International Journal of Naval Architecture and Ocean Engineering 12 (2020) 910–917

Contents lists available at ScienceDirect

International Journal of Naval Architecture and Ocean Engineering

journal homepage: http://www.journals.elsevier.com/international-journal-of-naval-architecture-and-ocean-engineering/

Application of reinforcement learning to fire suppression system of an autonomous ship in irregular waves

Eun-Joo Lee [a], Won-Sun Ruy [b, *], Jeonghwa Seo [b]

[a] Graduate School, Dept. of Naval Architecture & Ocean Engineering, Chungnam National University, Republic of Korea
[b] Department of Naval Architecture & Ocean Engineering, Chungnam National University, Republic of Korea

ARTICLE INFO

Article history:
Received 7 June 2020
Received in revised form
13 September 2020
Accepted 1 November 2020

Keywords:
Automatic fire suppression
Reinforcement learning
Position based fluids

ABSTRACT

In fire suppression, continuous delivery of water or foam to the fire source is essential. The present study concerns fire suppression in a ship under sea condition, by introducing reinforcement learning technique to aiming of fire extinguishing nozzle, which works in a ship compartment with six degrees of freedom movement by irregular waves. The physical modeling of the water jet and compartment motion was provided using Unity 3D engine. In the reinforcement learning, the change of the nozzle angle during the scenario was set as the action, while the reward is proportional to the ratio of the water particle delivered to the fire source area. The optimal control of nozzle aiming for continuous delivery of water jet could be derived. Various algorithms of reinforcement learning were tested to select the optimal one, the proximal policy optimization.

1. Introduction

As demand of autonomous ships is increasing, the relevant researches on the system design and core technology have been conducted (Sin, 2018; de Vos and Hekkenberg, 2019; DeFilippo et al., 2019; Fraize et al., 2019; Lee, 2019; Tang et al., 2019; Ko and Lee, 2019). One of the core technologies is autonomous damage control. Following the classification of autonomous ships by the International Maritime Organization (IMO), autonomous ships of autonomous degree 3 and 4 could be operated with absence of crews on board (Lee, 2019). For such degrees of autonomous operation, the automatic response to the maritime accidents or disasters should be provided (IMO, 2018). In particular, fire accident is the most crucial event on the safety of ships, and the autonomous detection and suppression of fire is quite important in the design and operation of unmanned vessels of degree 3 and 4.

According to a study on the application of automatic firefighting systems and fire suppression strategy, the success of firefighting depends on how fast the initial responses are taken (Yoon, 2019). There have been various studies on precise fire suppression in the initial stage, including the fire suppression using a large-scale monitor set outside (Miyashita et al., 2014) or the firefighting robot for short-distance fire suppression in concealed environments (McNeil and Lattimer, 2017).

In this study, a wall-mounted nozzle was selected as the fire extinguishing system for initial fire suppression. It can precisely deliver water to the fire source, while sprinklers, spraying water to an entire compartment and will result in a malfunction of on-board equipment in the compartment, and the ship cannot carry out continue operations. In addition, CO_2 extinguishers contain a clean extinguisher agent that does not leave behind any harmful residue or cause secondary damage to machines. It is sufficient for unmanned compartment without damage to sailors. However, the capability of onboard CO_2 generator and the number of fire suppression operation is limited, thus the system is applicable when the fire grew seriously, and a different approach is needed for initial fire suppression.

As an alternative to sprinkler and CO_2 extinguisher for initial fire suppression, the fire extinguishing system studied in this paper should be applied to various types of fire that can happen on sailing. For the wall-mounted nozzles, the extinguishing material can be selected from seawater and foam as generic fire suppression monitors on board. Also, it can locally deliver fire suppression material to the fire source, thus it is proper for initial suppression system.

In case of large-scale outdoor fire suppression, the effect of the wind is significant and the aiming should be corrected by the wind

* Corresponding author.
E-mail addresses: sik5646@naver.com (E.-J. Lee), wsruy@cnu.ac.kr (W.-S. Ruy), j.seo@cnu.ac.kr (J. Seo).
Peer review under responsibility of Society of Naval Architects of Korea.

https://doi.org/10.1016/j.ijnaoe.2020.11.001
2092-6782/

condition. In case of the fire suppression in ship compartments, however, the concealed compartment environment leads to restrict effect of wind, but the trajectory of water for fire suppression method focused on the six-degree-of-freedom motion of the compartment owing to the wave-induced force.

Aiming the nozzle in waves is a complex problem, and continuous aiming correction is needed. As the delivery of the water to the fire source has time delay from the aiming correction, prediction of the trajectory change of the water spray by the ship motion is required. For such prediction in complicated physical system, a numerical simulation of multi-body dynamics is applicable (Unity, 2020).

According to Kim et al. (2019), reinforcement learning is relatively superior to the Proportional Integral Differential (PID) controller in responding to unexpected environments and can make optimal control policies without any assumptions and dynamic modeling. Therefore, this study simulated the fire suppression case in a ship compartment moving under the sea condition and examined the correction of aiming of the monitor through the proximal policy optimization (Beck and Woolf, 1998; Juliani et al., 2018). Later, the study compared the trained model to the untrained fire suppression model with the initial aiming angle under various sea environments.

