Research on coordinated scheduling of straddle carriers and quay cranes in automated container terminals based on reinforcement learning

Zhenyu Fang*, Xiaolong Han
Shanghai Maritime University, Institute of Logistics Science & Engineering, Shanghai, China
*Fangzhenyu123a@163.com

Abstract. Aiming at the coordination and scheduling problem between Automated Straddle Carrier (Automated Straddle Carrier) and Quay Crane (QC) in automated container terminals, consider that the quay crane cannot cross the straddle carrier, and the straddle carrier and the quay crane cannot be in adjacent lanes at the same time. A mixed-integer optimization model with the objective function of minimizing the final completion time of the quay crane is established under the constraints of different operating speeds and different speeds of straddle carriers in different states. A genetic algorithm based on reinforcement learning is designed, and the initial population is generated by the Q-learning algorithm, and the genetic algorithm (GA) is iterated to increase the diversity of the initial population. Finally, taking 5 groups of experiments as examples, the model is compared and solved by GAMS solver, genetic algorithm and genetic algorithm based on reinforcement learning. The results of an example show that the genetic algorithm based on reinforcement learning can solve the model and make the value of the objective function smaller, thus verifying the feasibility of the modified algorithm.

Keywords: Automated container terminal; straddle carrier; quay crane; coordinate scheduling; reinforcement learning algorithm; genetic algorithm.

1. Introduction

With the continuous growth of container shipping volume in the world and the increasing size of container ships, the stability and efficiency, energy saving, environmental protection and low cost of container loading and unloading have gradually become the focus of port operators. [1] Compared with traditional container terminals, automated container terminals have greater advantages.

In the existing research, most scholars pay attention to the cooperative scheduling of AGV and quayside crane. Chen et al. [2] aiming at the charging problem in AGV scheduling, an AGV scheduling model considering charging strategy is established and solved by Gurobi and genetic algorithm. Liang Chengji et al. [3] aiming at the coordinated dispatching problem of AGV and double-trolley shore bridge, considering the limitation of transit platform and capacity of double-trolley shore bridge, taking the time window of double-trolley shore bridge as the constraint, a mixed integer programming model is established, and the solving effects of genetic algorithm and particle swarm optimization are compared. Huang Yongfu et al [4] aiming at the integrated scheduling problem of quayside crane, AGV and yard crane, with the goal of minimizing the total completion time of ship unloading task, an integrated scheduling mixed integer linear programming model considering buffer capacity is established, and a heuristic algorithm is designed to solve the problem. According to the influence of Andong et al. [5] on the operation mode and economic index of the terminal. YuChen et al. [6] established a mathematical model aimed at minimizing the total running time of automated straddle vehicles. Hamdi D [7] comprehensively models the storage location allocation problem. Skinner et al. [8] use genetic algorithm and sequential scheduling method to solve the model, which effectively reduces the overall time.

To sum up, in the literature related to the coordinated dispatching of various equipment in the wharf, a small part of the literature considers the coordinated dispatching of quay crane and straddle vehicles, and the solution method is relatively simple.
2. Problems Description

As shown in figure 1, the plane layout of the automatic container terminal based on straddling vehicles is mainly divided into berths, interaction area between straddling vehicles and shore crane, straddling vehicles waiting area and storage yard, in which straddling vehicles can complete the loading and unloading operation when interacting with the yard alone without the assistance of other equipment.

The quayside crane grabs the container and loads it to the designated shell on board, and then goes on to the next container task, while the unloading process is the opposite.

Fig.1 Plane layout of automated container terminal based on straddle carrier

This paper focuses on the coordinated scheduling of shore crane and straddle vehicle. In the solution method, the genetic algorithm based on reinforcement learning is designed and compared with GAMS solver and genetic algorithm, and the examples are analyzed.

