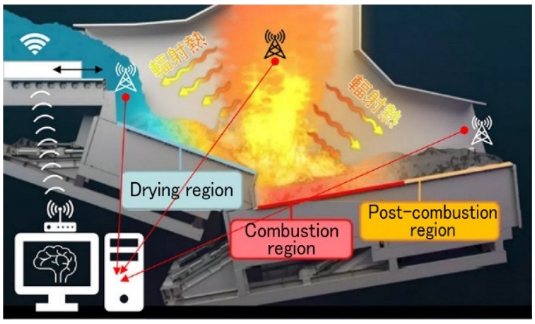
Feasibility Study of Effectiveness of Control Technology for Fuel Feeding Stabilization in Stoker Furnace Using Reinforcement Learning



TAKASHI IKEDA[*1] SHUNYA SASAKI[*2]
YUKI BABA[*2] TOSHIHIKO SETOGUCHI[*3]

Waste incineration plants are expected to contribute to the realization of the appropriate energy mix. In order to increase the profitability of such plants, automatic operation technology is needed. Since the fuel supply of waste incineration plants is greatly affected by variations in fuel characteristic, a prospective control technology for the fuel feeding device has not been established. We have developed a fuel feeding control technology based on reinforcement learning, which is expected to enable flexible selection of appropriate operational controls considering changes in fuel characteristic. The use of this technology on a fuel supply device simulator resulted in a reduction of 56% in the number of over-supply and 54% in the number of under-supply compared to the conventional operating conditions. This report presents the development of this technology.

1. Introduction

Waste incineration plants are products that contribute significantly to the energy diversification, decentralization, and increased self-sufficiency toward the achievement of the appropriate energy mix, and their automatic operation technology to improve plant profitability by reducing the load on workers and running costs through the improvement in efficiency is required. On the other hand, there is an issue that fluctuations in the fuel feeding due to variations in the fuel characteristics of waste cause uncontrollable combustion or extinction in the furnace. This requires manual intervention by operators to prevent harmful undesirable emissions such as carbon monoxide and nitrogen oxides that are generated unintentionally.

We have developed a reinforcement learning-based fuel feeding control technology for fuel feeding system for waste incineration plants. This technology is expected to enable flexible selection of appropriate operational controls considering short-term changes in fuel characteristics due to climate, etc. and long-term changes in fuel properties due to seasonal changes, etc. This report presents the development of this technology and the result of verifying the possibility to reduce fluctuations in fuel feeding.

2. Waste incineration stoker furnace and fuel feeding system

Waste incineration stoker furnaces are one of the most practical waste utilizing system in the world, in which waste is fed into the incinerator (stoker) by a table-shaped reciprocating fuel supply device. In the stoker, waste is conveyed downstream (toward the furnace end) while being mixed by the movable grate, and is burned when it comes into contact with air blown in from the bottom (**Figure 1**). The stoker is divided into a drying region, a combustion region, and a post-combustion region from upstream, where drying process, a pyrolysis process, and a char combustion process, respectively, occur in series. The fuel feeding system (feeder) of a waste incinerator is installed upstream of the stoker, below the hopper of waste and reciprocates hydraulically to feed fuel intermittently to the stoker. The controlling parameters of the fuel feeding system include the feeder stroke, feeder speed, and stop time between its advance and retraction of the feeder. Waste in the hopper is compressed by its own weight. But variations in fuel moisture content due to seasonal,

*1 Chief Staff Engineer, Business Intelligence & Innovation Department, Technology Strategy Office
*2 EPI Department, Digital Innovation Headquarters
*3 Chief Engineer, Combustion Research Department, Research & Innovation Center

climatic, and regional differences cause variations in the amount of fuel delivered, even for the same operating parameters. It is also caused by variations in shape and adhesion due to differences in the type of fuel. Therefore, it is difficult to conduct the control based on a predetermined rule, and a control technology for fuel feeding system that can flexibly follow the operating conditions with short-term and long-term changes in fuel characteristics is needed.