2. Fire suppression modeling and simulation of a ship compartment motion under the sea condition

2.1. Compartment modeling

To describe the compartment arrangement, a body-fixed Cartesian coordinate system was used. The origin was located at the intersection of FP and the baseline. x-, y-, and z-direction is defined as the longitudinal, port, and upward direction. The compartment model was realized by utilizing Unity's internal 3D objects. Its size is set to $13.40(\mathrm{W}) \times 21.30(\mathrm{L}) \times 8.50(\mathrm{H})$ m, and its location is

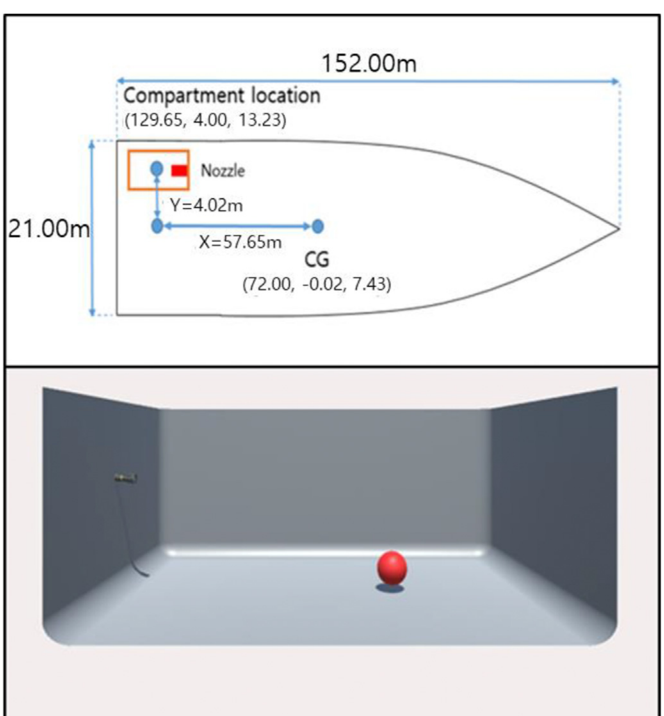

Fig. 1. Compartment's location and modeling.

(130.00, 4.00, 13.23), which is (57.65, 4.02, 5.80) apart from the mass center of the ship (72.00, −0.02, 7.43). Fig. 1 shows the location of the compartment an nd modeling of the compartment in Unity 3D.

2.2. Compartment motion

Following the interface standard for shipboard systems, the ship motion in waves was estimated. Major motions were surge, heave, roll, and pitch, and the magnitude and period with respect to the sea state is suggested (Department of Defense, 1986). As it suggests the maximum motion and modal period of the motion, a motion spectrum should be derived from the motion criteria. The spectrum of ship motion in waves (Perez, 2006) are discretized for generation of time-history of the motion, as shown in Eqs. (1) and (2). In the present study, the number of discretized domains, n, is set to 10.

$$Angular\ motion: \sum_{i=1}^{n} A_i \sin(\omega_i t + \phi_i) \tag{1}$$

$$Translational\ motion: \sum_{i=1}^{n} \frac{a_i}{{\omega_i}^2} \sin(\omega_i t + \phi_i) \tag{2}$$

$A =$ *Maximum anglular motion amplitude*

$a =$ *Maximum linear acceration amplitude*

$\omega =$ *angle velocity*

$\phi =$ *phase*

2.3. Modeling of water spray

To model water jet from the extinguishing nozzle, the study used Obi Asset in Unity3D (Obi user manual, 2020). Obi Assets can realize the fluid motion by defining the physical features of fluid through the particles based on the position-based fluids (Macklin and Müller, 2013), so that they could model high viscosity fluid, gas, or fresh water. This study used the property value defined as fresh water in Obi library to model the water droplets sprayed from nozzles.

2.4. Nozzle modeling

To model the exit water flow of the nozzle, the study referred to the actual nozzle dimension (Elkhart Brass, 2020). The nozzle was located at the center of the transverse wall of the compartment. It can rotate vertically and horizontally, from −180° to 180°. Note that 0° means the nozzle is aligned along the x-direction. The designed flow rate at the operation condition was 0.078 m^3/s, and the outer and inner diameters of the ring-type nozzle exit were 0.063 m and 0.05 m, respectively. Considering the baffle disk to be located in front of the nozzle, 70% of the total outlet area was allocated as the effective water spraying area. Thus, the effective area and diameter of the nozzle was 0.0013 m^2 and 0.041 m, respectively. The exit speed of water was 57 m/s according to Eq. (3).

$$v = Q/A = 57\ [m/s] \tag{3}$$

where, v, Q, and A means the flow velocity, flow rate, and effective area of the nozzle exit, respectively.

The water spray trajectory simulated by the modeling was validated by the trajectory database of the nozzle (SM-1250 SERIES REACH, 2020). The modeling considered the atmospheric drag terms by Obi library. The value of atmospheric drag was set to be 3.9, to make similar trajectory of the water. Fig. 2 shows the comparison of the modeled trajectory to the actual trajectory.

