

# A new solution to distributed permutation flow shop scheduling problem based on NASH Q-Learning

Ren, J.F.<sup>a,b</sup>, Ye, C.M.<sup>a,\*</sup>, Li, Y.<sup>a</sup>

<sup>a</sup>School of Business, University of Shanghai for Science and Technology, Shanghai, P.R. China

<sup>b</sup>School of Computer and Information Engineering, Henan University of Economics and Law, Zhengzhou, P.R. China

## ABSTRACT

Aiming at Distributed Permutation Flow-shop Scheduling Problems (DPFSPs), this study took the minimization of the maximum completion time of the workpieces to be processed in all production tasks as the goal, and took the multi-agent Reinforcement Learning (RL) method as the main frame of the solution model, then, combining with the NASH equilibrium theory and the RL method, it proposed a NASH Q-Learning algorithm for Distributed Flow-shop Scheduling Problem (DFSP) based on Mean Field (MF). In the RL part, this study designed a two-layer online learning mode in which the sample collection and the training improvement proceed alternately, the outer layer collects samples, when the collected samples meet the requirement of batch size, it enters to the inner layer loop, which uses the Q-learning model-free batch processing mode to proceed, and adopts neural network to approximate the value function to adapt to large-scale problems. By comparing the Average Relative Percentage Deviation (ARPD) index of the benchmark test questions, the calculation results of the proposed algorithm outperformed other similar algorithms, which proved the feasibility and efficiency of the proposed algorithm.

## ARTICLE INFO

### Keywords:

Flow shop scheduling;  
Distributed scheduling;  
Permutation flow shop;  
Reinforcement learning;  
NASH Q-learning;  
Mean field (MF)

### \*Corresponding author:

171910083@st.usst.edu.cn  
(Ye, C.M.)

### Article history:

Received 29 July 2021  
Revised 8 September 2021  
Accepted 26 September 2021



Content from this work may be used under the terms of the Creative Commons Attribution 4.0 International Licence (CC BY 4.0). Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

## References

- [1] Behnamian, J., Fatemi Ghomi, S.M.T. (2016). A survey of multi-factory scheduling, *Journal of Intelligent Manufacturing*, Vol. 27, No. 1, 231-249, [doi: 10.1007/s10845-014-0890-y](https://doi.org/10.1007/s10845-014-0890-y).
- [2] Somashekara, S.C.H., Setty, A.K.Y., Sridharmurthy, S.M., Adiga, P., Mahabaleshwar, U.S., Lorenzini, G. (2019). Makespan reduction using dynamic job sequencing combined with buffer optimization applying genetic algorithm in a manufacturing system, *Mathematical Modelling of Engineering Problems*, Vol. 6, No. 1, 29-37, [doi: 10.18280/mmep.060104](https://doi.org/10.18280/mmep.060104).
- [3] Zhang, Y.Q., Zhang, H. (2020). Dynamic scheduling of blocking flow-shop based on multi-population ACO algorithm, *International Journal of Simulation Modelling*, Vol. 19, No. 3, 529-539, [doi: 10.2507/IJSSIM19-3-C015](https://doi.org/10.2507/IJSSIM19-3-C015).
- [4] Wang, L., Deng, J., Wang, S.-Y. (2016). Survey on optimization algorithms for distributed shop scheduling, *Control and Decision*, Vol. 31, No. 1, 1-11.
- [5] Fernandez-Viagas, V., Framinan, J.M. (2015). A bounded-search iterated greedy algorithm for the distributed permutation flowshop scheduling problem, *International Journal of Production Research*, Vol. 53, No. 4, 1111-1123, [doi: 10.1080/00207543.2014.948578](https://doi.org/10.1080/00207543.2014.948578).
- [6] Gao, J., Chen, R. (2011). An NEH-based heuristic algorithm for distributed permutation flowshop scheduling problems, *Scientific Research and Essays*, Vol. 6, No. 14, 3094-3100.
- [7] Gao, J., Chen, R., Deng, W. (2013). An efficient tabu search algorithm for the distributed permutation flowshop scheduling problem, *International Journal of Production Research*, Vol. 51, No. 3, 641-651, [doi: 10.1080/00207543.2011.644819](https://doi.org/10.1080/00207543.2011.644819).

