Regulating control of in-pipe intelligent isolation plugging tool based on adaptive dynamic programming

Xingyuan Miao

China University of Petroleum, Beijing

Hong Zhao (✉ hzhao_cn@163.com)

China university of Petroleum, Beijing

Original Article

Keywords: Exoskeleton arm teleoperation, Magnetic drive pump, Pneumatic force-feedback, Hybrid fuzzy control

Posted Date: August 23rd, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-810332/v1>

License: This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Version of Record: A version of this preprint was published at Journal of Pipeline Systems Engineering and Practice on May 1st, 2022. See the published version at [https://doi.org/10.1061/\(ASCE\)PS.1949-1204.0000635](https://doi.org/10.1061/(ASCE)PS.1949-1204.0000635).

Title page

Regulating control of in-pipe intelligent isolation plugging tool based on adaptive dynamic programming

Xing-Yuan Miao, born in 1996, is currently a PhD student at *College of Mechanical and Transportation Engineering, China University of Petroleum, Beijing, China*. He received his bachelor degree from *China University of Petroleum, Beijing, China*, in 2019. His research interests include in-pipe robotics, deep reinforcement learning, electromechanical system design and automation.
Tel: +86-15027650591; E-mail: miaoxingyuan96@126.com

Hong Zhao, born in 1974, is currently a Professor in *China University of Petroleum, Beijing, China*. She received her Ph.D. degree from *Xi'an Jiaotong University, Xi'an, Shannxi Province, China*, in 2003. Her research interests include mechanical and electrical transmission control, virtual instrument measurement and control technology.
E-mail: hzhao_cn@163.com

Corresponding author: Hong Zhao E-mail: hzhao_cn@163.com

ORIGINAL ARTICLE

Regulating control of in-pipe intelligent isolation plugging tool based on adaptive dynamic programming

Xing-Yuan Miao¹ • Hong Zhao^{1*}

Received June xx, 2021; revised February xx, 2021; accepted March xx, 2021

© Chinese Mechanical Engineering Society and Springer-Verlag Berlin Heidelberg 2021

Abstract: In-pipe intelligent isolation plugging tools (IPTs) are crucial in pipeline maintenance. During the plugging process, the flow field around the IPT changes drastically, resulting in vibration and instability of the plugging process. Therefore, three foldable spoilers were designed at the tail of the IPT to reduce the vibration of the IPT. First, a disturbing flow experiment of IPT with spoilers was designed. A mathematical model of the pneumatic spoiler control system was established to regulate the spoiler angles. Second, based on the experimental data, a Bi-LSTM (bidirectional long short-term memory) neural network predictor between the plugging states, the spoiler angles, and the pressure gradient was established. Then, an adaptive dynamic programming controller was designed to select the optimal control action for each plugging state, thereby reducing the pressure gradient. Finally, Python and Matlab/Simulink were used for simulation. The results showed that the controller could reduce the pressure gradient during the plugging process by an average of 25.94%, which alleviated the vibration of the IPT and achieved a smooth plugging operation.

Keywords: Exoskeleton arm teleoperation • Magnetic drive pump • Pneumatic force-feedback • Hybrid fuzzy control

1 Introduction

Pipeline transportation has become the main transportation method for oil, natural gas, and many petrochemical products [1]. The oil and gas pipelines that show corrosion and leakage have been in service for a long time, creating significant safety hazards [2]. Therefore, pipeline maintenance work is essential. Pressure hole-opening and plugging technology is widely used in land

pipeline maintenance [3]. This method requires opening up a hole in the pipe section to be replaced, which reduces the efficiency of transportation and fails to adapt to the harsh marine environment. In-pipe intelligent plugging technology can repair pipeline leaks without interrupting transportation, greatly improving work efficiency. In addition, this approach can effectively complete the replacement of pipelines and valves, for the maintenance requirement of submarine pipelines [4].

The American company TDW and the British Stats Group are the leaders in in-pipe intelligent plugging. The IPTs are designed to complete high-pressure plugging operations with remote communication capability, allowing for use in submarine environments. Angus designed a spherical IPT that can partially plug a pipeline [5]. Li created a double-sealing IPT [6] and optimized the vital components to improve the working performance of the IPT [7-8]. Zhang developed a new type of spherical double-sealing IPT [9], along with a connection mechanism [10], control system, and communication system. The trajectory and accuracy of the rotation process were studied [11-12]. Zhao studied the changing laws of the flow field during the plugging process of the IPT [13]. The response surface method was used to optimize the shape and size of the end face of the IPT [14-5], thereby improving the flow field. In addition, the pneumatic plugging method was proposed [16]. Wu designed a velocity-tracking control system for an IPT [17], in which energy recovery could be carried out during the plugging process. The reinforcement learning method was adopted to improve energy efficiency [18].

