Research on Scheduling Method for Uncertainty of Hit Rate of Molten Steel Based on Q Learning

Liangliang Sun, Tianyi Lu, Shuya Sha, Wanying Zhu, Qiuxia Qu, and Baolong Yuan

Abstract Aiming at the problem of uncertain scheduling of molten steel hit rate in the steel refining process, taking into account the multi-stage, multi-equipment, and multi-constrained production process conditions of refining production and the process of refining process due to the uncertainty of molten steel hit rate during the refining process, in order to obtain a scientific and feasible approximate optimal scheduling plan in a short period of time, the system state and system state transfer rules of the steel production process are defined, and the random evolution scheduling optimization system model of steel production refining based on the discrete-time Markov chain is established. At the same time, in the refining process scheduling optimization problem, the complexity of the solution will increase exponentially with the increase of the number of reprocessing processes, and a stochastic dynamic programming algorithm based on heuristic simulation strategy and improved Q learning is designed to solve the problem. Aiming at the uncertain scheduling problem of molten steel hit rate under different process production paths, simulation experiments using actual production data of a large domestic steel mill verify the effectiveness of the proposed model and algorithm.

Keywords Production scheduling · Molten steel hit rate · Markov chain · Q learning

1 Introduction

The typical steel production process includes iron making, steelmaking-continuous casting, and rolling [1]. Steelmaking-continuous casting production is the bottleneck of the entire steel production which includes three stages of steelmaking, refining, and continuous casting. The refining stage consists of RH, CAS, KIP, and other refining stations [2].

L. Sun (✉) · T. Lu · S. Sha · W. Zhu · Q. Qu · B. Yuan
Faculty of Information and Control Engineering, Shenyang Jianzhu University, Shenyang, China
e-mail: swinburnsun@163.com

Y. Li et al. (eds.), *Advances in Simulation and Process Modelling*,
Advances in Intelligent Systems and Computing 1305,
https://doi.org/10.1007/978-981-33-4575-1_15

As the refining process of the steelmaking-continuous casting intermediate link, from the perspective of the process, the process undertakes the ironmaking production process and closely follows the rolling production process, and its productivity is weaker than other production processes [3], so it is considered to be steel production bottleneck link; from the perspective of management level, not only Enterprise Resource Planning (ERP) and Process Control System (PCS), but also the core link of Manufacture Executive System (MES) [4]. In the refining stage, steel production enterprises usually arrange molten steel into refining equipment for different impurity removal methods for production according to customers' requirements for molten steel composition. The steel scheduling in the refined production process is based on the number of furnaces. According to the corresponding production process paths obtained by different furnaces in different attributes of the batch planning process in daily units, considering the constraints of the steelmaking-continuous casting production process, select each types of refining equipment corresponding to the production time of each furnace and the different equipment serial numbers under the same type of refining equipment [5]. As the core link of steel production, the steelmaking-continuous casting production process is disturbed by many uncertain factors [6]. After the occurrence of the molten steel hit rate that does not meet the standard with a high frequency occurs, it is necessary to determine the content of sulfur, phosphorus, and carbon in the molten iron in the unit of the furnace after the production of a certain equipment is completed at a certain stage. If the content of the component meets the standard, the next process is processed according to the established process path and static scheduling compilation rules; if the content of the component does not meet the standard, the accurate content of the component needs to be judged. However, due to the difficulty of guaranteeing the stability of the molten steel component treatment with the furnace as the production unit in the refining production equipment (RH type, CAS type, KIP type, etc.), the steel hit rate may not reach the standard (the composition of molten steel from a certain refining process in units of furnace times does not meet the established requirements). As a result, the molten steel needs to be returned to one or more of the previous processes for refining. If a batch of refining occurs due to unqualified molten steel quality, it will seriously affect the normal production of subsequent batches. As a result, the refining production process cannot be produced in strict accordance with the production scheduling plan formulated before production, which is difficult to guarantee. The entire steelmaking-continuous casting production rhythm affects the efficiency of steel production, increases the energy consumption of steel production, and increases the use of personnel costs. To this end, this paper proposes a study on the uncertain scheduling method of molten steel hit rate in the refining production process to ensure the smooth progress of steel production.