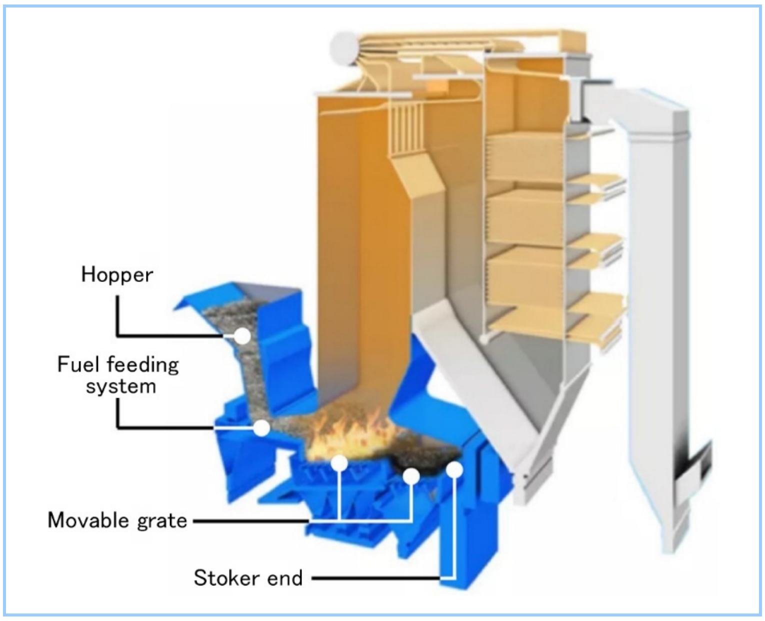


Figure 1 Waste incineration stoker furnace

3. Development of reinforcement learning-based fuel feeding system control technology

The most basic algorithm of reinforcement learning is the Q-learning, in which the action to be taken for the current state is determined based on the Q-table, which shows the expectation reward obtained by learning. In order to determine appropriate actions with the Q-learning, it is necessary to improve the comprehensiveness of the combinations of states and actions in the Q-table. In terms of the future deployment in actual equipment, we applied the Deep-Q-Network[(1)], a reinforcement learning method that predicts the Q-table using a neural network was applied to control the fuel feeding system. The Deep-Q-Network learns the results of a control with the neural network and reflects them in the Q-table. Those contribute to compensate for excess or deficiency of learning data for constructing the Q-table, and is expected to be able to update continuously the Q-table to fit the latest environmental circumstance.

In this report, we verified the feasibility of the Deep-Q-Network to the fuel feeding system using a simulator that reconstructs the behavioral characteristics of a fuel feeding system. In the simulator, which is two-dimensional the amount of fuel by discretizing the space from the hopper to the fuel feeding port is calculated, and it's fulfilling the mass balance between the discretized inspection volumes. The simulator considers compression and shear by the fuel feeding system. The variation in adhesion due to changes in characteristics, and consequently the fluctuations in the fuel feeding rate, which is the control target of the reinforcement learning shown in this paper, can be qualitatively considered in the simulator.

Table 1 shows the conditions of the reinforcement learning. To achieve appropriate control using reinforcement learning, it is necessary to set the appropriate observation subjects that define the current state of the control target and rewards and penalties for evaluation values obtained as a result of any action. The observation subjects are specified so that they can be directly and indirectly measured by image analysis, actual sensors, etc., from the view point of the future application to actual equipment. The action for control changes depending on the observation subjects, rewards, and penalties. So, we conducted verification in the three cases where the observation items, rewards, and penalties are set as shown in **Table 2**. These three cases were set as ones expected to provide acceptable control based on preliminary examinations.