3. Establishment Of Mathmatical Model

When establishing the mathematical model, we first put forward several reasonable assumptions: all the container tasks are known in the container area, the time for unloading the container across the truck is fixed, and the time for the quayside crane to grab the container is fixed.

3.1 Notations

(a). Collection and index: K is the quayside crane set, k, l ∈ K; V is a collection of straddling vehicles, v ∈ V; I is a collection of container tasks; i, j ∈ I;

(b). Parameters: S1 is the no-load speed of the straddle vehicle, in m/s; S2 is the heavy-duty speed of the straddle vehicle, in m/s; tq is the pick-up time of quayside crane car, in s; ts is the time for placing containers for straddling vehicles, in s.

(c). Variables: cdtkv is the distance from the end point of the previous task to the container area of the I container task of the quayside crane k plan, in m;

bdtkv is the distance from the container area of the I container mission to the front of the wharf for the cross carrier v of the I container mission in the quayside crane k plan, in m;

eik for quayside crane k car to carry out I task loading and return to the seaside for use;

Aikv is the arrival time of the straddle vehicle v belonging to the I container mission in the quayside crane k plan to the front of the wharf;

Qik is the time when the I container task in the quayside crane k plan begins to be unloaded to the driveway;

Fikv is the time when the straddle vehicle v belonging to the I container task in the quayside crane k plan ends the unloading operation and leaves the driveway;

Stlk is the time when the first container task in plan k of quayside crane begins to be captured by quayside crane trolley;

bik is the loading time of the I container mission in the quayside crane k plan;

Hik for the time when the first container task in the quayside crane k plan is grabbed by the quayside crane trolley and left the seaside;
$W_{ik}^v$ is the waiting time of the straddle vehicle belonging to the I container in the quayside crane k plan; 
$Z_{ijk}$ is 0-1 variable, the first container task in the quayside crane k plan is captured by the quayside crane trolley, it ends at 1 before the departure of the straddle vehicle belonging to the j container task, otherwise it is 0; 
$C_{ijk}$ is a variable of 0-1, and the cross-vehicle unloading operation of the first container task in plan k ends at 1 before the j container task is grabbed by the quayside crane trolley, otherwise it is 0; 
$L_{ijkl}$ is a variable of 0-1, and the lane number assigned by the 1st container task of quayside crane k is less than that of the j container task of quayside crane l, otherwise it is 0; 
$X_{ik}$ is the decision variable, the lane number assigned to the I container task in the quayside crane k plan; 

$X_{ik}^jv$ is the decision variable. After completing the quayside crane k container task I, the straddle vehicle v completes the quayside crane l container task j, otherwise it is 0.

3.2 Mathematical model

Objective function:

Min $F$  

(1)

Constraints:

\[ F \geq e_{ik}, k \in K, i \in I \]

(2)

\[ \sum_{t \in K} \sum_{j \in I} A_{ik}^j = 1, k \in K, i \in I, v \in V \]

(3)

\[ \sum_{k \in K} \sum_{t \in I} A_{ik}^j = 1, k \in K, i \in I, v \in V \]

(4)

\[ e_{ik} \geq H_{ik} + b_{ik}, k \in K, i \in I \]

(5)

\[ H_{ik} \geq S_{ik} + t_q, k \in K, i \in I \]

(6)

\[ S_{ik} \geq A_{ik} + W_{ik}^v + ts, k \in K, i \in I \]

(7)

\[ A_{ik}^v \geq cd_{ik}^v/s_1 + bd_{ik}^v/s_2, k \in K, i \in I, v \in V \]

(8)

\[ W_{jk}^v \geq e_{jk} - A_{ik}^v - ts + (X_{ij}^jv - 1)M, k, l \in K, i, j \in I, v \in V \]

(9)

\[ Q_{jk} - H_{ik} \leq M \cdot Z_{ijk}, k \in K, i, j \in I \]

(10)

\[ H_{ik} - Q_{jk} \leq M(1 - Z_{ijk}), k \in K, i, j \in I \]