2 Scheduling Modeling

In general, when defining the system state of the refining production process scheduling problem, it is necessary to consider the number of furnaces that the steel production enterprise simultaneously performs refining production on the same set of refining equipment, the production status of each furnace at a certain time, and the number of ways for each refining process to complete the production task and the situation of each type of refining equipment being occupied at a certain time. According to the above requirements for the definition of the status of the scheduling system of the steel-refining production process, we assume that a steel production enterprise simultaneously carries out D refining production tasks on the same set of refining production equipment, and the completion of D refining production tasks requires E types refining equipment, the number of each refining equipment is $H_k(k \in \{1 \ldots E\}$ (which is a positive integer). The state of the refining production process scheduling system is defined as follows:

$$X = [s_1, s_2, \ldots, s_D, q_1, q_2, \ldots, q_D, R_1, R_2, \ldots, R_E, t]^T \tag{1}$$

where s_i is a positive integer, indicating the current refining production status of the i-th heat, $i \in \{1, \ldots, D\}$; q_i is an integer which represents the number of ways to complete production tasks in a refining production process of the i-th furnace; R_j represents the number of refining and processing equipment of the j-th type that is not occupied at the current moment, and $R_j \in \left\{0, \ldots, H_j\right\}$ is a positive integer ($j \in \{1, \ldots, E\}$, j is a positive integer).

In general, before defining the execution status of the refining production process in the refining production process, it must be ensured that at least one refining production and processing equipment required to execute this refining production process is available to execute this refining production process. According to the above requirements for the execution status of the production process in the refining production process, we define the execution status of the production process in the refining production process as:

$$U = [\beta_1, \beta_2, \ldots, \beta_D]^{\mathrm{T}} \tag{2}$$

where U represents the production execution state of each furnace process in the refining production process; β_i represents the execution state of a production process of the i-th furnace, $\beta_i = 1$ or 0, ($i \in \{1, \ldots, D\}$, i is positive Integer), $\beta_i = 1$ means to execute a certain process of the i-th heat and $\beta_i = 0$ means not to execute a certain refining production process of the i-th heat.

During the refining production process, when a certain refining production process within a certain furnace is started or the task of a refining production process is completed, the state of the refining production process scheduling system will shift. If in the same state of refining production scheduling system, the production task of a production process of a certain batch of refining production is completed first, the state of refining production scheduling system will be transferred to the state of temporary refining production scheduling system. We assume that the initial state of the refining production scheduling system is:

$$x(0) = [1, \ldots, 1, 0, \ldots, 0, R_1, \ldots, R_E, 0]^{\mathrm{T}} \tag{3}$$

where $x(0)$ represents the initial state of the refining production scheduling system; 1 represents the state of the production scheduling system where each furnace is ready to perform the first refining process; 0 represents the completion of each furnace performing the first refining process; $R_j \in \{0, \ldots, H_j\}$, ($j \in \{1, \ldots, E\}$, j is a positive integer, R'_j is a positive integer), which represents the number of the j-th refining and processing equipment that is not occupied at the current moment.

$$x(0) = [1, \ldots 1, 0, \ldots, 0, R_1, \ldots, R_E, 0]^{\mathrm{T}} \tag{4}$$

Make reasonable arrangements for the refining equipment required for each refining production and the start processing time of each refining production process in each furnace. Therefore, in this paper, 'the minimum sum of the waiting time of the processing in each process of the refining production process' and 'the minimum difference between the ideal opening time and the actual opening time of each furnace in the steel-refining production process' are taken as the optimization goals, with 'the same refinement equipment handles adjacent furnaces at the same time without conflicting furnaces' as a constraint, and establishes the following mathematical objective function, $Q(x(k), u(k))$ means that refining is performed in the refining production state $x(k)$. The starting processing time for processing by the production action is $u(k)$, where $Q(x(k+1), u(k+1))$ represents the set of all refining production actions that can be selected for the refining production state $x(k+1)$. The starting processing time for production, $g(x(k), u(k), x(k+1))$ is the steel-refining production action $u(k)$ from the refining production state $x(k)$ to the steel-refining production. Time of state $x(k+1)$:

