Preprints of the 21st IFAC World Congress (Virtual)
Berlin, Germany, July 12-17, 2020

A New Reinforcement Learning for Multi-Train Marshaling with Time Evaluation

Yoichi Hirashima *

* *Osaka Institute of Technology, 1-79-1, Kitayama, Hirakata, Osaka, 573-0196, Japan. (e-mail: hirash-y@ is.oit.ac.jp).*

Abstract: In this paper, a new reinforcement learning method is proposed to solve a train marshaling problem for assembling several outgoing trains simultaneously. In the addressed problem, the order of the incoming freight cars is assumed to be random. Then, the freight cars are classified into several sub-tracks. The cars on sub-tracks are rearranged to the main track by a certain desirable order. In the proposed method, each set of freight cars that have the same destination make a group, and the desirable group layout constitutes the best outgoing trains. When a rearrangement operation is conducted, the best number of sub-tracks used in the operation is obtained by a reinforcement learning system, as well as the best layout of groups in the trains, the best order to rearrange cars by the desirable order, and the best sub-track for the car to be removed. The marshaling plan that consists of series of removal and rearrangement operations are generated based on the processing time of movements of freight cars. The total processing time required to assemble outgoing trains can be minimized by the proposed method.

Keywords: Train Marshaling, Scheduling, Reinforcement Learning, Freight Car

1. INTRODUCTION

Freight train distribution plays an important role in terms of the sustainability in logistics, because railway distribution has a smaller impact on the environment than truck transportation (Li et al. (2007)). Although, freight cars cannot be delivered to areas without railways, flexible transportation is possible by selecting trucks and railways corresponding to infrastructure at the destination of freight. Freight trains consist of multiple freight cars, each carrying containers. Containers are commonly used to pack goods, and each container that makes up a freight train has its own destination. In intermodal transportation using both truck and rail, trucks bring containers into freight stations. Then, the containers are located on the freight cars in the order of arrival. In addition, the cars are also brought to the station by train, so that the initial layout of the freight cars is random. On the other hand, when constructing a outgoing train, the goal is to arrange freight cars with adjacent freight cars with the same destination in order to simplify the distributing procedure. If the arrangement of freight cars differs from the desired one in the outgoing train, relocate operation is required to freight cars. At this time, the total processing time required for relocation can be reduced by selecting the proper procedure in consideration of the initial placement and destination of the freight cars.

In the addressed problem, the marshaling process is conducted in a freight yard that consists of a main-track and several sub-tracks. In the process, freight cars are initially carried into the yard in the random order. Then, they are classified into sub-tracks, pulled out from sub-trackes in the desirable order, and lined on the main track in order to assemble an outgoing train. In this case, the number of arrangements of freight cars increases by the exponential rate with increase of total count of cars. Although similar problems are treated by mathematical programming and genetic algorithm (Blasum et al. (2000); Kroon et al. (2008); He et al. (2000); Dahlhaus et al. (2000); Jacob et al. (2007); Adlbrecht et al. (2015)), they do not consider road portages. Thus, conventional methods cannot apply directly to the addressed problem that assumes the initial order of incoming freight cars is random.

Recently, a reinforcement learning method to derive a marshaling plan for a single outgoing train based on the total processing time of marshaling has been proposed (Hirashima (2014)). The method can evaluate total processing time for a marshaling plan by autonomous learning methos.

In this paper, a new scheduling method is proposed in order to classify, rearrange and line freight cars by the desirable layout onto the main track. In the proposed method, the focus is centered on to reduce the total processing time to achieve desirable layout of multiple trains on the main track. The optimal arrangement of freight cars in the main track is derived based on the destination of freight cars, in order to minimize the total processing time. A position of the freight car to be moved is selected from several candidates, considering the destination of cars to be moved, based on the evaluation values assigned to each arrangement of cars on sub-track. Each evaluation value reflects the smallest processing time to achieve the best arrangement of cars in the outgoing train on the main track. The learning algorithm is enhanced based on the reinforcement learning (Watkins and Dayan (1992)).

In order to show effectiveness of the proposed method, computer simulations are conducted for two cases.

