Proceedings of the 2022 Winter Simulation Conference
B. Feng, G. Pedrielli, Y. Peng, S. Shashaani, E. Song, C.G. Corlu, L.H. Lee, E.P. Chew, T. Roeder, and P. Lendermann, eds.

SIMULATIONS FOR OPTIMIZING DISPATCHING STRATEGIES IN SEMICONDUCTOR FABS USING MACHINE LEARNING TECHNIQUES

Benjamin Kovács

Alpen-Adria-Universität Klagenfurt
Universitätsstraße 65-67
Klagenfurt am Wörthersee 9020, AUSTRIA

ABSTRACT

Optimizing operations of semiconductor manufacturing plants is a tremendous challenge due to the complexity and scale of real-world problem instances. Simulations are widely used for prototyping, evaluating, comparing, and verifying new control strategies, reducing the costs, risks, and time required for the development. We present an open-source, customizable discrete-event simulator tool for the semiconductor industry, simulating real-world-scale problems based on open datasets incorporating the challenging characteristics and constraints of the process. The simulator provides a general, customizable interface to allow benchmarking of various methods. A reinforcement learning environment is also bundled with the toolbox. Using our software, we develop evolutionary algorithms and reinforcement learning-based dispatching strategies and compare them to heuristics widely adapted in the industry.

1 INTRODUCTION

Wafer fabrication is one of the most complex manufacturing processes concerning several hundreds of machines and thousands of simultaneous lots in the production process. Completing a lot usually requires several hundred steps involving re-entrant flow and dynamic routes due to stochastic sampling steps and rework requirements. Coupling constraints, unplannable machine breakdowns, scheduling maintenance of machines, and sequence-dependent setups make the problem challenging to model and solve with exact planning methods.

Due to the scale and dynamicity of the problem, the prevailing solving methods are various dispatching heuristics with local optimization on some critical parts of the process. Several simulation solutions and datasets have been developed to evaluate dispatching algorithms. However, most large-scale simulator tools are closed-source commercial solutions, making them difficult to apply in research due to licensing issues and limitations of the available interfaces. To provide a performant, customizable, and open alternative to commercial solutions, we introduced *PySCFabSim* (Kovács et al., 2022). The tool is designed for AI researchers and implements the model of the *SMT2020* dataset (Kopp et al., 2020). Additionally, industrial datasets can be integrated with the tool, as we use a more general input format.

To investigate the dispatching strategies' potential, we researched metaheuristics and developed improved algorithms using evolutionary algorithms (EAs) like genetic programming. Additionally, a reinforcement learning (RL) environment has been created to experiment with multi-agent approaches.

2 SIMULATION

We developed our discrete event simulator in Python. The core simulator uses no external libraries, allowing accelerated execution using just-in-time compilation. The implementation is optimized for performance to allow the rapid collection of samples in large quantities required for training and evaluating RL and EA approaches.

The designed tool simulates the challenging characteristics of semiconductor manufacturing processes (Mönch et al., 2011), including batch processing and cascading machines, sequence-dependent setups, stochasticity in machine availability, routes, and machine dedications. Based on development requirements, the mentioned advanced features can be deactivated.

The simulator was validated by implementing the dispatching strategies introduced by the authors of the base dataset. The experiments demonstrated that the key performance indicators match the reference implementation, and simulating a 2-year horizon is finished in a reasonable 20-minute time frame.

We provide extensions for schedule visualization using charts and integrate a web-based experiment monitoring framework for live-tracking the most important performance metrics.

Parallel, large-scale training, and evaluation of policies are enabled by quick instantiation and initialization (< 3 seconds) and a small memory footprint (around 200 MB). The simulator's state can be saved and reloaded to explore the effects of stochasticity in the future.

3 METHODS

Considering industrial requirements, the first phase of the research focuses on improving the dispatching strategies. The improved rules can be efficiently verified and integrated into the existing infrastructure. We improve existing heuristics by augmenting the hierarchical rules with new sorting criteria and developing separate strategies for each workstation or machine. Metaheuristics are an effective way to design novel dispatching strategies automatically. We employ genetic programming to evolve heterogenous strategies for the workstations in our simulated factory. The resulting methods outperform the reference by a high margin on the observed performance indicators.

However, the solution space explorable by this metaheuristic approach is limited. For example, decisions are taken greedily based on local measures making global optimization impossible. Therefore, we investigate additional machine learning techniques with representation learning capabilities.

Recently, RL-based methods achieved several breakthroughs in sequential decision-making problems and are increasingly used to solve combinatorial optimization problems, including scheduling (Shyalika et al., 2020). Therefore, we developed a new, customizable RL environment based on our simulator tool. The framework allows the creation of environments with custom observation space composed of a pre-defined set of features. Developers can also add custom features using data extracted from the core simulator using plugins. The same flexibility is provided for the design of the reward function. We also provide multiple pre-defined action spaces, allowing agents to dispatch by either choosing a machine, a job, or a heuristic.

4 FUTURE WORK

In the future, we aim to improve the simulator by adding more datasets based on real-world factories. Additionally, we plan to extend the RL framework to provide graph observations of the production process, allowing the development of policies based on graph neural networks.

REFERENCES

B. Kovács, P. Tassel, R. Ali, M. El-Kholany, M. Gebser and G. Seidel, "A Customizable Simulator for Artificial Intelligence Research to Schedule Semiconductor Fabs," 2022 33rd Annual SEMI Advanced Semiconductor Manufacturing Conference (ASMC), 2022, pp. 1-6, doi: 10.1109/ASMC54647.2022.9792520.

L. Mönch, J. Fowler, S. Dauzère-Pérès, S. Mason and O. Rose, "A survey of problems, solution techniques, and future challenges in scheduling semiconductor manufacturing operations", Journal of Scheduling, vol. 14, no. 6, pp. 583-599, 2011. Available: 10.1007/s10951-010-0222-9.

D. Kopp, M. Hassoun, A. Kalir and L. Mönch, "SMT2020—A Semiconductor Manufacturing Testbed," in IEEE Transactions on Semiconductor Manufacturing, vol. 33, no. 4, pp. 522-531, Nov. 2020, doi: 10.1109/TSM.2020.3001933.

C. Shyalika, S. Thushari and K. Asoka. "Reinforcement learning in dynamic task scheduling: A review." SN Computer Science 1.6 (2020): 1-17, doi: 10.1007/s42979-020-00326-5.