasks may be released and that there may be preventive maintenance. To deal with this integrated optimization problem, an algorithm (NSGA-III/ARV) was proposed. The global performance of the proposed algorithm was verified through internal analysis and comparison with other intelligent algorithms.
[22]	In this research, the authors developed robust models for performing online flight scheduling, attempting to increase responsiveness to demand uncertainty. It was proved, through experiments using data from a large US airline, that the generated scheduling can achieve significant profit improvements.
[23]	In this paper, online scheduling in flexible job shop is studied. A multi-objective mathematical model for the problem, based on a genetic algorithm, is developed. The experimental results show that the proposed algorithm can achieve the optimal solutions for small problems and near optimal solutions for the medium dimension problems.
[24]	In this paper, the intelligent manufacturing service scheduling problem is analysed. A method based on deep reinforcement learning is used. Two case studies are used to show the efficiency of the proposed method.
[25]	In this study, an approach is proposed to solve the online scheduling problem that reduces energy consumption in a flow shop. A new algorithm based on the Particle Swarm Optimization (PSO) was developed. The algorithm was tested, concluding that it can obtain exceptional energy savings for small and medium scale problems.
[26]	This paper addresses the online scheduling problem in a flexible job shop system, considering new job releases. To handle continuous production states and learn the most appropriate action (i.e., dispatching rule) at each rescheduling point, a Deep Q Network (DQN) was developed to solve this problem.

As can be seen, the case studies presented in Table 1 cover different fields. There is for example the online scheduling problem in the transport area with the vehicle routing problem; the health area, with the scheduling of

oncologic treatments; services, with the flight schedules; and industry. This shows the scope and importance of research in the area of online scheduling.

#### IV. CONCLUSIONS

The main theme of this paper is the online scheduling problem. Initially, the basics of the scheduling problems were explored. Next, scheduling strategies were discussed, defining deterministic scheduling and stochastic scheduling. The offline and online scheduling were also defined. In the former, it is assumed that all tasks are known at the moment of scheduling, while the latter is a continuous process in which the tasks appear stochastically. Finally, case studies related to online scheduling were analysed. First, a bibliometric analysis was performed that allowed analysing that the theme of online scheduling has gained more importance and attention in the 2000s. The areas where there has been greater development of this theme are the industry and the health area for the online scheduling of nurses. In most of the articles, approximate techniques and in some cases, hyper-heuristics are also used successfully. Sometimes, hybrid algorithms are used, combining exact and approximate techniques. In order to act on the online aspect of the problems, in many cases, algorithms have been developed that run periodically or whenever new tasks arrive, so that they can be included into the schedule. In some articles there is a rescheduling, and the sequence of task may change due to the arrival of new tasks. In other cases, the new tasks are introduced in blocks of time that do not affect the previous schedule. This last technique is limited, since, with the arrival of new tasks, the schedule is not re-evaluated. For example, if the task that arrives is urgent, with this methodology its importance will not be immediately transmitted to the system, so it will be placed in the next free time block and not in a priority position. In this sense, it is thought that, in future work, within the scope of online scheduling, when new tasks or other significant changes enter the system, the system should be evaluated, allowing the necessary changes to be made to the existing schedule. In another sense, it is believed that with the evolution of the implementation of Industry 4.0 it is possible to increase the number of case studies of online scheduling. With Industry 4.0 and the evolution of IoT, Deep Learning and Machine Learning concepts it is increasingly possible to interconnect the various elements of the system, with a continuous and real-time flow of information, which favours the implementation of real-time online scheduling. This is a branch that should be explored in future works. This article is an initial research for a master's thesis on online scheduling. Next, more specific research will be conducted and a tool to support online scheduling in the context of an organization will be developed.

#### ACKNOWLEDGMENT

This work has been supported by *FCT – Fundação para a Ciência e Tecnologia* within the R&D Units Project Scope: UIDB/00319/2020.