In the simulation, the number of particles determines the simulation resource requirements. Therefore, two approaches were used to reduce the number of water particles while keeping accuracy. The first approach was to reduce the water particles at the center of the nozzle exit. By distributing the water particles around the edge rather than disk, the boundary of the water jet trajectory can be kept. Fig. 2 shows comparison of water jets of disk and ring type. Under the same condition, the number of particles to be analyzed in disk (370,500) was 13 times greater than that of edge (28,500). Second, the lifetime of tracer particle is limited. It means that the particle motion is not tracked for the whole simulation time, but till the arrival at the compartment floor. It was set to disappear 0.01 s later after reaching the target, so as to reduce the number of active particles in the simulation domain.

By reducing simulation resource, multiple simulation can be conducted at the same time. Asynchronous Advantage Actor-Critic (A3C) (Mnih et al., 2016) is the method of learning asynchronously by multiple actor runners. Using this method, more stable learning is possible without bias. As a result, A3C can be learned with six models at the same time.

2.5. Initial aiming angle

To set the initial angle of the nozzle aiming, the horizontal and vertical angles need to be determined. The y- and z-directional rotational angle of the nozzle was initially set, assuming parabolic trajectory without air resistance. Eq. (4) cited from Miyashit et al. (2014) shows the water's trajectory without air resistance. Using this equation, the study set the initial vertical angle of the nozzle by examining the angle satisfying the equation based on the locations of fire and nozzle. Brent's Search method (Brent, 1973) was used to examine the nozzle angle, which determines θ in Eq. (4).

$$h = -\frac{\rho g}{4P\cos^2\theta} l^2 + \tan\theta l + h_0 \tag{4}$$

$h =$ *height of water discharge trajectory* $[m]$

$h_0 =$ *nozzle height* $[m]$

$\theta =$ *discharge angle* $[rad]$

$l =$ *distance from nozzle* $[m]$

$P =$ *discharge pressure* $[Pa]$

3. Reinforcement learning

Reinforcement learning (Juliani et al., 2018) is one area of the machine learning methods where an agent under a certain environment selects a behavior or behavioral order to maximize the current reward. The agent in reinforcement learning is the component that makes the decision of what action to take and learn the policy by itself.

3.1. Reinforcement learning condition

To reflect the practical motion of ship in irregular waves, the phase lag, ϕ_i and sea-state of the motion equation at the beginning of each episode were randomly set between 0–2π and 4–8, respectively. Also, the location of the target was randomly located in the compartment during the training process. The initial learning process progresses by taking random actions under such random environment. Each episode is set to 5000 steps so as to seek the maximum cumulative reward within 100 s. A total of six models were used along with A3C to maintain a more stable learning process than the one with only one model. Fig. 3 indicates the learning process after the final modeling. Table 1 shows the parameters required for the learning. The total duration for the learning was 3 h 40 min.

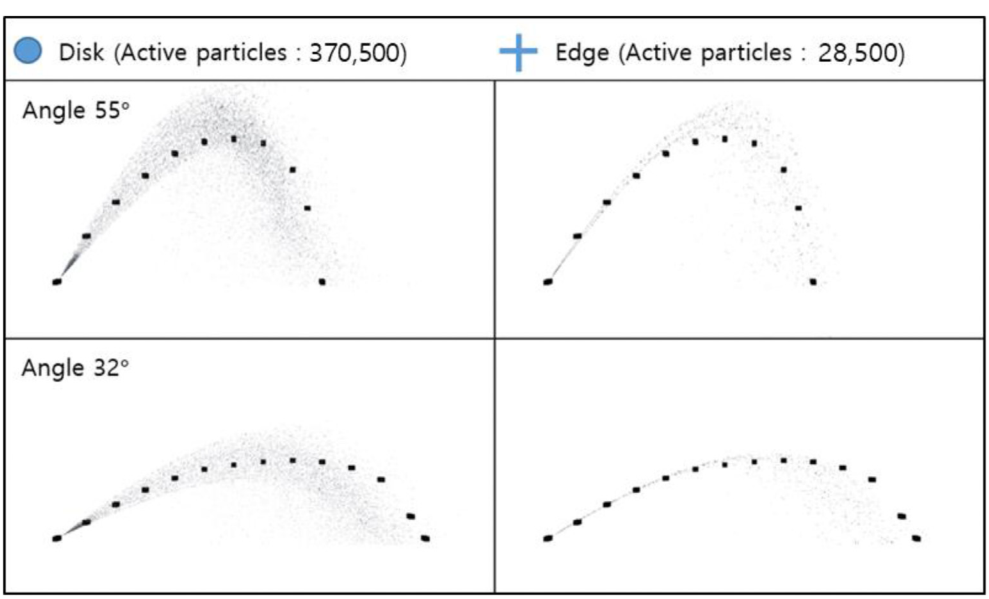

Fig. 2. Comparison with actual trajectory.