(11)

\[ S_{ik} - F_{jk}^v < M \cdot C_{ijk}, k \in K, i, j \in I, v \in V \]

(12)

\[ F_{jk}^v - S_{ik} \leq M(1 - C_{ijk}), k \in K, i, j \in I, v \in V \]

(13)

\[ x_{jl} - x_{ik} < M \cdot b_{ijk}, k \in K, l \in K, i, j \in I \]

(14)

\[ x_{ik} - x_{lj} \leq M(1 - L_{ijkl}), k, l \in K, i, j \in I \]

(15)

Formula (1) is the objective function to minimize the final quayside crane completion time; (2) the final task completion time is not less than the time for any quayside crane to complete the loading task and return to the task;

4. Design Of Ga Algorithm Based On Q-Learning

formula (3) and formula (4) have only one pre-order task and one post-order task for any container task. Type (5) the time for the quayside crane to complete the loading task (6) the time of grabbing the box and leaving the front of the wharf is determined by the time of the quayside crane trolley starting to grab the box and the time it takes to unload the box; (7) the starting time of the quayside crane trolley is determined by the time the cross-carrier arrives at the front of the wharf, the waiting time and the time it takes for the cross-carrier to unload. Formula (8) is determined by the time that the straddle vehicle arrives at the front of the wharf before the end of the task and the time it takes from the box area to the front of the wharf; formula (9) is the waiting time of the straddle carrier; the sum of formula (10) and formula (11) to realize that the I task in the quayside crane k plan is captured by the quayside crane trolley before the departure of the straddle vehicle v of the j task is 1, otherwise it is 0. Formula (12) and formula (13) realize that the straddle vehicle v falling box operation of the I
task in the k plan of the quayside crane is 1 when the quayside crane trolley grasping operation of the j task ends, otherwise it is 0. Formula (14) and formula (15) are assigned in the same lane when the value of $L_{j,k+1} + L_{j,k}$ equals 0.

Q-learning algorithm is a model-independent algorithm in reinforcement learning [9]. As shown in figure 2, in state $S_t$, the agent Agent gets the reward value $\alpha_t$ at tweak 1 time by performing the action. Each time, the learning agent selects an action according to the strategy, then perceives the next state and reward value, adjusts the strategy according to experience and finds the optimal strategy. Because the traditional GA is highly dependent on the initial population and easy to fall into the local optimal solution, this paper combines the Q-learning algorithm with the GA algorithm, and uses the Q-learning algorithm to generate the initial population into GA for iteration. The flow chart is shown in figure 3.

![Fig.2 Agent-environment interaction diagram](image)

4.1 Q-learning algorithm steps

According to the problems studied in this paper, environment refers to automatic container terminal, which mainly includes three aspects: yard distribution, quayside crane and automatic straddle vehicle. In this paper, considering the operation characteristics of straddle vehicle and shore bridge and the waiting time for each other, the reward value is constructed according to the minimum task completion time, which is expressed by formula (16).

$$R(s, a) = -e_{ik}$$  \hspace{1cm} (16)

Step1 Initialize algorithm parameters. Step 2 Determine that the current state of the environment is $S_t$. Step 3 According to the current state $S_t$, the greedy strategy ($\varepsilon$-greedy) is used to select action $\alpha_t$. Step 4 After the execution of action $\alpha_t$, the environmental state $S_{t+1}$ and the reward value $r$ are determined at the next moment. This paper defines the reward value as -1 * (the time required for the cross-carrier and shore crane to carry out the container task), that is, the longer the time required, the smaller the reward value. Step 5 Update Q value table, $t = t + 1$, the update formula is as follows

$$Q(s_t, a_t) = Q(s_t, a_t) + \alpha [r_{t+1} + \gamma \max Q(s_{t+1}, a_t) - Q(s_t, a_t)] \quad 0 < \alpha < 1; 0 < \gamma < 1$$  \hspace{1cm} (17)