- [8] Naderi, B., Ruiz, R. (2014). A scatter search algorithm for the distributed permutation flowshop scheduling problem, *European Journal of Operational Research*, Vol. 239, No. 2, 323-334, doi: [10.1016/j.ejor.2014.05.024](https://doi.org/10.1016/j.ejor.2014.05.024).
- [9] Lin, S.-W., Ying, K.-C., Huang, C.-Y. (2013). Minimising makespan in distributed permutation flowshops using a modified iterated greedy algorithm, *International Journal of Production Research*, Vol. 51, No. 16, 5029-5038, doi: [10.1080/00207543.2013.790571](https://doi.org/10.1080/00207543.2013.790571).
- [10] Yang, S.L., Xu, Z.G., Li, G.Z., Wang, J.Y. (2020). Assembly transport optimization for a reconfigurable flow shop based on a discrete event simulation, *Advances in Production Engineering & Management*, Vol. 15, No. 1, 69-80, doi: [10.14743/apem2020.1.350](https://doi.org/10.14743/apem2020.1.350).
- [11] Wang, S.-Y., Wang, L., Liu, M., Xu, Y. (2013). An effective estimation of distribution algorithm for solving the distributed permutation flow-shop scheduling problem, *International Journal of Production Economics*, Vol. 145, No. 1, 387-396, doi: [10.1016/j.ijpe.2013.05.004](https://doi.org/10.1016/j.ijpe.2013.05.004).
- [12] Rifai, A.P., Nguyen, H.-T., Dawal, S.Z.M. (2016). Multi-objective adaptive large neighborhood search for distributed reentrant permutation flow shop scheduling, *Applied Soft Computing*, Vol. 40, 42-57, doi: [10.1016/j.asoc.2015.11.034](https://doi.org/10.1016/j.asoc.2015.11.034).
- [13] Wang, K., Huang, Y., Qin, H. (2016). A fuzzy logic-based hybrid estimation of distribution algorithm for distributed permutation flowshop scheduling problems under machine breakdown, *Journal of the Operational Research Society*, Vol. 67, No. 1, 68-82, doi: [10.1057/jors.2015.50](https://doi.org/10.1057/jors.2015.50).
- [14] Komaki, M., Malakooti, B. (2017). General variable neighborhood search algorithm to minimize makespan of the distributed no-wait flow shop scheduling problem, *Production Engineering*, Vol. 11, No. 3, 315-329, doi: [10.1007/s11740-017-0716-9](https://doi.org/10.1007/s11740-017-0716-9).
- [15] Chen, W., Hao, Y.F. (2018). Genetic algorithm-based design and simulation of manufacturing flow shop scheduling, *International Journal of Simulation Modelling*, Vol. 17, No. 4, 702-711, doi: [10.2507/IJSIMM17\(4\)C017](https://doi.org/10.2507/IJSIMM17(4)C017).
- [16] Lebbar, G., El Abbassi, I., Jabri, A., El Barkany, A., Darcherif, M. (2018). Multi-criteria blocking flow shop scheduling problems: Formulation and performance analysis, *Advances in Production Engineering & Management*, Vol. 13, No. 2, 136-146, doi: [10.14743/apem2018.2.279](https://doi.org/10.14743/apem2018.2.279).
- [17] Rathinam, B., Govindan, K., Neelakandan, B., Raghavan, S.S. (2015). Rule based heuristic approach for minimizing total flow time in permutation flow shop scheduling, *Tehnički Vjesnik - Technical Gazette*, Vol. 22, No. 1, 25-32, doi: [10.17559/TV-20130704132725](https://doi.org/10.17559/TV-20130704132725).
- [18] Babu, K.S., Vemuru, S. (2019). Spectrum signals handoff in LTE cognitive radio networks using reinforcement learning, *Traitement du Signal*, Vol. 36, No. 1, 119-125, doi: [10.18280/ts.360115](https://doi.org/10.18280/ts.360115).
- [19] Foerster, J., Nardelli, N., Farquhar, G., Afouras, T., Torr, P.H.S., Kohli, P., Whiteson, S. (2017). Stabilising experience replay for deep multi-agent reinforcement learning, In: *Proceedings of 34<sup>th</sup> International Conference on Machine Learning*, Sydney, Australia, 1146-1155.
- [20] Tampuu, A., Matiisen, T., Kodelja, D., Kuzovkin, I., Korjus, K., Aru, J., Vicente, R. (2017). Multiagent cooperation and competition with deep reinforcement learning, *PLoS ONE*, Vol. 12, No. 4, Article No. e0172395, doi: [10.1371/journal.pone.0172395](https://doi.org/10.1371/journal.pone.0172395).
- [21] Xu, X., Jia, Y., Xu, Y., Xu, Z., Chai, S., Lai, C.S. (2020). A multi-agent reinforcement learning-based data-driven method for home energy management, *IEEE Transactions on Smart Grid*, Vol. 11, No. 4, 3201-3211, doi: [10.1109/TSG.2020.2971427](https://doi.org/10.1109/TSG.2020.2971427).
- [22] Zerguine, N., Mostefai, M., Aliouat, Z., Slimani, Y. (2020). Intelligent CW selection mechanism based on Q-learning (MISQ), *Ingénierie des Systèmes d'Information*, Vol. 25, No. 6, 803-811, doi: [10.18280/isi.250610](https://doi.org/10.18280/isi.250610).
- [23] Sukhbaatar, S., Szlam, A., Fergus, R. (2016). Learning multiagent communication with backpropagation, In: *Proceedings of 30<sup>th</sup> Conference on Neural Information Processing Systems (NIPS 2016)*, Barcelona, Spain, 2252-2260.
- [24] Foerster, J.N., Assael, Y.M., de Freitas, N., Whiteson, S. (2016). Learning to communicate with deep multi-agent reinforcement learning, In: *Proceedings of 30<sup>th</sup> Conference on Neural Information Processing Systems (NIPS 2016)*, Barcelona, Spain, 2137-2145.
- [25] He, H., Boyd-Graber, J., Kwok, K., Daumé III, H. (2016). Opponent modeling in deep reinforcement learning, In: *Proceedings of 33<sup>rd</sup> International Conference on Machine Learning*, New York, USA, 1804-1813.
- [26] Palmer, G., Tuyls, K., Bloembergen, D., Savani, R. (2018). Lenient multi-agent deep reinforcement learning, In: *Proceedings of 17<sup>th</sup> International Conference on Autonomous Agents and MultiAgent Systems, AAMAS 2018*, Stockholm, Sweden, 443-451.
- [27] Omidshafiei, S., Pazis, J., Amato, C., How, J.P., Vian, J. (2017). Deep decentralized multi-task multi-agent reinforcement learning under partial observability, In: *Proceedings of the 34<sup>th</sup> International Conference on Machine Learning (ICML 2017)*, Sydney, Australia, 2681-2690.
- [28] Sunehag, P., Lever, G., Gruslys, A., Czarnecki, W.M., Zambaldi, V., Jaderberg, M., Lanctot, M., Sonnerat, N., Leibo, J.Z., Tuyls, K., Graepel, T. (2018). Value-decomposition networks for cooperative multi-agent learning based on team reward, In: *Proceedings of the 17<sup>th</sup> International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2018)*, Stockholm, Sweden, 2085-2087.
- [29] Li, C., Tang, J., Ma, T., Yang, X., Luo, Y. (2020). Load balance based workflow job scheduling algorithm in distributed cloud, *Journal of Network and Computer Applications*, Vol. 152, Article No. 102518, doi: [10.1016/j.jnca.2019.102518](https://doi.org/10.1016/j.jnca.2019.102518).
- [30] Neary, C., Xu, Z., Wu, B., Topcu, U. (2021). Reward machines for cooperative multi-agent reinforcement learning, In: *Proceedings of the 20<sup>th</sup> International Conference on Autonomous Agents and MultiAgent Systems, AAMAS '21*, Virtual Event, United Kingdom, 934-942.
- [31] Li, R., Han, Y., Ma, T., Liu, H. (2020). Nash-Q learning-based collaborative dispatch strategy for interconnected power systems, *Global Energy Interconnection*, Vol. 3, No. 3, 227-236, doi: [10.1016/j.gloi.2020.07.004](https://doi.org/10.1016/j.gloi.2020.07.004).