2. PROBLEM DESCRIPTION

2.1 Marshaling

In the addressed problem, the freight yard consist of 1 main track and m sub-tracks. Define k as the number of freight cars placed on the sub-tracks, and they are carried to the main track by the desirable order based on their destination. The freight cars are moved by a single power car. The power car is at the head of the train, the freight cars are carried into the yard from the tail by the random order, decoupled, and determined its location. When all the freight cars are located on one of the sub-tracks, the initial arrangement of the freight cars is obtained for marshalling. In the marshaling, a power car moves freight cars from sub-track to sub-track or from sub-track to main track. The movement of freight cars from sub-track to sub-track is called removal, and the car-movement from sub-track to main track is called rearrangement. Then, set $k \le n{\cdot}m-(n-1)$ to leave room for removals. For simplicity, the maximum number of freight cars that each sub-track can have is assumed to be n, the ith car is recognized by an unique symbol c_i $(i = 1, \cdots, k)$, and the number of sub-tracks is l. Marshaling is started from initial arrangement of the cars on the sub-tracks, and completed by moving k freight cars to the main track. The solution is a sequence of car movements during marshaling, and the sequence of car movements that minimizes the total processing time is the optimal solution. .

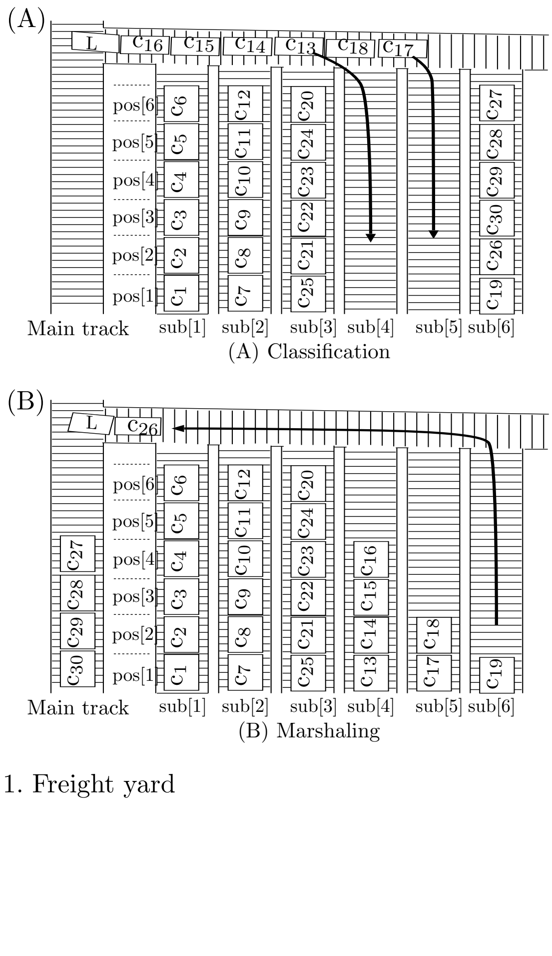

Fig. 1. Freight yard

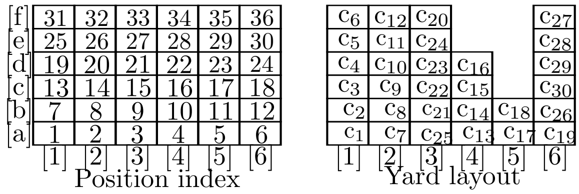

Fig. 2. Example of position index and car arrangement

2.2 Freight yard

Fig.1 shows the outline of freight yard in the case $k = 30, m = n = 6$. In the figure, a main track is at the left end, and other tracks sub[1]-sub[6] are sub-tracks. The main track is linked with sub-tracks by a joint track, which is used for moving cars between sub-tracks, or for moving them from a sub-track to the main track. In the figure, freight cars are moved from sub-tracks, and lined in the main track by the descending order, that is, rearrangement starts with c_{30} and finishes with c_1. When the power car L moves a certain car, other cars locating between the power car and the car to be moved must be removed to other sub-tracks. This operation is called removal.

In Fig.1(A), the arrangement of freight cars carried into the yard are random, and the cars are decoupled when they are located on a specific sub-tracks. In Fig.1(B), cars c_{30}–c_{27} are rearranged into the main track, and the car c_{26} is pulled out to be reaaranged into the main track.

In each sub-track, positions of cars are defined by n rows. Every position in sub-tracks has unique position number represented by $m \cdot n$ integers. The position number of cars located in the main track is 0. Fig.2 shows an example of position index for $m = n = 6$, and the yard layout according to Fig.1.

In Fig.2, the position located at row [a] in the sub-track [1] has the position number 1, and the position "[f][6]" has the position number 36. For unified representation of layout of car in sub-tracks, cars are placed from the row [a] in every track, and newly placed car is located at the adjacent position of the neighboring car. In the figure, in order to rearrange c_{25}, cars c_{20}, c_{21}, c_{23} and c_{24} in front of c_{25} have to be removed to other sub-tracks. Then, since $k \le n \cdot m - (n-1)$ is satisfied, c_{25} can be moved even when all the other cars are placed in sub-tracks.