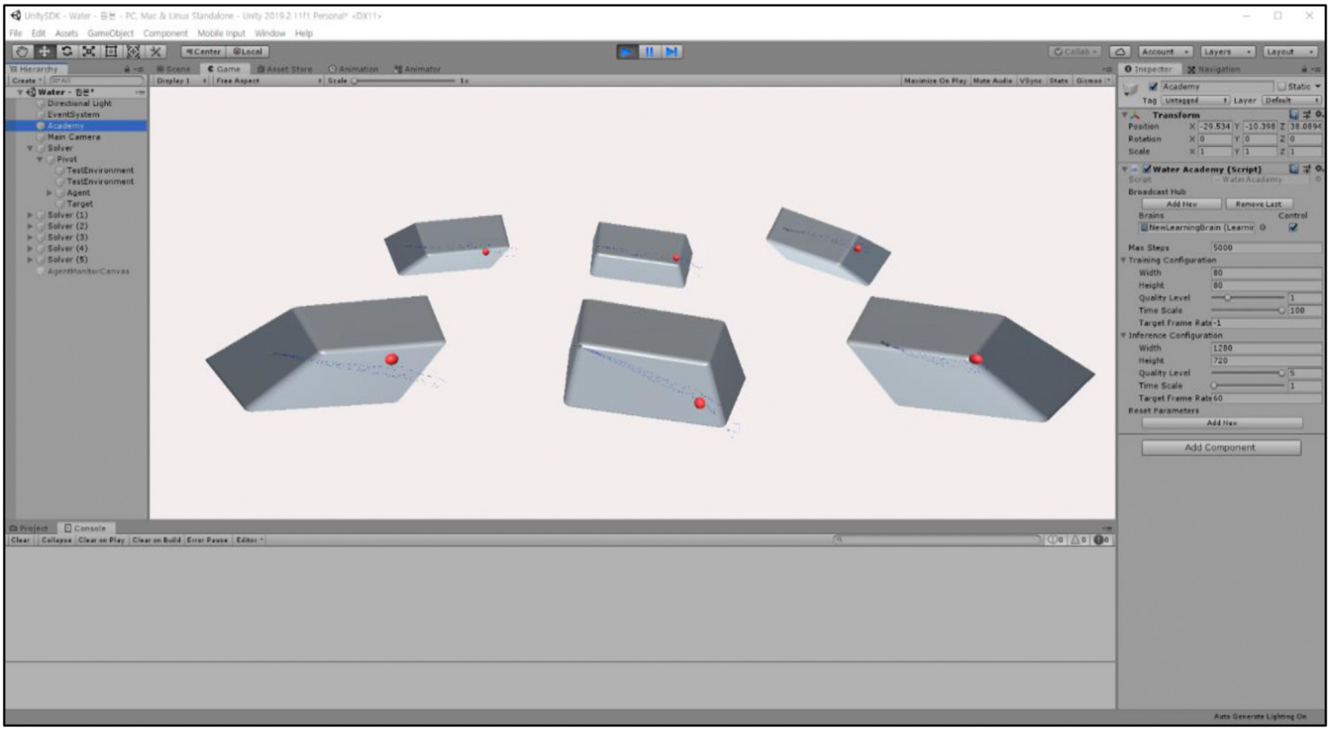

Fig. 3. Training environment of reinforcement learning.

Table 1
Simulation parameters.

Batch size	1024
Beta	0.001
Buffer size	10240
Learning rate	3.0e-4
Max step	1.0e6
Time horizon	64
Sequence length	64
Summary frequency	10000

3.2. State

In reinforced learning, state is defined as the currently observable information. In ship dynamics, the translational motion is usually measured by accelerometers. The rotational velocity is acquired by the gyroscope. It was assumed that the gyroscope and accelerometer were located 1 m away from the mass center of the ship in the lateral direction for the measurement.

In addition, the location of the target object, initial aiming angle, and the current angle of the nozzle can be acquired. Thus, as for the state, the speed and acceleration of roll, heave, and pitch, the initial aiming angle of the nozzle, and the current angle of the nozzle were determined as the state.

3.3. Action

The action is the behavior that the agent can take under a certain state. In the present study, the action is aiming control. An agent at the initial stage of learning does not have the information about whether a certain action is beneficial or not. Therefore, the agent at first takes a series of random actions, and later the probability that the agent would take a specific action will increase as the learning continues. The continuous value between $-2°$ and $2°$ at each step was selected for the action so that it could rotate vertically and horizontally. It was assumed that there is no delay in operation of the monitor motion by control input.

3.4. Reward

The reward is the only evaluative information available for the agent to learn. Through a cumulated reward value, the agent can evaluate the actions taken and determine which actions are beneficial. The goal of the reinforcement learning is to find the policy in which the agent maximizes the cumulated reward value in each environment. In the present study, the reward is estimated by the number of reached water particles to the fire region. Eq. (5) shows the formula for calculating the reward value in this study. A summed turning value of particle number is needed to separate penalties and reward.; if it is below the summed turning value of particle number, 10 particles per frame, a penalty should be given. Later, 0.01 was multiplied to normalize the reward value. On average, it was realized that 100 water particles per frame reached to the fire source.

$$R = (P - 10) \times 0.01 \tag{5}$$

Where, R = *reward*, P = *number of the particles reaching the target*

4. Analysis of the results

Before discussion of the reinforced learning results, the outcome was summarized. The criteria for determination of success or fail status of reinforcement learning listed below are referred from ML-Agent (ML-Agents, 2020).