$$\begin{aligned} &\mathrm{Min}(Qx(k), u(k)) = \alpha \mathrm{Min}_{u(k+1)\in U^{x(k+1)}} \mathrm{E}[Q(x(k+1), u(k+1))] \\ &+ (1-\gamma) Q(x(k), u(k)) + \gamma \{[g(x(k), u(k), x(k+1)\} \\ &+ |Q(x(k), u(k)) - T_i| \end{aligned} \tag{5}$$

where $\mathrm{Min}\,Q(x(k), u(k))$ represents the minimum processing time for the refining production system to perform the refining production action $u(k)$ in the refining production state $x(k)$; $Q(x(k), u(k))$ represents the processing time for the current refining production system to perform the refining production action $u(k)$ in the refining production state $x(k)$; $\mathrm{g}(x(k), u(k), x(k+1))$ represents the time from the refining production state $x(k)$ to perform the refining production action $u(k)$ for refining production to the refining production state $x(k+1)$; $\mathrm{Min}_{u(k+1)\in U^{x(k+1)}} \mathrm{E}[Q(x(k+1), u(k+1))]$ means to perform the refining production action $u(k+1)$ processing in the refining production state $x(k+1)$, the minimum value of the expected value of time mathematics; γ represents the discount factor ($\gamma \in \{0, \ldots, 1\}$); α represents the learning coefficient ($\alpha \in \{0, \ldots, 1\}$); $|Q(x(k), u(k)) - T_i||$ means that the refining production system performs refining in the refining production state $x(k)$, the difference between the ideal opening time of the production action $u(k)$ and the actual opening time; T_i represents the furnace i ideal opening pouring time ($k \in \{1, \ldots, E\}$, $i \in \{1, \ldots, D\}$, k and i are positive integers), and the ideal opening time of furnace i needs to meet:

$$T_i > T_\varphi + T_\omega. \tag{6}$$

T_φ represents the processing time of the heat on the refining equipment; T_ω represents the sum of the waiting time of the adjacent processes of the heat on the refining equipment. Through Eq. (5), the performance indexes 'sum of waiting times of refining production process heat in each process' and 'difference between the ideal opening time and actual opening time of each heat in refining production process' are converted into optimization goals, Next, another performance index, 'processing adjacent furnaces on the same refining equipment within the same time without 'operation conflict,' is converted into constraint conditions by Eq. (7).

$$\begin{aligned} &Q(x(k+1), u(k+1)) > Q(x(k), u(k)) + \mathrm{g}(x(k), u(k), x(k+1)), \\ &k \in \{1, \ldots, E\}; i \in \{1, \ldots, D\} \end{aligned} \tag{7}$$

In Eq. (7), $\mathrm{g}(x(k), u(k), x(k+1))$ is the refining production action $u(k)$ from the refined state $x(k)$ to the refined state $x(k+1)$'s production time.

3 Solution Methodology

The problem of refining production scheduling under the uncertain environment of molten steel hit rate is a large-scale flow shop scheduling problem, which is to solve the NP problem. Considering that the Q learning algorithm in reinforcement learning has adaptive, greedy search, and can quickly search for the optimal solution, but the traditional Q learning algorithm has the disadvantage that it cannot accurately select the next optimal state. Taking this problem into consideration, improvements are made on the basis of the traditional Q learning algorithm. Iterative calculation using the improved Q learning algorithm will save processing time in the refining production stage, improve the efficiency of refining production scheduling, and be able to better cope with actual refining production scheduling during the process, and a sudden situation occurs that the steel composition cannot meet the production requirements and need to be reprocessed. According to the scheduling mathematical model built in this paper, a method for solving the uncertain scheduling problem of molten steel hit rate in the refining production process using the improved Q

learning method is proposed. The following is the procedure of improving the Q learning algorithm.