In the freight yard, define $x_i (1 \le x_i \le n \cdot m, i = 1, \cdots, k)$ as the position number of the car c_i, and $\boldsymbol{s} = [x_1, \cdots, x_k]$ as the state vector of the sub-tracks. For example, in Fig.2, the state is represented by $\boldsymbol{s} = [1, 7, 13, 19, 25, 31, 2, 8, 14, 20, 26, 32, 3, 9, 15, 21, 4, 10, 6, 18, 24, 5, 30, 36, 12, 0, 0, 0, 0, 0]$. A trial of the rearrange process starts with the initial layout, rearranging freight cars according to the desirable layout in the main track, and finishes when all the cars are rearranged to the main track.

In this model, there are $k!$ freight car arrangements for one combination of freight cars on each sub-track, and the number of states increases at exponentially as linear increase of the car counts. In addition, since it is possible to transition between arbitrary 2 states, there is a possibility that arbitrary wagon arrangement will appear in the movement of freight cars from the initial arrangement to the desired arrangement.

2.3 Desirable arrangements

In the main track, freight cars that have the same destination are placed at one of adjacent positions. In this case, without additional operations are not required at the destination regardless of arrangement of these cars. In order to consider this feature in the desired arrangement in the outgoing train on the main track, a group is organized by cars that have the same destination, and these cars can be placed at any positions in the group. Then, a group is generated for each destination, and the order of groups lined as the outgoing train on the main track is predetermined by destinations. This feature yields several desired layouts in the main track.

Fig.3 depicts examples of desirable layouts of cars and the desired layout of groups in the main track. In the figure, freight cars $c_1, \cdots, c_6$ to the destination A make $group_1$, $c_7, \cdots, c_{14}$ to the destination B make $group_2$, and $c_{15}, \cdots, c_{22}$ to the destination C make $group_3$ $Groups_{1,2,3}$ are lined by ascending order in the main track, which make a desirable layout.

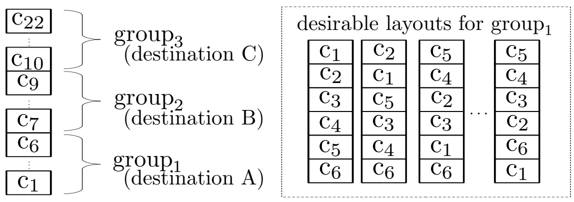

Fig. 3. Example of desirable layout and arrangement in a group

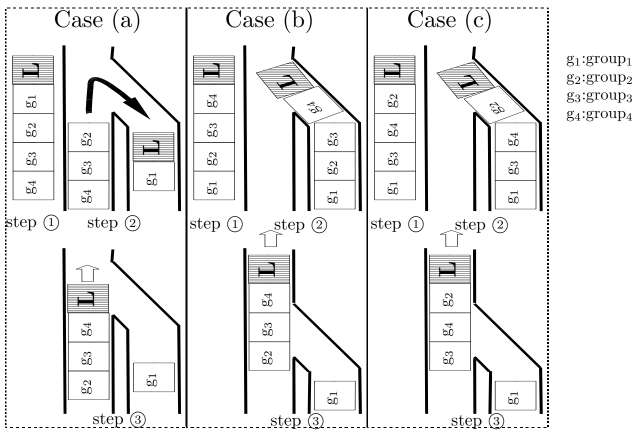

Fig. 4. Group arrangements

Also, the layout of groups lined by the reverse order do not yield additional removal actions at the destination of each group. Thus, in the proposed method, the layout lined groups by the reverse order is regarded as one of desired layouts, and the candidates of desired layouts are derived by the following:

(1) Determine the order in which each group is decoupled based on the destination,
(2) Determine whether each group is located in the formar half or the latter half of the train,
(3) Connect the formar half of the cars in ascending order and the latter of the cars in descending order,
(4) Connect the rear end of the formar half and the head of the latter half to obtain a desirable arrangement of a candidate.

In step1, if the group separation order is uniquely determined, the arrangement order of the first half and the second half determined in step2 is also uniquely determined. Then, all the candidates of desired layout can be extracted by a combination of whether each group belongs to the formar half or the latter half. However, for the two cars located at the end of the formar half and the head of the latter half, if one is determined, the other is uniquely determined. In other words, the number of candidates h is 2^{r-1}, where r is the total number of groups.