- Environment/Cumulative Reward: The average reward value that the agent acquired from the episode.
- Environment/Episode Length: The average episode duration of all agents in each environment.

- Policy/Value Estimate: The average estimated value of all states that the agent visited, which should be increased for a successful learning process.
- Losses/Value Loss-value function: The mean loss of the value function update. It correlates to how well the model can predict the value of each state. This might fluctuate while the agent is learning due to lack of training data, and then decrease once the reward stabilizes.

Fig. 4 shows the learning progresses. The environment/cumulative reward stably increases and converges, which means that the reinforced learning improves fire suppression performance. In addition, environment/episode length means the duration of the action by the agent, which shows that the agent gradually completes the episode. The gradual increase of policy/value estimate shows that the learning has been progressed successfully. The successful learning reduces Losses/value loss.

To verify the learning level of the reinforced trained model, the present study compared the results of fire suppression with and without control by the reinforced learning. The former is called trained model and the latter is called untrained model. For comparison of two models, the number of tracer water particles reaching the target for 5000 steps and the time without delivery of water particle were measured. The time where the water particles fail to reach the target object was obtained from the number of frames based on the ‘Fixed Update’ function in Unity3D, which was executed 50 times per second.

4.1. Comparison with several different sea-state

During the simulation, the time that water was not delivered to the fire (t_U) and the number of the particle reached the fire region (n_R) were recorded. The location of the fire source was (20, 0, −2) from the fire monitor. For comparison, t_U and n_R in calm water were first derived. t_U in calm water was 0.46 s, which implies the time to the initially projected water jet reaches the fire region. n_R in calm water was 403,641 on average.

To verify the result, the sea-state condition and ϕ were randomly generated. The location of the target object was fixed on (20, 0, −2) from the nozzle, and three randomly generated time-history of ship motion for a sea-state were applied. The sea state ranged from 4 to 8.

Table 2 shows simulation results comparison by sea-states. In sea-state 4, t_U was 0.46 s for trained and untrained models, and the fire-fighting water continuously reached the fire source. The total sum of the average number of particles was improved by 7% from 384,144 to 409,319. n_R in both untrained and trained model decreased in high sea-states, but the reduction of untrained model was remarkable.

Fig. 5 shows the summary of simulation results. As the sea-state increases, n_R of the untrained model decreased while t_U increased. It implies that the fire suppression performance significantly degrades for untrained model in rough seas. In the trained model, however, the decrease of n_R was far smaller than that by the untrained model and the fire-fighting water could continuously reach the fire source.

4.2. Comparison by the location variation of the fire source

By fixing the sea-state and ϕ, the fire suppression performance change by the location variation of the fire source was examined. The simulation condition was selected from one case of sea-state 8, where the difference between the trained and untrained models was largest. Based on the location of the nozzle, a total of 15 locations (5 x-axial locations [2 m, 5 m, 10 m, 15 m, 20 m] × 3 y-axial locations [−5 m, 0 m, 5 m]) were compared. The height of the fire

Table 2
Comparison by sea-state.

Sea state		Untrained model	Trained model	Improvement
4	t_U	0.46 s	0.46 s	0.00 s (0%)
	n_R	384,144	409,319	25,175 (7%)
5	t_U	1.15 s	0.46 s	0.69 s (60%)
	n_R	345,890	404,539	58,649 (17%)
6	t_U	7.93 s	0.46 s	7.47 s (94%)
	n_R	289,907	400,013	110,106 (38%)
7	t_U	20.87 s	0.46 s	20.41 s (98%)
	n_R	245,625	387,789	142,164 (58%)
8	t_U	52.18 s	0.46 s	51.72 s (99%)
	n_R	136,000	343,294	207,294 (152%)