✉ Hong Zhao
hzhao_cn@163.com

¹ College of Mechanical and Transportation Engineering, China University of Petroleum, Beijing, 102249, China

A long short-term memory (LSTM) neural network has significant advantages in processing time series, but can only access a specific point in time. The relationship between the current and future moments is not involved. A bidirectional LSTM (Bi-LSTM) can solve the problem of long-term dependence, and can process time-series data in forward and reverse directions simultaneously. This model can provide past and future information for each moment in a time series. In addition, more data features can be extracted to improve the accuracy of the prediction rule. Zeng proposed a fault prediction method based on a Bi-LSTM neural network [19], according to the data collected by aero-engine sensors. Compared with other prediction methods for time series, the error was reduced by 33.58%, achieving high prediction accuracy. Kang used a Bi-LSTM neural network to build a prediction model for the remaining service life of rolling bearings [20], which improved the convergence speed and prediction accuracy of the model. Zhou proposed a real-time human posture recognition method based on a Bi-LSTM neural network [21]. The Bi-LSTM neural network was used as a classifier to recognize the occluded and unobstructed situations, improving the recognition accuracy. Therefore, the Bi-LSTM neural network has great advantages in predicting the flow field parameters around the IPT. It can accurately establish the relationship between the plugging states, the spoiler angles, and the flow field parameters.

Adaptive dynamic programming (ADP) is a valuable method in optimal control theory, in which a neural network or other functions approximate the optimal objective function. An offline iteration or online update can define the optimal control strategy, which is widely used in robot control and other fields. Chen used an ADP algorithm to optimize the online parameters of the GH4169 superalloy thermal deformation process [22], obtaining a more uniform and finer microstructure. Huang proposed an ADP controller based on Lyapunov to realize more accurate and faster navigation of unmanned surface vessels [23]. Ruan designed an optimal online-learning fin-stabilizer controller based on ADP [24], which reduced the error of the ship roll model and the influence of external disturbances. Liu used the ADP algorithm to plan the smooth movement of a spacecraft solar panel stretching process [25], to make the solar panel system move smoothly to the end state. It can be seen that ADP can regulate the pneumatic spoiler control system during the plugging process, according to the flow field parameter information extracted by the Bi-LSTM predictor. The pressure gradient was optimized in real time, reducing the impact of the fluid on the IPT.

In our previous research, we found that the spoiler angles

had a significant impact on the flow field around the IPT [26]. In the past, we discretized the entire plugging process, and regulated the spoiler angles in each discrete interval [27] optimizing the flow field parameters. However, the actual plugging process of the IPT is continuous [28], so the spoilers should be constantly regulated. In this paper, we proposed a continuously regulating control system based on a Bi-LSTM neural network and ADP algorithm. First, we carried out a disturbing flow experiment of an IPT for different spoiler angles. A mathematical model was established for the pneumatic spoiler control system, and a linear analysis was carried out. Second, a Bi-LSTM neural network was used to establish the predictor of the control system. A non-linear relationship between the plugging states, spoiler angles, and pressure gradient was established. In addition, a random search algorithm was used to optimize the Bi-LSTM predictor model. The best hyperparameters were selected for the highest prediction accuracy. Third, we designed an ADP controller, which consisted of critic network and action network. The critic network was used to approximate the optimal objective function. The action network obtained the optimal control signals. The pressure gradient between the upstream and downstream of the IPT was reduced by regulating the pneumatic spoiler control system. Finally, we conducted an interactive simulation between Python and Matlab/Simulink to verify the optimization effect of the controller.

2 Model and experiment of the control system of the IPT with spoilers

2.1 Structure model of the IPT with spoilers

The structure of the IPT with spoilers is shown in Fig.1, which mainly includes a pressure head, sealing ring, squeeze bowl, sliders, actuator plate, pushing tube, spoiler device and a pneumatic control system. After receiving the work instruction, the IPT moved to the designated position under the push of the medium in the pipe. The pushing tube and the actuator plate were driven by a pneumatic cylinder [29], moving the sliders upward along the surface of the squeezing bowl. The thread on the surface of the sliders pierced the pipe wall such that the whole mechanism was fixed. The sealing ring was squeezed, causing it to expand radially to seal against any leakage of the pipeline. In this research, a pneumatic plugging method was adopted, which avoids environmental pollution and saves energy. The main function of the pneumatic control system was to provide the power for the plugging process and to regulate the spoiler angles. The chassis moved inwardly to compress the spring

under the push of the pneumatic cylinder. The spoilers were driven to open outwards by the link mechanism in a flipping

motion. The principle of spoiler' motion is shown in Fig.2.