Fig.4 depicts examples of material handling operation for extended layout of groups at the destination of $group_1$. In the figure, cars in $group_1$ lined by the order shown in the step ① in case (a) are separated at the main track, and moved to a sub-track by the locamotive L at step ②. In cases (b)(c), cars lined by the layout shown in the step ① are carried in a sub-track, and $group_1$ is separated at the sub-track. Then, in cases (a)(b)(c), $group_1$ can be located without any removal actions for cars in each group. Thus, all the layouts of groups in the cases are regarded as candidate for desired one in the learning process of the proposed method.

2.4 Construction of multiple outgoing train

If there is enough time to construct multiple trains, marshaling of these trains can be couducted simultaneously in the same yard. Now, w trains are configured at the same time, and the ith train is represented by F_i $(i = 1, 2, \cdots, w)$. Then, in marshalling, both the train composition order on the main track can be ascending order $F_1, F_2, \cdots, F_w$ and descending order $F_w, F_{w-1}, \cdots, F_1$ can satisfy the train delivery order. In other words, if trains are constructed in ascending order, the trains are sent when each train configuration is completed on the main track. If trains are constructed in descending order, each train can be sent in a predetermined order after all trains are assembled by keeping the constructed trains on the main track. Therefore, $h = 2^r$ where h is the number of candidates of sequence that moves r groups from the sub-track into main track.

Fig.5 shows examples for $w = 2$. In the figure, dashed squares denote cars for a train2 (F_2) in sub-tracks, a rectagle printed F_2 is train2, and a rectagle printed F_1 is train1 in each example. In Fig.5a, ① a train F_1 is assembled first, ② departed before ③ assembling a train F_2, so that the construction order of trains is F_1, F_2. Also, in Fig.5b, a train is assembled in the order of F_2, F_1. That is, ① the construction process for F_2 starts first, the process for F_1 starts after the process for F_2 is finished, then, ③ F_1 is departed before F_2. In both cases, F_1, F_2 can be sent in order from the front of the main track. Since the two construction methods can generate different rearrangement orders for cars in the sub-track, the total processing time can be reduced by selecting the appropriate construction method. The selection is conducted by using a reinforcement learning method explained in section 5.

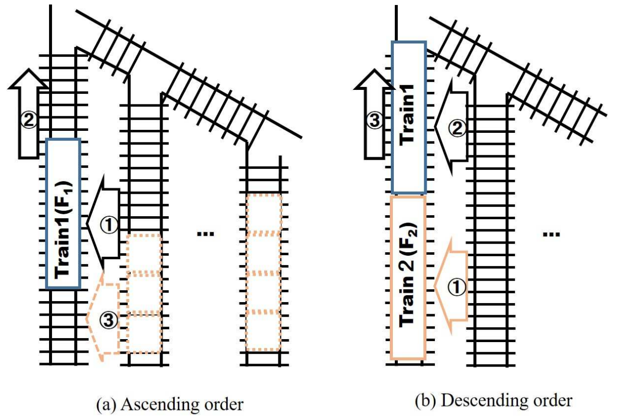

Fig. 5. Example of assembling order of trains

3. REARRANGEMENT PROCESS

When a rearranging car exists and it has no car to be removed in front of it, its rearrangement precedes any removals. In the case that several cars can be rearranged without a removal, rearrangements are repeated until all the candidates for rearrangement requires at least one removal. If several candidates for rearrangement require no removal, the order of selection is random, because any orders satisfy the desirable layout of groups in the main track. This operation is called direct rearrangement. When a car in a certain sub-track can be rearrange directly to the main track and several cars located adjacent positions in the sub-track satisfy the group layout of main track, they are jointed and applied direct rearrangement. The rearrangement process for cars consists of following 7 operations :

(I) selection of constraction order for trains and the layout of groups
(II) classification of the incoming freight cars into sub-tracks,
(III) rearrangement for all the cars that can apply the direct rearrangement into the main track,
(IV) selection of a freight car to be rearranged into the main track,
(V) selection of a removal destinations of the cars in front of the car selected in (IV),
(VI) selection of the number of cars to be moved in (V),
(VII) removal of the cars determined in (VI) to the sub-track selected in (V),

In the operation (I), each candidate for desirable arrangement on the main track including all the trains is defined as u_{j_1} $(1 \leq j_1 \leq h, h = 2^r)$. In the operation (II), candidates are sub-tracks where freight cars are carried in, and each candidate is represented by u_{j_2} $(h+1 \leq j_2 \leq h+m)$. Let u_{j_3} $(h+m+1 \leq j_3 \leq h+m+v_{\mathrm{g}})$ be a candidate for the number of freight cars to be located on the sub-track selected among u_{j_2}. In the operation (IV), each group has the predetermined position in the main track. Then, the car to be rearranged is defined as c_T, and candidates are determined by the number of freight cars that have already rearranged to the main track and the group layout in the main track. v_{O} is defined as the number of freight cars in group to be rearranged, and each candidate is represented by u_{j_4} $(h+m+v_{\mathrm{g}}+1 \leq j_4 \leq h+m+v_{\mathrm{g}}+v_{\mathrm{O}})$.