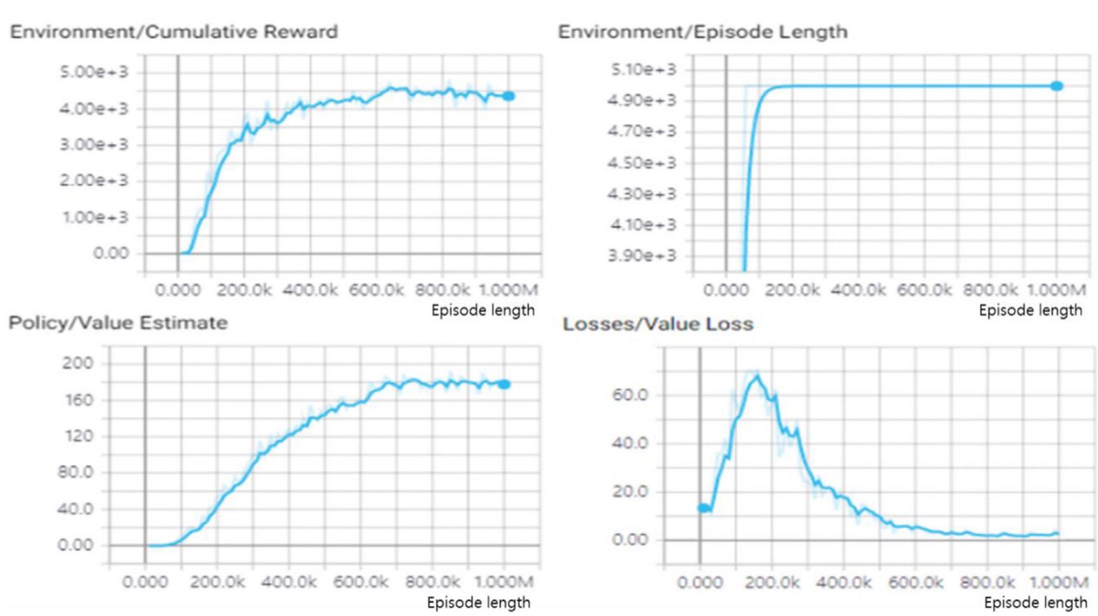

Fig. 4. Summary statistics of the learning process.

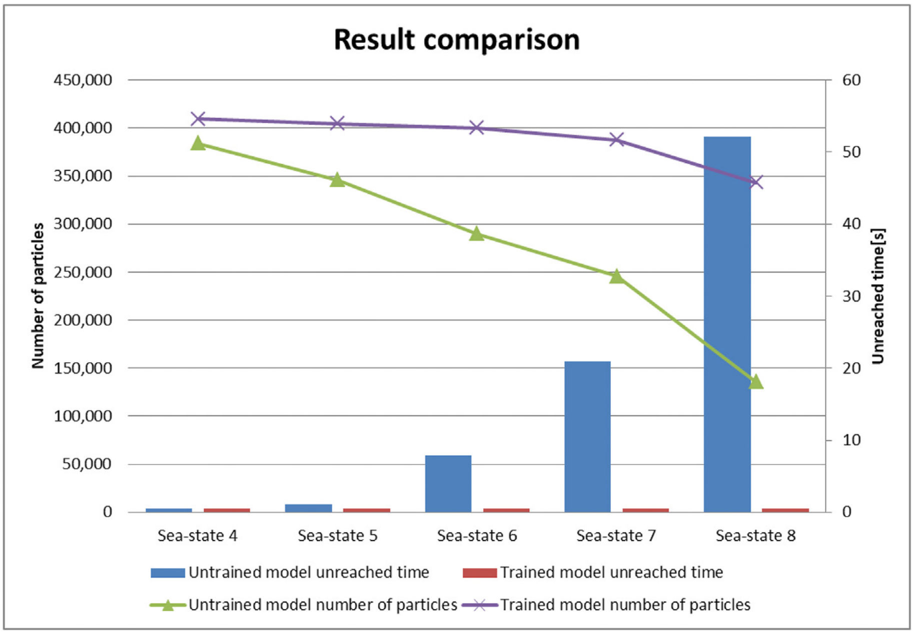

Fig. 5. Comparison of the results with several different sea-state.

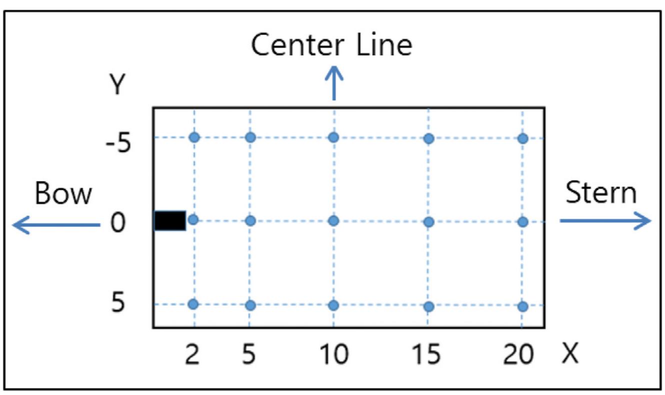

Fig. 6. Compartment coordinates.

source was assumed to be on top of a facility (2 m). Fig. 6 shows the locations of the fire sources in the compartment.

Table 3 shows comparison of n_R between the trained and untrained models. The improvement of water delivery by trained control model increased as the location moves in x-direction and further away from the midship. When the location of the fire source is at (2, 0, −2), n_R showed no meaningful difference between the untrained model (533,791) and the trained model (533,811). However, at (20, 5, −2) which was the furthest away from the nozzle and ship longitudinal center line, the number of particles reaching the fire source was improved by 267% by the trained model.

Table 4 summarizes t_U the fire source using the trained and untrained models. Same as n_R, t_U grew as the distance in x-direction increased. In the case of the untrained model, if the target location is (20, 5, −2), for 59.8 s out of 100 s, the fire-fighting water particles could not reach the target object. On the contrary, in trained model, the t_U was 0.52 s, similar to the baseline, 0.46 s.

Table 3
Comparison of the number of reached particles.