Step 1: Define the state action pair of refining production process $(x(k), u(k))$, $x(k)$ represents the production state of each furnace in the refining production stage, $u(k)$ represents each furnace in the refining production stage for the same kind of refining production and processing equipment selection status ($k \in \{1, \ldots, E\}$ is a positive integer); and need to build a matrix of uncertain molten steel hit rate in the refining production process, the refining production process experience matrix R and refining production process learning matrix Q.

Step 2: Initialize the learning matrix Q of the refining production process and set the learning coefficient α;

Step 3: Select the state action pair $(x(1), u(1))$ of the first refining process in the refining production process as the initial state;

Step 4: Use the objective function formula to calculate;

Step 5: When refining production occurs on the same refining production equipment at the same time for adjacent processes of different times, select the probability by calculating the action, such as Eq. (8), arrange the order in which adjacent refining processes of different furnace times enter the refining equipment for production

$$P(a_i/S_t) = Q(x_k, u_k)/\sum_j Q(x_k, u_k) \tag{8}$$

where $Q(x_k, u_k)$ is the processing time matrix for production on the refining equipment u_k in the production state x_k; $\sum_j Q(x_k, u_k)$ means that in the j-th furnace production state x_k is selected in the refining equipment. The sum of the processing time matrices of all the processes performed on u_k, $u_k \in \{U\}$;

Step 6: Determine the update status of the learning matrix Q matrix in the refining production process. If the objective function is greater than 0, the learning matrix Q in the refining production process is updated. Otherwise, it will go back to step 4 and use the selected next production state as the initial state to perform iterative calculation using the objective function;

Step 7: Determine the convergence of the learning matrix Q in the refining production process. The conditions for the learning matrix Q to converge are: $\max_j \left|(Q_j^{i+1} - Q_j^i)/Q_j^i \leq 1\right|$.

Step 8: Obtain the learning matrix Q for the condensed refined production process, then the algorithm ends.

At end of improved Q learning algorithm, it enters the manual online decision-making process, the dispatcher uses the Q learning algorithm to obtain the condensed refined production learning matrix Q. The molten steel produced by each process is individually selected according to the different production requirements of each furnace:

$$u^*(k) = \arg \max_{u(k) \in U^{x(k)}} Q(x(k), u(k)) \tag{9}$$

where $u^*(k)$ represents the optimal refining production equipment selection state; $\arg \max_{u(k) \in U^{x(k)}} Q(x(k), u(k))$ indicates the state action pair that maximizes the value of $Q(x(k), u(k))$.

4 Simulation Study

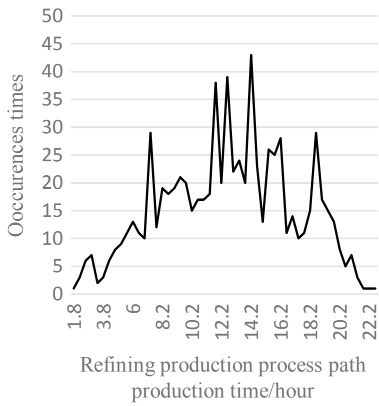
A steel production enterprise simultaneously undertakes three refining production tasks with different process paths and furnaces. The equipment and corresponding quantities used by the enterprise to complete this refining production task are: two RH refining equipments, two KIP refining equipments, two LF refining equipments, and one CAS refining equipment, and steel production enterprises simultaneously undertake three different process paths for refining production flowcharts.

The improved Q learning algorithm solution strategy proposed in this paper is used to verify and solve the problem of uncertain scheduling optimization of molten steel hit rate in the refining production process of different process paths, and the learning matrix Q of the convergent refining production process is obtained. When it is necessary to carry out refining, the condensed refining production process learning matrix Q is used for manual online decision making to obtain a scientific and effective scheduling plan.

Using heuristic strategy simulation, the refining production process, the furnaces in each process have a small waiting time, and the small furnace production process first enters the refining production equipment for production and the furnace process with early opening time first enters the refining production equipment for production. Each time 1000 simulations are performed, and 2000 refining production process paths can be obtained. The simulation results are shown in Figs. 1 and 2.