In the operation (V), the $m-1$ sub-tracks except for the sub-track where the cars to be removed exist are candidates for destination of removal. The sub-track index of each candidate is expressed as $u_{j_5}, h+2m+v_{\mathrm{g}}+v_{\mathrm{O}}+1 \leq j_5 \leq h+3m+v_{\mathrm{g}}+v_{\mathrm{O}}-1$. the removal destination of car located in front of the car to be rearranged is defined as r_{M}.

In the operation (VI), defining w_1 as the number of removal cars required to rearrange c_T, and defining w_2 as the number of removal cars that can be located on the sub-track selected in the operation (V), the candidates for the number of cars to be moved are determined by $1 \leq u_{j_7} \leq \min\{w_1, w_2\}, h+3m+v_{\mathrm{g}}+v_{\mathrm{O}} \leq j_6 \leq h+3m+v_{\mathrm{g}}+v_{\mathrm{O}}+\min\{w_1, w_2\}-1$.

4. TRANSFER DISTANCE OF POWER CAR

When a power car transfers freight cars, the process of the unit transition consists of following 6 elements: (E1) starts without freight cars, and reaches to the joint track, (E2) restarts in reverse direction to the target car to be moved, (E3) joints them, (E4) pulls out them to the joint track, (E5) restarts in reverse direction, and transfers them to the indicated location, and (E6) disjoints them from the powercar. Then, the transfer distance of a power car in (E1), (E2), (E4) and (E5) is defined as D1, D2, D3 and D4, respectively. Also, define the unit distance of a movement for cars in each sub-track as $D_{\min_v}$, the length of joint track between adjacent sub-tracks, or, sub-track and main track as $D_{\min_h}$. The location of the power car at the end of above process is the start location of the next movement process of the selected car. The initial position of the power car is located on the joint track nearest to the main track.

Fig.6 shows an example of transfer distance. In the figure, $m = n = 6$, $D_{\min_v} = D_{\min_h} = 1, k = 18$, (a) is position index, and (b) depicts movements of power car and freight car. Also, the power car starts from position 8, the target c_1 is located on the position "18", the destination of the target is "4", and the number of cars to be moved is "2". Since the power car moves without freight cars from "8" to "24", the transfer distance is $D_1 + D_2 = 12$ $(D_1 = 5, D_2 = 7)$, whereas it moves from "24" to "16" with 2 freight cars, and the transfer distance is $D_3 + D_4 = 13$ $(D_3 = 7, D_4 = 6)$.

4.1 Processing time for the unit transition

In the process of the unit transition, the each time for (E3) and (E6) is assumed to be the constant t_{E}. The processing times for elements (E1), (E2), (E4) and (E5) are calculated by using the transfer distance of the power car $D_i (i = 1, 2, 3, 4)$, the weight of the freight cars W moved in the process, and the performance of the power car. Then, the time each for (E1), (E2), (E4) and (E5) is assumed to be obtained by the function $f()$ derived considering dynamics of the power car, limitation of the velocity, and control rules. Thus, the processing time for the unit transition t_U is calculated by

$$t_U = t_{\mathrm{E}} + \sum_{i=1}^{2} f(D_i, 0) + \sum_{i=3}^{4} f(D_i, W) \qquad (1)$$

The maximum value of t_U is defined as $\mathrm{t_{max}}$ and calculated by

$$\begin{aligned}\mathrm{t_{max}} = \mathrm{t}_E + f(k\mathrm{D}_{\mathrm{min}_v}, 0) + f(m\mathrm{D}_{\mathrm{min}_h}, 0) \\ + f(m\mathrm{D}_{\mathrm{min}_h} + n, \mathrm{W_{max}}) \\ + f(k\mathrm{D}_{\mathrm{min}_v}, \mathrm{W_{max}}) \qquad (2)\end{aligned}$$

where, $\mathrm{W_{max}}$ is the largest weight as n freight cars.