## REFERENCES

- [1] R. W. Conway, W. L. Maxwell, and L. W. Miller, *Theory of scheduling*, 1st ed. Boston: Addison Wesley, 1967.
- [2] M. L. Pinedo, *Scheduling Theory, Algorithms and Systems*, 5th ed. New York: Springer, 2012. doi: 10.1007/978-3-319-26580-3.
- [3] J. Behnamian, "Survey on fuzzy shop scheduling," *Fuzzy Optimization and Decision Making*, vol. 15, no. 3, pp. 331–366, Sep. 2016, doi: 10.1007/s10700-015-9225-5.
- [4] D. Dubois, H. Fargier, and P. Fortemps, "Fuzzy scheduling: Modelling flexible constraints vs. coping with incomplete knowledge," *Eur J Oper Res*, vol. 147, no. 2, pp. 231–252, 2003.
- [5] A. L. Guiffrida and R. Nagi, "Fuzzy set theory applications in production management research: a literature survey," *J Intell Manuf*, vol. 9, pp. 39–56, 1998.
- [6] K. Baker and D. Trietsch, *Principles of Sequencing and Scheduling*, 2<sup>o</sup>. John Wiley & Sons, Inc. Edition, 2019. Accessed: Dec. 07, 2022. [Online]. Available: <https://www.ptonline.com/articles/how-to-get-better-mfi-results>
- [7] K. Lee, F. Zheng, and M. L. Pinedo, "Online scheduling of ordered flow shops," *Eur J Oper Res*, vol. 272, no. 1, pp. 50–60, 2019, doi: 10.1016/j.ejor.2018.06.008.
- [8] "Software Dimensions ." 2022. Accessed: Dec. 27, 2022. [Online]. Available: <https://www.dimensions.ai/>
- [9] "Software R Studio." 2022. Accessed: Dec. 27, 2022. [Online]. Available: <https://posit.co/>
- [10] W. J. Gutjahr and M. S. Rauner, "An ACO algorithm for a dynamic regional nurse-scheduling problem in Austria," *Comput Oper Res*, vol. 34, no. 3, pp. 642–666, 2007.
- [11] A. Legrain, J. Omer, and S. Rosat, "An online stochastic algorithm for a dynamic nurse scheduling problem," Elsevier B.V., 2020. doi: 10.1016/j.ejor.2018.09.027.
- [12] A. Baykasoglu and F. B. Ozsoydan, "Dynamic scheduling of parallel heat treatment furnaces: A case study at a manufacturing system," *J Manuf Syst*, vol. 46, pp. 152–162, Jan. 2018, doi: 10.1016/j.jmsy.2017.12.005.
- [13] W. Yang, W. Li, Y. Cao, Y. Luo, and L. He, "An information theory inspired real-time self-adaptive scheduling for production-logistics resources: Framework, principle, and implementation," *Sensors (Switzerland)*, vol. 20, no. 24, pp. 1–21, Dec. 2020, doi: 10.3390/s20247007.
- [14] L. Ferreira et al., "A dynamic selection of dispatching rules based on the Kano model satisfaction scheduling tool," in *Innovation, Engineering and Entrepreneurship*, Springer, 2011, pp. 339–346.
- [15] T. Le-Anh, R. B. M. De Koster, and Y. Yu, "Performance evaluation of dynamic scheduling approaches in vehicle-based internal transport systems," *Int J Prod Res*, vol. 48, no. 24, pp. 7219–7242, Dec. 2010, doi: 10.1080/00207540903443279.
- [16] J. C. Pemberton and L. G. Greenwald, "On the need for dynamic scheduling of imaging satellites," *International Archives of Photogrammetry Remote Sensing and Spatial Information Sciences*, vol. 34, no. 1, pp. 165–171, 2002.
- [17] X. Li and B. I. Epureanu, "Intelligent agent-based dynamic scheduling for military modular vehicle fleets," in *IIE Annual Conference. Proceedings*, Institute of Industrial and Systems Engineers (IISE), 2017, pp. 404–409.
- [18] W. B. Russell and V. H. Pascal, "Scenario-based planning for partially dynamic vehicle routing with stochastic customers," *Oper Res*, vol. 52, no. 6, pp. 977–987, Nov. 2004, doi: 10.1287/opre.1040.0124.
- [19] A. Legrain, M. A. Fortin, N. Lahrichi, and L. M. Rousseau, "Online stochastic optimization of radiotherapy patient scheduling," *Health Care Manag Sci*, vol. 18, no. 2, pp. 110–123, Jun. 2015, doi: 10.1007/s10729-014-9270-6.
- [20] A. Cervin and P. Alriksson, "Optimal on-line scheduling of multiple control tasks: A case study," in *Proceedings - Euromicro Conference on Real-Time Systems*, 2006, pp. 141–150. doi: 10.1109/ECRTS.2006.18.
- [21] Y. An, X. Chen, K. Gao, Y. Li, and L. Zhang, "Multiobjective Flexible Job-Shop Rescheduling with New Job Insertion and Machine Preventive Maintenance," *IEEE Trans Cybern*, 2022, doi: 10.1109/TCYB.2022.3151855.
- [22] H. Jiang and C. Barnhart, "Robust airline schedule design in a dynamic scheduling environment," *Comput Oper Res*, vol. 40, no. 3, pp. 831–840, Mar. 2013, doi: 10.1016/j.cor.2011.06.018.
- [23] P. Fattahi and A. Fallahi, "Dynamic scheduling in flexible job shop systems by considering simultaneously efficiency and stability," *CIRP J Manuf Sci Technol*, vol. 2, no. 2, pp. 114–123, 2010, doi: 10.1016/j.cirpj.2009.10.001.
- [24] L. Zhou, L. Zhang, and B. K. P. Horn, "Deep reinforcement learning-based dynamic scheduling in smart manufacturing," in *Procedia CIRP*, Elsevier B.V., 2020, pp. 383–388. doi: 10.1016/j.procir.2020.05.163.
- [25] D. Tang, M. Dai, M. A. Salido, and A. Giret, "Energy-efficient dynamic scheduling for a flexible flow shop using an improved particle swarm optimization," *Comput Ind*, vol. 81, pp. 82–95, Sep. 2016, doi: 10.1016/j.compind.2015.10.001.
- [26] S. Luo, "Dynamic scheduling for flexible job shop with new job insertions by deep reinforcement learning," *Applied Soft Computing Journal*, vol. 91, Jun. 2020, doi: 10.1016/j.asoc.2020.106208.