Proactive Data-driven Scheduling of Business Processes

Francesca Meneghello[1,2,3], Arik Senderovich[4], Massimiliano Ronzani[3], Chiara Di Francescomarino[5] and Chiara Ghidini[2]

[1]*Sapienza University of Rome, Via Ariosto, 25, 00185, Rome, Italy*
[2] *Free University of Bozen-Bolzano, via Bruno Buozzi, 1 - 39100, Bozen-Bolzano, Italy*
[3]*Fondazione Bruno Kessler, Via Sommarive, 18, POVO - 38123, Trento, Italy*
[4]*York University, 4700 Keele St, Toronto, ON M3J 1P3, Canada*
[5]*University of Trento, Via Sommarive, 9, 38123 Trento, Italy*

Abstract

Business processes are structured sets of activities that organizations execute to achieve various objectives, such as completing a sale, providing a service, or managing a supply chain [1]. Unlike processes in manufacturing facilities, which typically involve predictable durations, and stable resource availability, business processes exhibit significant uncertainty and variability [2, 3]. Specifically, scheduling business processes is notoriously complex due to the stochastic nature of activity durations caused by human behavior, as well as the time-varying availability, overlapping responsibilities, and vacations. These challenges make traditional deterministic scheduling methods inadequate [4].

In response to this limitation, proactive scheduling techniques generate robust offline schedules [5, 6]. These methods explicitly account for the stochastic nature of activity durations and aim to compute a *proactive optimal solution* that achieves a predefined confidence level [5]. For example, a solution with a confidence level of $1 - \alpha$ ensures that the returned Makespan remains below an optimal threshold in at least $(1 - \alpha)$ of cases. While proactive scheduling has made significant progress in addressing variability through probabilistic activity durations [7, 8, 9, 10, 5], critical gaps remain. Notably, the challenge of *planned* resource unavailability is rarely addressed. Although studies incorporate the planned unavailability of resources, some overlook the uncertainty of duration [11], while others fail to handle the complexity and variability typical of business processes [12, 13].

Beyond methodological gaps, many real-world business processes face an additional practical challenge: the lack of readily available data on key scheduling parameters. Information such as activity durations, their probabilistic distributions, and resource availability calendars is often incomplete or missing. Without access to this data, even advanced scheduling techniques struggle to perform effectively in real-world settings. To bridge this gap, we employ *process mining*, a data-driven approach designed to extract knowledge and insights from event logs [14].

Building on this foundation, we formulate the Business Process Scheduling Problem (BPSP) and establish its relationship to the well-known Resource-Constrained Project Scheduling Problems (RCPSP). Our approach leverages process mining and proactive scheduling methods to tackle both the inherent uncertainty in business processes, and lack of knowledge of the problem. Specifically, we develop a framework that combines data-driven simulation [15], and constraint programming [11]: we construct a constraint programming (CP) model that generates deterministic schedules for the BPSP, verify them using the simulator, and select the best performing schedules. Specifically, the schedules returned by the CP model are evaluated using the simulator to identify the one that minimizes the uncertain Makespan while meeting a predefined confidence level. This process ensures robustness against both variability in process execution and resource constraints.

We evaluate our approach in two phases to demonstrate both its effectiveness and applicability. In the first phase, we conduct experiments using synthetic data to demonstrate the adaptability of our solution to varying uncertainty levels, problem sizes, and resource configurations, achieving effective optimization of the probabilistic Makespan. This controlled environment ensures that the method can adapt to diverse scheduling scenarios and reliably optimize outcomes under uncertainty. In the second phase, we validate the applicability of our approach by applying it to real-world event logs from a healthcare process, achieving an optimization of the process Makespan by an average of 5% to 14%. Using process mining techniques, we extract BPSP parameters from the logs and employ our framework to generate optimal schedules. This highlights the practical value of our method, showcasing its ability to address complex, data-driven scheduling challenges in a real-world domain.

Keywords

Business Process Scheduling Problem, Constraint Programming, Business Process Simulation

francesca.meneghello@uniroma1.it (F. Meneghello); sariks@yorku.ca (A. Senderovich); mronzani@fbk.eu (M. Ronzani); c.difrancescomarino@unitn.it (C. D. Francescomarino); chiara.ghidini@unibz.it (C. Ghidini)

References

[1] M. Dumas, L. M. Rosa, J. Mendling, A. H. Reijers, Fundamentals of business process management, Springer, 2018.

[2] M. Shoush, M. Dumas, When to intervene? prescriptive process monitoring under uncertainty and resource constraints, in: International Conference on Business Process Management, Springer, 2022, pp. 207–223.

[3] J. Xu, C. Liu, X. Zhao, S. Yongchareon, Z. Ding, Resource management for business process scheduling in the presence of availability constraints, ACM Transactions on Management Information Systems (TMIS) 7 (2016) 1–26.

[4] M. L. Pinedo, Scheduling, volume 29, Springer, 2012.

[5] J. C. Beck, N. Wilson, Proactive algorithms for job shop scheduling with probabilistic durations, Journal of Artificial Intelligence Research 28 (2007) 183–232.

[6] T. Chaari, S. Chaabane, N. Aissani, D. Trentesaux, Scheduling under uncertainty: Survey and research directions, in: 2014 International conference on advanced logistics and transport (ICALT), IEEE, 2014, pp. 229–234.

[7] U. Satic, P. Jacko, C. Kirkbride, Performance evaluation of scheduling policies for the dynamic and stochastic resource-constrained multi-project scheduling problem, International Journal of Production Research 60 (2022) 1411–1423.

[8] D. Liu, Z. Xu, A multi-pr heuristic for distributed multi-project scheduling with uncertain duration, IEEE Access 8 (2020) 227780–227792.

[9] V. A. Hauder, A. Beham, S. Raggl, S. N. Parragh, M. Affenzeller, Resource-constrained multi-project scheduling with activity and time flexibility, Computers & Industrial Engineering 150 (2020) 106857.

[10] H. Chen, G. Ding, J. Zhang, S. Qin, Research on priority rules for the stochastic resource constrained multi-project scheduling problem with new project arrival, Computers & Industrial Engineering 137 (2019) 106060.

[11] S. Kreter, A. Schutt, P. J. Stuckey, J. Zimmermann, Mixed-integer linear programming and constraint programming formulations for solving resource availability cost problems, European Journal of Operational Research 266 (2018) 472–486.

[12] D. Yang, X. Zhou, Z. Yang, Q. Jiang, W. Feng, Multi-objective optimization model for flexible job shop scheduling problem considering transportation constraints: A comparative study, in: 2020 IEEE Congress on Evolutionary Computation (CEC), IEEE, 2020, pp. 1–8.

[13] P. Winklehner, V. A. Hauder, Flexible job-shop scheduling with release dates, deadlines and sequence dependent setup times: A real-world case, Procedia Computer Science 200 (2022) 1654–1663.

[14] W. van der Aalst, Process Mining: Data Science in Action, 2nd ed., Springer Publishing Company, Incorporated, 2016.

[15] F. Meneghello, C. D. Francescomarino, C. Ghidini, M. Ronzani, Runtime integration of machine learning and simulation for business processes: Time and decision mining predictions, Inf. Syst. 128 (2025) 102472. URL: https://doi.org/10.1016/j.is.2024.102472. doi:10.1016/J.IS.2024.102472.