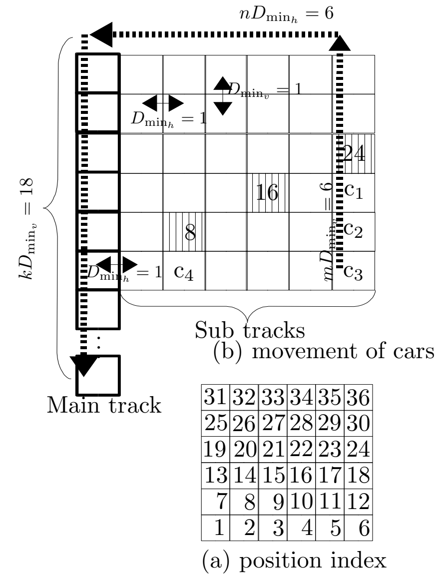

Fig. 6. Calculation of transfer distance

5. LEARNING ALGORITHM

Now, constraction order for trains is defined as m_c. Desired pair of layout of groups in the main track described by G_O and m_c is selected among candidates u_{j_1} $(1 \leq j_1 \leq h)$. Then, evaluation value for $(G_\mathrm{O}, m_\mathrm{c})$ is defined by $Q_1(G_\mathrm{O}, m_\mathrm{c})$ that is updated by the following equation when one of desired layout is achieved in the main track:

$$Q_1(G_\mathrm{O}, m_\mathrm{c}) \leftarrow \max \left\{ \begin{array}{c} Q_1(G_\mathrm{O}, m_\mathrm{c}), \\ (1-\alpha) Q_1(G_\mathrm{O}, m_\mathrm{c}) + \alpha V_1 \end{array} \right\} \quad (3)$$

$$V_1 = \mathrm{R} \prod_{i=1}^{l} \gamma_i \quad (4)$$

where l denotes the total movement counts required to achieve the desired layout, α is learning rate, γ_i is discount factor calculated for each movement, R is reward that is given only when one of desired layout is achieved in the main track.

Define r_M as the sub-track selected as the destination for the removed car, p_M as the movement counts of freight cars, and $\boldsymbol{s}'$ as the state that follows $\boldsymbol{s}$. In the classification stage, Q2,Q3 are defined as evaluation values for $(\boldsymbol{s}_\mathrm{a}, u_{j_2})$, $(\boldsymbol{s}_\mathrm{b}, u_{j_3})$ respectively, where $\boldsymbol{s}_\mathrm{a} = [\boldsymbol{s}, (G_\mathrm{O}, m_\mathrm{c})]$, $\boldsymbol{s}_\mathrm{b} = [\boldsymbol{s}, r_\mathrm{M}, (G_\mathrm{O}, m_\mathrm{c})]$. When one of desired layout is achieved in the main track, Q_1, Q_3 are received the reward, and Q_2, Q_3 are updated by following rules:

$$Q_2(\boldsymbol{s}_\mathrm{a}, r_\mathrm{M}) \leftarrow \max_{u_{j_3}} Q_3(\boldsymbol{s}_\mathrm{b}, u_{j_3}) \quad (5)$$

$$Q_3(\boldsymbol{s}_\mathrm{b}, p_\mathrm{M}) \leftarrow (1-\alpha) Q_3(\boldsymbol{s}_\mathrm{b}, p_\mathrm{M}) + \alpha V_2$$

$$V_2 = \begin{cases} \mathrm{R} \prod_{i=1}^{l} \gamma_i \\ \text{(all cars assigned)} \\ \gamma \max_{u_{j_2}} Q_2(\boldsymbol{s}_\mathrm{a}, u_{j_2}) \\ \text{(otherwise)} \end{cases} \quad (6)$$

In the marshaling, Q_4, Q_5 and Q_6 are defined as evaluation values for $(\boldsymbol{s}_\mathrm{a}, u_{j_4})$, $(\boldsymbol{s}_\mathrm{a}, u_{j_5})$, $(\boldsymbol{s}_\mathrm{c}, u_{j_6})$, respectively, where $\boldsymbol{s}_\mathrm{c} = [\boldsymbol{s}, \mathrm{c}_T]$, $\boldsymbol{s}_\mathrm{d} = [\boldsymbol{s}, \mathrm{c}_T, r_\mathrm{M}]$. Q_4, Q_5, Q_6 are updated by following rules:

$$Q_4(\boldsymbol{s}_\mathrm{a}, \mathrm{c}_T) \leftarrow \max_{u_{j_5}} Q_5(\boldsymbol{s}_\mathrm{c}, u_{j_5}), \quad (7)$$

$$Q_5(\boldsymbol{s}_\mathrm{c}, r_\mathrm{M}) \leftarrow \max_{u_{j_6}} Q_6(\boldsymbol{s}_\mathrm{d}, u_{j_6}), \quad (8)$$