In Figs. 1 and 2, the simulation strategy 2 is that the furnace process with the early opening time first enters the refining production equipment for production. The

Fig. 1 Simulation results 1



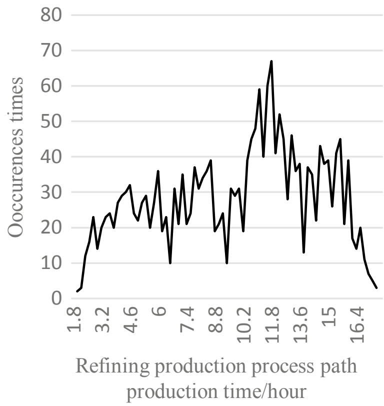


Fig. 2 Simulation results 2

simulation results are concentrated around the average production time of the refining production process path and the refining production process path in the simulation process. The average production time is 4.6 h ahead of simulation strategy 1.

5 Conclusion

This paper first describes the two performance indicators 'the sum of the waiting time of the furnace process in each process is the smallest' and 'the deviation between the ideal opening time and the actual opening time is small'; the Markov chain is used to describe the hit rate of molten steel uncertainty, define the refining production state, refining production process execution status, and refining production state transfer rules; establish a scheduling optimization mathematical model for the uncertain scheduling problem of molten steel hit rate in the refining production process. Then, a solution strategy for the refining production scheduling problem under uncertain environments of molten steel hit rate is proposed. First, the solution strategy is to use heuristic strategy simulation to reduce the optimal process path selection range and obtain the initial Q value of the iterative calculation of the subsequent improved Q learning method; second, using improved Q learning method to iteratively solve the refining production scheduling problem under uncertain environments of molten steel hit rate. In order to overcome the problem that the traditional Q learning algorithm cannot accurately select the optimal state action pair for each iteration calculation when solving the scheduling problem, use the Pareto solution set optimization solution idea introduces the action selection probability; finally, manual online decision making is used to obtain a scientific and feasible production scheduling scheme for the refining stage under the uncertain environment of molten steel hit rate. At the end of the paper, the solution strategy proposed in this paper is used to verify the case of molten steel hit rate uncertain scheduling under different process production paths. The verification results show that the proposed solution is feasible to solve the problem of scheduling optimization in the uncertain environment of molten steel hit rate in the refining production process, which promotes the research of industrial scheduling theory with the characteristics of large-scale mixed flow shop scheduling problems.

Acknowledgements The research is sponsored by the National Natural Science Foundation of China (61873174, 61503259), China Postdoctoral Science Foundation Funded Project (2017M611261), Science and Technology Projects of Ministry of Housing and Urban Rural Development (2018-K1-019), Liaoning Provincial Natural Science Foundation of China (20180550613, 2020-KF-11-07), Young science and technology innovation talent support plan (RC200003), and Liaoning Revitalization Talents Program (XLYC1807115).

References

1. Pang, X.F., Jiang, Y.C., Yu, S.P., Li, H.B., Gao, L., Che, Z.H.: Flexible job shop rescheduling method of steelmaking-continuous casting base on human-computer cooperation. Comput. Integr. Manuf. Syst. **24**(10), 2415–2427 (2018) (in Chinese)
2. Mao, K.: Lagrange Relaxation Level Optimization Method and its Application to Production Scheduling of the Steelmaking-Continuous. Northeastern University (2014) (in Chinese)
3. Liu, G., Luh, P.B., Resch, R.: Scheduling permutation flow shops using the Lagrangian relaxation technique. Ann. Oper. Res. **3**(4), 171–189 (1997)
4. Zheng, Z., Long, J.Y., Gao, X.Q., Gong, Y.M., Hu, W.Z.: Present situation and prospect of production control technology focusing on planning and scheduling in iron and steel enterprise. Comput. Integr. Manuf. Syst. **20**(11), 2660–2674 (2014) (in Chinese)
5. Sun, L.L.: Research on the optimal scheduling method for the productive process of steelmaking-refining-continuous casting. Northeastern University (2014) (in Chinese)
6. Tan, Y.Y., Huang, Y.L., Liu, S.X.: Two-stage mathematical programming approach for steelmaking process scheduling under variable electricity price. J. Iron Steel Res. (International) **12**(20), 1–8 (2013)