X [m]		Y [m]		
		5	0	−5
2	Untrained model	478,399	533,791	484,595
	Trained model	552,771	533,811	524,172
	Improvement	74,372 (16%)	20 (0%)	39,577 (8%)
5	Untrained model	406,931	485,188	438,585
	Trained model	574,405	565,771	556,187
	Improvement	167,474 (41%)	80,583 (17%)	117,602 (27%)
10	Untrained model	250,922	311,931	311,117
	Trained model	580,358	574,756	545,157
	Improvement	329,436 (131%)	262,825 (84%)	234,040 (75%)
15	Untrained model	154,751	185,570	198,887
	Trained model	497,109	549,487	484,441
	Improvement	342,358 (221%)	363,917 (196%)	285,554 (144%)
20	Untrained model	96,296	113,679	126,749
	Trained model	353,129	423,260	396,085
	Improvement	256,833 (267%)	309,581 (272%)	269,336 (212%)

Table 4
Comparison of unreached time.

X [m]		Y [m]		
		5	0	−5
2	Untrained model	0.12 s	0.06 s	0.14 s
	Trained model	0.12 s	0.06 s	0.14 s
	Improvement	0.00 s (0%)	0.00 s (0%)	0.00 s (0%)
5	Untrained model	5 s	0.12 s	2.62 s
	Trained model	0.16 s	0.12 s	0.16 s
	Improvement	4.84 s (97%)	0.00 s (0%)	2.46 s (94%)
10	Untrained model	22.76 s	11.14 s	9.64 s
	Trained model	0.26 s	0.24 s	0.28 s
	Improvement	22.50 s (99%)	10.90 s (98%)	9.36 s (97%)
15	Untrained model	42.44 s	34.18 s	28.44 s
	Trained model	0.38 s	0.36 s	0.44 s
	Improvement	42.06 s (99%)	33.82 s (99%)	28.00 s (98%)
20	Untrained model	59.8 s	51.02 s	46.78 s
	Trained model	0.52 s	0.46 s	0.56 s
	Improvement	59.28 s (99%)	50.56 s (99%)	46.22 s (99%)

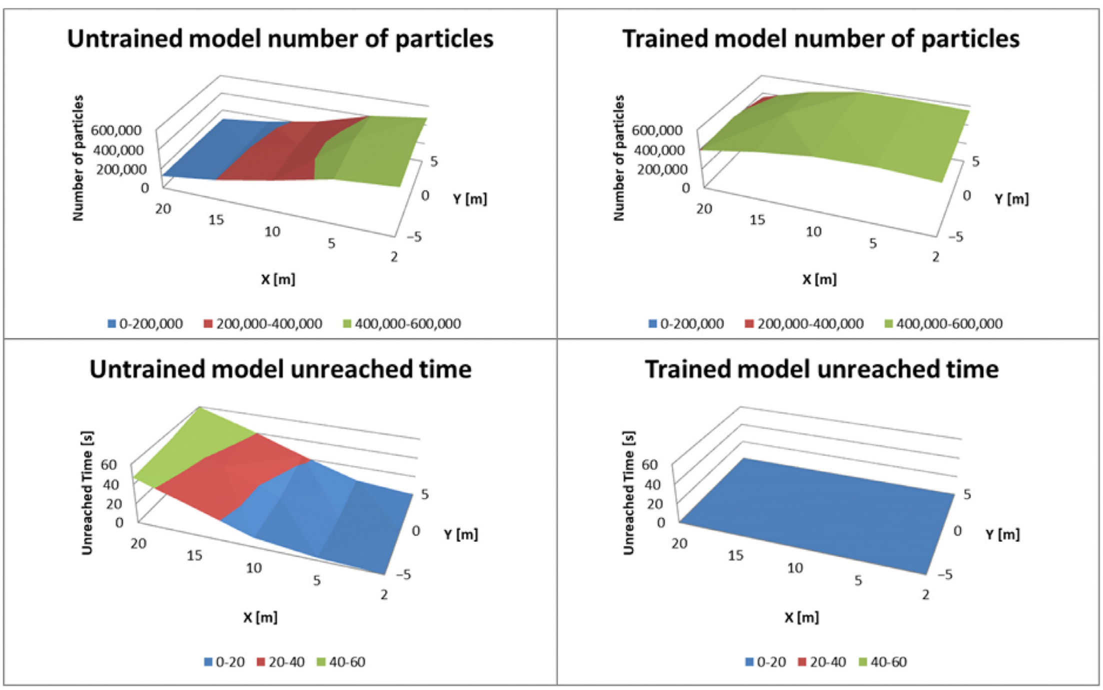

Fig. 7. Comparison of trained and non-trained model: n_R and t_U.

Fig. 7 displays the surface graphs of n_R and t_U. It was intuitively determined that as the location of the fire source moves further away from the nozzle, the differences between the trained and untrained models in the number of particles reaching the target object, as well as the time at which the water cannot reach the target, increased. Thus, it can be determined that the trained model can play a role under rough sea-states.

5. Conclusion

This study concerns the development of a fire suppression solution based on reinforcement learning for application to the compartment of a ship with six-degree-of-freedom motion in irregular waves. Using Obi library during the simulation, the behavior of the fire-fighting water was modeled as closely as in reality. The re-aiming steps of the nozzle was improved through reinforcement learning.