$$Q_6(\boldsymbol{s}_\mathrm{d}, p_\mathrm{M}) \leftarrow \quad (9)$$

$$\begin{cases} (1-\alpha) Q_6(\boldsymbol{s}_\mathrm{d}, p_\mathrm{M}) + \alpha \left[\mathrm{R} + V_3 \prod_{i=1}^{q} \gamma_i \right] \\ \quad (u \text{ is a rearrangement}) \\ (1-\alpha) Q_6(\boldsymbol{s}_\mathrm{d}, p_\mathrm{M}) + \alpha [\mathrm{R} + \gamma V_4] \\ \quad (u \text{ is a removal}) \end{cases}$$

$$V_3 = \max_{u_{j_4}} Q_4(\boldsymbol{s}'_\mathrm{a}, u_{j_4}), V_4 = \max_{u_{j_5}} Q_5(\boldsymbol{s}'_\mathrm{c}, u_{j_6})$$

where, q is the number of direct movements conducted sequentially. γ_i is used to reflect the total processing time of marshaling into evaluation values and calculated by the following equation: γ_i $(i = 1, \cdots, l)$

$$\gamma_i = \delta \frac{\mathrm{t_{max}} - \beta t_U}{\mathrm{t_{max}}}, \ 0 < \beta < 1, 0 < \delta < 1 \quad (10)$$

5.1 Action selection

Each u_j, $(1 \leq j \leq\leq h + 3m + v_\mathrm{g} + v_\mathrm{O} + \min\{w_1, w_2\} - 1)$ is selected by the Soft-Max selection method(Watkins and Dayan (1992)). In the proposed method, Q_i $(i = 1, \cdots, 6)$ are normalized into $\tilde{Q}_i$ $(i = 1, \cdots, 6)$, and probability P for selection of each candidate is calculated as follows(Hirashima (2014)):

$$\tilde{Q}_i(\boldsymbol{s}_i, u) = \frac{Q_i(\boldsymbol{s}_i, u) - \min_{u_x \in u_{j_i}} Q_i(\boldsymbol{s}_i, u_x)}{\max_{u_x \in u_{j_i}} Q_i(\boldsymbol{s}_i, u_x) - \min_{u_x \in u_{j_i}} Q_i(\boldsymbol{s}_i, u_x)},$$

$$P_i(\boldsymbol{s}_i, u) = \frac{\exp(\tilde{Q}_i(\boldsymbol{s}_i, u))/\xi}{\sum_{u_x \in u_{j_i}} \exp(\tilde{Q}_i(\boldsymbol{s}_i, u_x))/\xi}$$

$$(i = 1, 2, 3, 4, 5, 6, 7) \quad (11)$$

where ξ is a thermo constant.

6. COMPUTER SIMULATIONS

Computer simulations are conducted for $m = 12, n = 6$, $k = 15$ for train1 and $k = 6$ for train2 in order to compare learning performaces of the following 2 methods:

(A) Proposed method to minimize the total processing time for trains 1，2,

(B) A method to minimize the processing time train 1 and train 2 separately(Hirashima (2014)).

The initial arrangement of incoming train is described in Fig.7. The original group layout for train 1 is $\mathrm{group}_1, \mathrm{group}_2, \mathrm{group}_3, \mathrm{group}_4$, and for train 2 is $\mathrm{group}_5, \mathrm{group}_6$. Members in each group are set as follows: Cars $c_1, \cdots, c_5$ are in group_1, $c_6, \cdots, c_9$ are in group_2, $c_{10}, \cdots, c_{13}$ are in group_3, $c_{14}, \cdots, c_{15}$ are in group_4, $c_{16}, \cdots, c_{18}$ are in group_5, and $c_{19}, \cdots, c_{21}$ are in group_6. Other parameters are set as $\alpha = 0.9, \beta = 0.2, \delta = 0.9, \mathrm{R} = 1.0$, $\xi = 0.1$, $D_{\min_v} = D_{\min_h} = 20\mathrm{m}$.

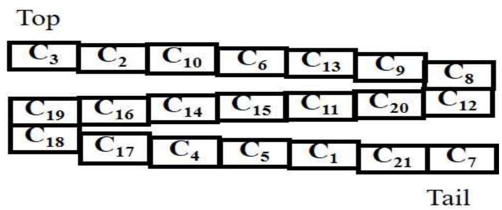

Fig. 7. Initial arrangement of cars for 2 Trains

The power car assumed to accelerate and decelerate the train with the constant force 100×10^3N and to be 100×10^3kg in weight. Also, all the freight cars have the same weight 10×10^3kg. The velocity of the power car is limited to no more than 10m/s. Then, the power car accelerates the train until the velocity reaches 10m/s, keeps the velocity, and decelerates until the train stops within the indicated distance D_i $(i = 1, 2, 3, 4)$. When the velocity does not reach 10m/s at the half way point of D_i, the power car starts to decelerate immediately.