Table 1 Setting parameters of reinforcement learning

Item	Value
NN intermediate layer	3
Number of NN intermediate layer nodes	16
NN activation function	Relu (intermediate layer), Identity function (output layer)
NN optimization function	Adam method
ε-greedy random rate	1.0 (start) → 0.1 (finish)
Discount factor	0.99

Table 2 Observation subjects and rewards design for verification cases

Case No.	Observation information	Reward for feeding *1	Penalty for over-feeding *2	Penalty for over-feeding *3
Case 1	Fuel supply Current stroke	+5	-1	0
Case 2	↑	+1	0	0
Case 3	↑	+1	0	-1

*1: When the supply per stroke is within ±10% of the target value, the reward is given.
*2: [Rate of the number of times of under-supply (supply less than a half of the target value) to 5 strokes] x [Penalty value] is given. (When the number of times of under-supply per 5 strokes is 5, the penalty of -1 is given.)
*3: When the supply per stroke is greater than the target value +20%, the penalty is given.

4. Effectiveness verification result of fuel feeding system and its discussion

We verified the effectiveness of the fuel feeding system for a relatively small waste incineration stoker furnace in MHI. In order to evaluate over-feeding and under-feeding of fuel, this verification counted the number of over-feeding, where the fuel injection volume was more than 10% of the target supply, and the number of under-feeding, where the fuel injection volume was less than 50% of the target volume, were evaluated against the target fuel injection volume of 90 t/d. In this verification, the control target was the stroke of the fuel feeding system, which is considered to have the largest impact on the fuel feeding. For the objective of comparison, conventional operating conditions with the feeder stroke set constant value were used. The changes in fuel characteristics were considered by changing the density and the adhesion in the hopper. The former simulates the different types of waste (fuel), while the latter simulates the different moisture content of the waste. They were changed every 10 times of reciprocation of the fuel feeding system, as shown in **Figure 2**. Because fuel is fed from the waste pit into the hopper by cranes in waste incinerators, this verification assumes that the density and adhesion change each time the fuel is fed by the crane. In order to exclude the influence of the initial stage of the calculation, the evaluation period was set from the timing when the fuel feeding work 100 to 150 stokes.

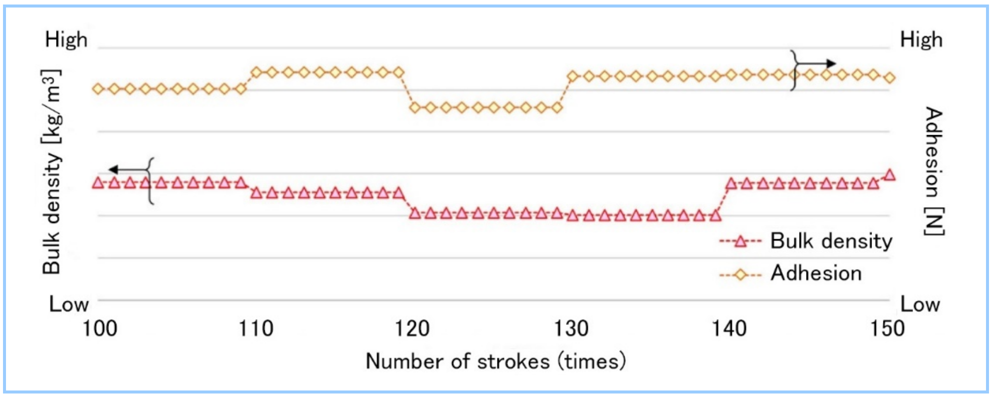


Figure 2 Changes in fuel properties given for calculation

Figure 3 shows the result of feeding fuel amount per stroke in Case 3. **Table 3** shows the integrated results of the number of over-feeding and under-feeding during the evaluation period. The all results for the three cases in Table 2 shows similar acceptable control results in terms of reducing the numbers of over-feeding and under-feeding. Figure 3 shows Case 3 as an example, which showed the best results of feeding control among the three cases. While under the conventional operating conditions with the stroke, over-feeding and under-feeding are repeated as shown in Figure 3, in Case 3, where the Deep-Q-Network is applied, the stroke is changed by the control with the Deep-Q-Network every 10 stroke(i.e., at the timing when the density and adhesion were changed). As a result of the fuel feeding control with the Deep-Q-Network, the number of times of over-supply is reduced by up to 56% and the number of times of under-supply is reduced by up to 54% compared to the conventional operating conditions.