- [32] Li, J., Xiao, Z. (2020).  $H_\infty$  control for discrete-time multi-player systems via off-policy Q-learning, *IEEE Access*, Vol. 8, 28831-28846, doi: [10.1109/ACCESS.2020.2970760](https://doi.org/10.1109/ACCESS.2020.2970760).
- [33] Shah, D., Somani, V., Xie, Q., Xu, Z. (2020). On reinforcement learning for turn-based zero-sum Markov games, In: *Proceedings of the 2020 ACM-IMS on Foundations of Data Science Conference, FODS '20*, Virtual Event, USA, 139-148, doi: [10.1145/3412815.3416888](https://doi.org/10.1145/3412815.3416888).
- [34] Feng, Q., Shang, W.-P., Jiao, C.-W., Li, W.-J. (2020). Two-agent scheduling on a bounded parallel-batching machine with makespan and maximum lateness objectives, *Journal of the Operations Research Society of China*, Vol. 8, No. 1, 189-196, doi: [10.1007/s40305-019-00258-9](https://doi.org/10.1007/s40305-019-00258-9).
- [35] Littman, M.L. (1994). Markov games as a framework for multi-agent reinforcement learning, In: *Proceedings of the Eleventh International Conference on International Conference on Machine Learning, ICML'94*, New Brunswick USA, 157-163, doi: [10.1016/B978-1-55860-335-6.50027-1](https://doi.org/10.1016/B978-1-55860-335-6.50027-1).
- [36] Hu, J., Wellman, M.P. (1998). Multiagent reinforcement learning: Theoretical framework and an algorithm, In: *Proceedings of the Fifteenth International Conference on Machine Learning, ICML '98*, Madison, USA, 242-250.
- [37] Xiao, L., Bahri, Y., Sohl-Dickstein, J., Schoenholz, S., Pennington, J. (2018). Dynamical isometry and a mean field theory of cnns: How to train 10,000-layer vanilla convolutional neural networks, In: *Proceedings of the 35<sup>th</sup> International Conference on Machine Learning*, Stockholm, Sweden, 5393-5402.
- [38] Yang, Y., Luo, R., Li, M., Zhou, M., Zhang, W., Wang, J. (2018). Mean field multi-agent reinforcement learning, In: *Proceedings of the 35<sup>th</sup> International Conference on Machine Learning*, Stockholm, Sweden, 5571-5580.
- [39] Ruiz, R., Pan, Q.-K., Naderi, B. (2019). Iterated Greedy methods for the distributed permutation flowshop scheduling problem, *Omega*, Vol. 83, 213-222, doi: [10.1016/j.omega.2018.03.004](https://doi.org/10.1016/j.omega.2018.03.004).