The results are shown in Fig.8. In the figure, horizontal axis shows the number of trials and the vertical axis shows the minimum processing time found in the past trials. Each result is averaged over 20 independent simulations.

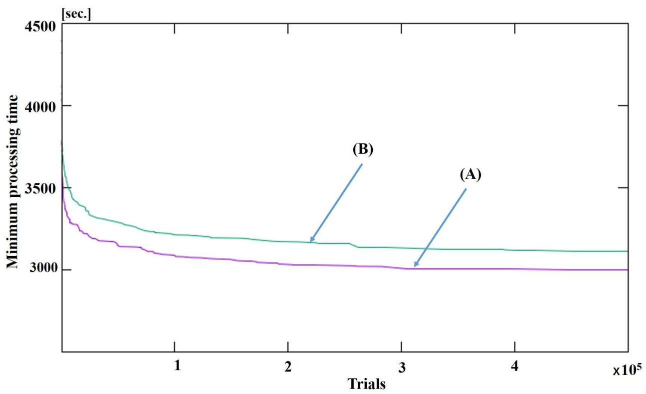

Fig. 8. Performance comparison

In the figure, the total processing time for (A) is smaller as compared to that for (B). Since (A) optimizes multiple trains at the same time, the arrangement of cars on subtrack is better for marshaling of train 2 when a marshaling for the train 1 is completed. In other words, (B) spoils initial arrangement of marshaling for train 2 in order to minimize the processing time of marshaling for train 1.

Table 1. Total Processing Time

methods	processing time (sec.) best	average	worst
method (A)	2897.208	2999.473	3092.410
method (B)	2976.529	3109.618	3316.142

Total processing time for each method at the 1.0×10^6th trial are described in Table 1 for each method.

7. CONCLUSIONS

A new method for deriving a marshaling plan of freight cars considering multiple trains has been proposed. In the proposed method, the total processing time for classification and marshaling is evaluated and the learning algorithm based on the reinforcement learning is designed. In order to reduce the total processing time, the proposed method yields classification and marshaling plan for multiple trains simultaneously, and computer simulations show effectiveness of the plan generated by the proposed method as compared to the repetitive conduction of the optimization method for single train. Moreover, the layout of multiple trains, the arrangements of cars in the classification, the rearrange order of cars, the position of each removal car, the number of cars to be removed, and the group layout in the outgoing train has been learned autonomously so that the learning performance of the proposed method has been improved.

REFERENCES

Adlbrecht, J.A., Hüttler, B., Zazgornik, J., and Gronalt, M. (2015). The train marshalling by a single shunting engine problem. *Transportation Research Part C: Emerging Technologies*, 58, 56–72.

Blasum, U., Bussieck, M.R., Hochstättler, W., Moll, C., Scheel, H.H., and Winter, T. (2000). Scheduling trams in the morning. *Mathematical Methods of Operations Research*, 49(1), 137–148.

Dahlhaus, E., Manne, F., Miller, M., and Ryan, J. (2000). Algorithms for combinatorial problems related to train marshalling. *Proceedings of Australasian Workshop on Combinatorial Algorithms 2000*, 7–16.

He, S., Song, R., and Chaudhry, S. (2000). Fuzzy dispatching model and genetic algorithms for railyards operations. *European Journal of Operational Research*, 124(2), 307–331.

Hirashima, Y. (2014). A new reinforcement learning for train marshaling with generalization capability. *Applied Mechanics and Materials*, 974, 269–273.

Jacob, R., Marton, P., Maue, J., and Nunkesser, M. (2007). Multistage methods for freight train classification. *7th Workshop on Algorithmic Methods and Models for Optimization of Railways*, 158–174.

Kroon, L., Lentink, R., and Schrijver, A. (2008). Shunting of passenger train units: an integrated approach. *Transportation Science*, 42, 436–449.

Li, G., Muto, M., Aihara, N., and Tsujimura, T. (2007). Environmental load reduction due to modal shift resulting from improvements to railway freight stations. *Quarterly Report of RTRI*, 48(4), 207–214.

Watkins, C.J.C.H. and Dayan, P. (1992). Q-learning. *Machine Learning*, 8, 279–292.