Step 6 Whether the maximum number of search iterations is reached, and if so, continue with Step7, otherwise return Step2. Step 7 Generate a chromosome according to the Q value table. Step 8 If the number of chromosomes generated reaches the initial set population size, it will be transferred to Step9, otherwise Step2.
4.2 GA algorithm

(a). Chromosome coding: the chromosome code in the initial population generated according to the above Q-learning steps is shown in figure 4, the chromosome length is 2 * the number of container tasks, the left side is the straddle carrier number to which the task belongs, and the right side is the corresponding shore crane number. If task 1 is completed by No. 1 straddle carrier and No. 1 shore crane.

(b). Fitness function: the fitness function in this paper is the reciprocal of the minimum maximum completion time of the quay crane in the model, 1/F.

(c). Choice, this paper adopts the classical roulette selection method.

(d). Crossover, this paper adopts multi-point crossover method. As shown in figure 5.

(e). In this paper, exchange variation is used in this paper. As shown in figure 6.

Parameter setting of a numerical example

In order to verify the accuracy of the above model and the effectiveness of the algorithm, this paper uses GAMS solver and MATLAB software to analyze the examples. This paper sets the probability of action selection Egr=0.9; Number of exploration iterations Epi=20; Learning factor alpha=1; Discount factor gamma=0.8; Number of iteration Gen=500; Number of population Pop=200; Crossing rate PC=0.7; Mutation rate PM=0.05; Number of lanes is 4; The box grabbing time of quayside crane and straddle vehicle is 60s; The random number of the loading time of quayside crane after catching the box is between 100 and 140s; The driving speed of empty vehicle and heavy vehicle is 4m/s and 3m/s respectively.

4.3 Example analysis

Tasks of different sizes of five cross-transport vehicles is calculated by using GAMS solver, GA algorithm and GA algorithm based on Q-learning respectively as shown in Table 1.

<table>
<thead>
<tr>
<th>Number of tasks</th>
<th>GA results/s</th>
<th>Q-GA results/s</th>
<th>GAMS results/s</th>
<th>Q-GA better than GA</th>
</tr>
</thead>
<tbody>
<tr>
<td>20</td>
<td>2447.59</td>
<td>2354.074</td>
<td>2254.12</td>
<td>3.82%</td>
</tr>
<tr>
<td>50</td>
<td>6044.64</td>
<td>5771.22</td>
<td>5569.4</td>
<td>4.52%</td>
</tr>
<tr>
<td>70</td>
<td>9355.65</td>
<td>9002.84</td>
<td>8685.42</td>
<td>3.77%</td>
</tr>
<tr>
<td>100</td>
<td>13999.56</td>
<td>13358.47</td>
<td>12642.67</td>
<td>4.58%</td>
</tr>
<tr>
<td>150</td>
<td>23390.06</td>
<td>22194.06</td>
<td>21290</td>
<td>5.11%</td>
</tr>
<tr>
<td>200</td>
<td>32185.59</td>
<td>31578.67</td>
<td>30725.29</td>
<td>1.89%</td>
</tr>
</tbody>
</table>
As can be seen from Table 1, the Q-learning-based GA algorithm is 3% higher than the GA algorithm. Taking 20 container tasks is shown in figure 7. Figure 8 shows the Gantt chart of cross-vehicle and quayside crane scheduling under 20 container tasks.

![Fig.7 Convergence graph for 20 container tasks](image)

![Fig.8 scheduling Gantt chart of 20 container tasks](image)

5. Conclusion

In this paper, aiming at the coordinated scheduling problem of quayside crane and automatic straddle vehicle in automatic container terminal, with the goal of minimizing the final completion time of quayside crane, the model is solved by GAMS solver, GA and GA algorithm based on Q-learning. The analysis of a numerical example shows that the GA algorithm based on Q-learning can get a better solution, which proves the feasibility and effectiveness of the algorithm.

References