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.

MULTI-AGENT FRAMEWORK FOR INTELLIGENT DISPATCHING AND MAINTENANCE IN SEMICONDUCTOR ASSEMBLY AND TESTING

Christian John Immanuel S. Boydon
Bin Zhang
Cheng-Hung Wu

Institute of Industrial Engineering
National Taiwan University
Roosevelt Road, Da'an District
Taipei City, 10617, TAIWAN

ABSTRACT

This research presents a multi-agent framework for real-time dynamic dispatching and preventive maintenance (DDPM) of unrelated parallel machines in semiconductor assembly and testing. Uncertain job arrivals and random machine health deterioration are considered. Markov decision processes solve the DDPM problem, but are limited by the curse of dimensionality for large systems. To overcome this challenge, a multi-agent framework is designed to decompose the large DDPM problem into single-machine DDPM subproblems whose optimal strategies are then coordinated globally. Simulation studies show that cycle time is improved by at least 13.3% compared with traditional dispatching rules for large systems under a wide range of production settings.

1 INTRODUCTION

This research presents a multi-agent framework for real-time dynamic dispatching and preventive maintenance (DDPM) of unrelated parallel machines in semiconductor assembly and testing. The DDPM framework considers uncertain job arrivals and random deterioration of machine health. The machine health deterioration leads to lower production efficiency, e.g., higher calibration/setup time, higher rework rate due to lower yield quality. This results to environments where machines can process the same job types but have different performance rates, which is common in semiconductor assembly and testing (Fowler and Mönch 2017). To maximize productivity, jobs are efficiently allocated to the machines following a dynamic dispatching strategy that considers real-time factory status and probability of future random events (e.g., job arrival, machine deterioration). A preventive maintenance (PM) strategy for restoring machine efficiency by undergoing scheduled downtime is also devised to further improve long-run performance (Kao et al. 2018).

The joint DDPM framework is effective for achieving long-run productivity but becomes computationally challenging for large systems. Markov decision processes (MDP) are used to model the joint DDPM problem (Wu et al. 2020). The state space is given by WIP level and machine health, which grows exponentially with the number of job types and machines. The decision space gives the job type for dispatch and PM action for each machine, incurring the same dimensional growth. This makes the DDPM problem susceptible to the curse of dimensionality. The computational workload grows exponentially with the number of job types and machines, making it infeasible to solve optimal DDPM strategies for large systems. In practice, the optimal strategy can be solved as needed for systems with one machine. For systems with three or more machines, solving an optimal strategy can take weeks (Huang et al. 2021).

2 MULTI-AGENT FRAMEWORK

To overcome the curse of dimensionality, a modular framework is designed. The framework involves a two-level architecture enabled by a decomposition process. The architecture is presented in Figure 1. Each agent is described below:

- Job assignment agent: Decomposes the large DDPM problem by allocating job types to machines, resulting to DDPM subproblems that each deal with one machine and a small number of job types. The resulting single-machine MDP's are then computationally tractable, overcoming the curse of dimensionality issue.
- Machine agents: Solves local optimal strategies following the decomposition results of the job assignment agent. Since each subproblem deals with one machine, the machine agents can efficiently solve optimal single-machine DDPM strategies, which are then forwarded to the coordination agent.
- Coordination agent: Reconciles the individual solutions of each machine agent and generates a high-level DDPM strategy that achieves global production efficiency.

The algorithms of the agents are linearly scalable with the number of job types and machines, overcoming the curse of dimensionality issue. This enables solving an efficient DDPM strategy for large systems.

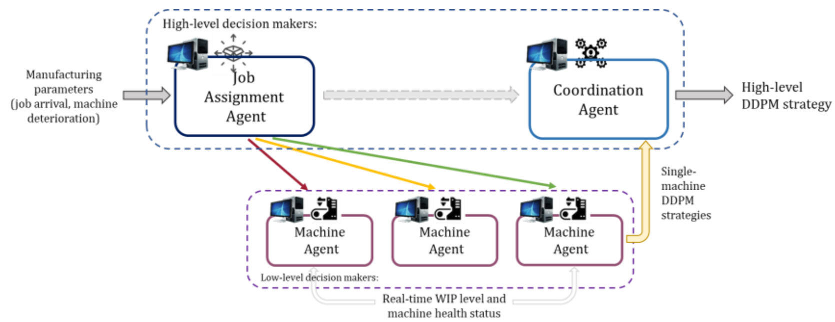

Figure 1: Architecture of the proposed multi-agent framework.

3 SIMULATION EXPERIMENTS

The framework is validated by an extensive simulation study on a system with five job types and five machines in 20 control problems with utilization rate of 0.7 and 0.9. Simulations are done with Markovian and non-Markovian production time distributions. Average cycle times are presented in Table 1, normalized to the results of the multi-agent framework. Compared with traditional dispatching rules, cycle time reduction of at least 13.3% is achieved in the simulation experiments.

Table 1: Average normalized cycle time under different utilization rates.

Markovian production time				Non-Markovian production time			
Utilization rate	**Multi-agent framework**	C_μ	First-come first-served	Utilization rate	**Multi-agent framework**	C_μ	First-come first-served
0.7	**1.000**	1.133	1.281	0.7	**1.000**	1.140	1.317
0.9	**1.000**	1.222	1.363	0.9	**1.000**	1.175	1.382

REFERENCES

Fowler, J. W., and L. Mönch. 2017. "Modeling and Analysis of Semiconductor Manufacturing". In *Advances in Modeling and Simulation*, 301-313. Springer.

Huang, J., D. G. Down, M. E. Lewis, and C.-H. Wu. 2021. "Dynamically Scheduling and Maintaining a Flexible Server". *Naval Research Logistics (NRL).*

Kao, Y.-T., S. Dauzère-Pérès, J. Blue, and S.-C. Chang. 2018. "Impact of Integrating Equipment Health in Production Scheduling for Semiconductor Fabrication". *Computers & Industrial Engineering* 120:450-459.

Wu, C.-H., Y.-C. Yao, S. Dauzère-Pérès, and C.-J. Yu. 2020. "Dynamic Dispatching and Preventive Maintenance for Parallel Machines with Dispatching-dependent Deterioration". *Computers & Operations Research* 113:104779.