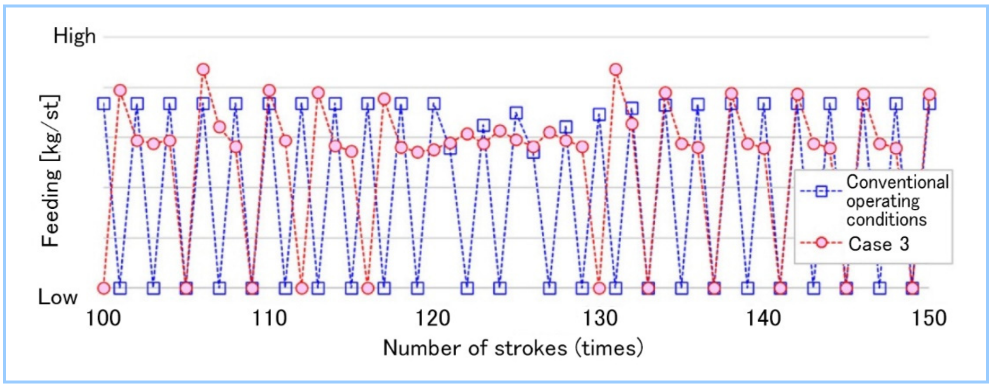


Figure 3 Supply per stroke

Table 3 Integrated results of the number of over- supply and under-supply

	Number of under-feeding	Number of over-feeding
Existing operating conditions	25	22
case 1	13 (-48%)	13 (-41%)
case 2	13 (-48%)	14 (-36%)
case 3	11 (-56%)	10 (-54%)

Figures in parentheses are the reduction ratio when the existing operating conditions are set to 100%.

The evaluation of the amount of change in the stroke of the fuel feeding system every 10 strokes which was the controlled by the Deep-Q-Network in Case 3 in response to the change in density and adhesion is comparable with the fuel feeding control by operator. Although the conditions of this study were simple, these results indicate the practicality of Deep-Q-Network to the fuel feeding system of incinerator. This shows that an appropriate fuel supply control device is able to be realized with the Deep-Q-Network.

In the case of actual incinerators, there are a large number of complicate factors that affect the variation in fuel feeding. Therefore, comprehensive verification and evaluation for managing to these complicate factors may require an enormous amount of time and resource. The consumption of them is also undesirable from the standpoint of obtaining the market. In order to advance the control accuracy presented in this report, it is needed to acquire extra data based on operation in actual furnace. For the future development to market this technology, we will set a minimum target which was required by the market for developing in order to achieve quick implementation of the technology in actual furnace, and it should be update the technology within the actual operation in actual furnace.

5. Conclusion

For waste incineration plants, the development of an automatic operation technology to reduce the workload of operators and achieve highly efficient operation is needed. For this, we are developing a control technology for fuel feeding system that reduce fluctuations in the fuel feeding. We have developed a reinforcement learning-based control technology, which is a promising control technology that allows fuel feeding system to flexibly follow operating conditions that are difficult to control based on a rule and involve short-term and long-term changes in fuel characteristics. Through verification using a simulator, we confirmed that this reinforcement learning is effective with up to 56% reduction in the number of over-supply and 54% reduction in the number of times of under-supply compared to the conventional operating conditions.

Waste incineration plants are products that contribute significantly to the energy diversification, decentralization, and increased self-sufficiency toward the achievement of the appropriate energy mix, and our competitors are also developing technologies continuously in an aggressive manner.

References

(1) Volodymyr Mnih, et al., Human-level control through deep reinforcement learning, Nature 518, 529-533.