For the state, the study set the location of the fire source, the initial and current angles of the fire monitor aiming, and speed and acceleration of the compartment. A reward was given by scaling the degree at which the number of fire-fighting water particles reaching the target object, and through the action, the nozzle was set to find the target object. It was determined that the learning progressed smoothly when the action was set to be able to move between $-2°$ and $2°$ in angular position. In the case of the untrained model, the number of particles reaching the fire source was decreased, and the time at which not even one particle reached the target object increased in rougher sea state condition. By the reinforcement trained model, the fire-fighting water reached to the fire source continuously regardless of the sea-state and the location of the fire source. The limitation of this approach is that this study did not reflect the operation time delay of in monitor control dynamics. Further research will examine a more practical model of monitor control by considering the delay in operation time of the monitor.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This research was financially supported by the Institute of Civil Military Technology Cooperation funded by the Defense Acquisition Program Administration and Ministry of Trade, Industry and Energy of Korean government under grant No. UM19304RD3.

References

Beck, J.E., Woolf, B.P., 1998. Using a learning agent with a student model. In: International Conference on Intelligent Tutoring Systems. Springer, pp. 6–15.
Brent, R.P., 1973. Chapter 4: an Algorithm with Guaranteed Convergence for Finding a Zero of a Function. Algorithms for Minimization without Derivatives. Prentice-Hall, Englewood Cliffs, NJ, ISBN 0-13-022335-2.
de Vos, J., Hekkenberg, R., 2019. Assessment of the required subdivision index for autonomous ships based on equivalent safety. Assessment 17–18.
DeFilippo, M., Robinette, P., Sacarny, M., Benjamin, M.R., 2019. The remote explorer IV: an autonomous vessel for oceanographic research. In: OCEANS 2019-Marseille, pp. 1–8.
Department of Defense, 1986. Interface Standard for Shipboard Systems: Ship Motion and Attitude (Metric), DOD-STD-1399(NAVY) Section 301A. Naval Sea Systems Command, Washington, DC.
Elkhart brass X-Stream, 2020. http://www.elkhartbrass.com/products/master-stream/x-stream/multimedia. (Accessed 13 March 2019).
Fraize, J., Ekman, F., Fürth, M., Hoffenson, S., Chell, B., 2019. Multi-objective reliability-based design optimization of an autonomous sailing vessel. In: 29th International Ocean and Polar Engineering Conference. International Society of Offshore and Polar Engineers.
IMO, 2018. Maritime safety committee (MSC) 100th session.
Juliani, A., Berges, V.P., Vckay, E., Gao, Y., Henry, H., Mattar, M., Lange, D., 2018. Unity: A General Platform for Intelligent Agents arXiv preprint arXiv 1809.02627.
Kim, J.B., Hwang, K.Y., An, C.H., Han, Y.H., 2019. Research trends on PID control and reinforcement learning control an drone. J. Adv. Technol. Res. 4 (1), 6–13. ISSN: 0974-6269.
Ko, E.S., Lee, M.H., 2019. A study on the legal status of unmanned maritime systems. Korean J. Mil. Art Sci. 75, 89–118.
Lee, G.I., 2019. Standardization trend of IEC for self-operating ships. In: Proceedings of the Korean Institute of Navigation and Port Research Conference. Korean Institute of Navigation and Port Research, pp. 222–224.
Macklin, M., Müller, M., 2013. Position based fluids. ACM Trans. Graph. 32 (4), 1–12.
McNeil, J.G., Lattimer, B.Y., 2017. Robotic fire suppression through autonomous feedback control. Fire Technol. 53, 1171–1199.
Miyashita, T., Sugawa, O., Imamura, T., Kamiya, K., Kawaguchi, Y., 2014. Modeling and analysis of water discharge trajectory with large capacity monitor. Fire Saf. J. 63, 1–8.
ML-Agents, 2020. https://github.com/Unity-Technologies/ml-agents/. (Accessed 12 May 2020).
Mnih, V., Puigdomenech, A., Badia, Mirza, M., Graves, A., Lillicrap, T.P., Harley, T., Silver, D., Kavukcuoglu, K., 2016. Asynchronous Methods for Deep Reinforcement Learning. ArXiv preprint arXiv:1602.01783.
Obi user manual, 2020. URL. http://obi.virtualmethodstudio.com/tutorials/.
Perez, T., 2006. Ship Motion Control: Course Keeping and Roll Stabilisation Using Rudder and Fins. Springer Science & Business Media.
Sin, I.S., 2018. Special Report-Trends in standardization of autonomous transport ships. TTA J 36–42.
SM-1250 SERIES REACH, 2020. URL. https://www.elkhartbrass.com/files/aa/downloads/performance/SM-1250%20SERIES%20REACH@HEIGHT.pdf.
Tang, H., Yin, Y., Shen, H., 2019. A model for vessel trajectory prediction based on long short-term memory neural network. J. Mar. Eng. Technol. 1–10.
Unity. https://docs.unity3d.com/Manual/index.html. (Accessed 13 January 2020).
Yoon, M.O., 2019. Expansion of the application of autonomous firefighting facilities and fire suppression strategy improvement plan for the reduction of fire damage. J. Kor. Inst. Fire Sci. Eng.