Fig. 1 Mechanical structure of the IPT with spoilers

Fig. 2 Schematic diagram of spoiler' motion

According to the law of cosines, the relationship between the angles and the lengths of the spoilers, the length of Link1, the distance between the spoilers and the center of the chassis, and the displacement of the pneumatic cylinder piston can be obtained, as shown in Eq. (1):

$$\begin{cases} y=0 & \alpha=0 \\ \left(x - \frac{y}{\tan \alpha}\right)^2 + \left(\frac{l}{2} - \frac{y}{\sin \alpha}\right)^2 - l_1^2 = 2\left(x - \frac{y}{\tan \alpha}\right)\left(\frac{l}{2} \cos \alpha - \frac{y}{\tan \alpha}\right) & 0 < \alpha < 180 \\ y=7.5 & \alpha=180 \end{cases} \quad (1)$$

where x is the distance between the bottom of the spoilers and the center of the chassis, mm; y is the displacement of the pneumatic cylinder piston, mm; α is the angle of the spoilers, °; l is the length of the spoilers, mm; l_1 is the length of Link1, mm.

2.2 Model of the pneumatic control system for IPT

The pneumatic control system of the IPT mainly included two pneumatic servo systems, which controlled the plugging process and the flipping motion of the spoilers, as

shown in Fig.3. The system on the left is the pneumatic plugging control system, which used the inner pneumatic cylinder for power. The system determined the plugging process by detecting the pneumatic cylinder piston displacement in real time.

The system on the right in Fig.3 is the pneumatic spoiler control system, which controlled the flipping of the spoilers by the pneumatic cylinder. The spoiler angle was determined by detecting the pneumatic cylinder piston displacement. The system was regulated to reduce the pressure gradient between the upstream and downstream of the IPT, so the research was mainly aimed at the pneumatic spoiler control system.

Fig. 3 Schematic diagram of pneumatic control system

The pneumatic cylinder piston force balance equation is as follows:

$$A_1 P_1 - A_2 P_2 = M \frac{d^2 y}{dt^2} + B \frac{dy}{dt} + F_L + F_f \quad (2)$$

The pressure-flow equation of the proportional valve (taking the air inlet as an example) is:

$$G(s) = \frac{135.9}{s(s^2 + 71.37s + 103.6)} \quad (5)$$

2.3 Disturbing flow experiment of in-pipe intelligent plugging

To study the relationship between the plugging states, the spoiler angles, and the pressure gradient, disturbing flow experiments were carried out for spoiler angles of 0°, 30°, 60°, 90°, 120°, 150°, and 180°. The IPT model with spoilers was simplified. The disturbing flow experiment system of in-pipe intelligent plugging is shown in Fig.6, which mainly included the water circulation system, power transmission, plugging system, and data acquisition system. The water circulation system used a submersible pump to transport

water to the plexiglass pipe and formed a water circulation loop through the water tank. The diameter of the pipe was 50 mm, and the flow in the pipe was controlled with throttle valve. The power transmission system provided power for the plugging process. To facilitate operation, a hybrid stepping motor was used for driving, and a ball screw nut was used for transmission. A stepping motor controller was selected for control. The PXI-6236 data acquisition system mainly collected the current signals transmitted by the pressure transmitters. The three pressure transmitters were installed at the measuring points A, B, and C on the outer wall of the pipeline, 100 mm apart, corresponding to the upstream, midstream and downstream of the IPT, respectively.

Fig. 6 Disturbing flow experiment system of in-pipe intelligent plugging

The pressure around the IPT during the plugging process of the original model (spoiler angles of 0°) is shown in Fig.7. During the plugging process, the pressure difference between the upstream and downstream of the IPT gradually increased, which means that the pressure gradient before and after the IPT also increased. The pressure from upstream to downstream of the flow field around the IPT presented a regional distribution. The pressure gradient between the regions was relatively large, which would disturb the normal

operation of the IPT, causing the fluid in the pipe to be unstable [32]. Additionally, the disturbance could impose vibration on the IPT [33], which is not conducive to the stability of the plugging.

Fig. 7 Pressure curves of the original model

Through numerical simulation and experimental analysis of the flow field around the IPT, the pressure gradient ($\Delta P/\Delta L$) between the upstream and downstream of the IPT had a significant impact on the normal operation of the IPT [34]. Therefore, this gradient could be a measurement standard of the flow field around the IPT and can be calculated from the measured pressure at points A and C, as shown in Eq. (6).

$$\frac{\Delta P}{\Delta L} = \frac{P_A - P_C}{\Delta L} \quad (6)$$

where P_A and P_C are the pressure values measured at points A and C respectively, kPa; ΔL is the distance between points A and C, which is 200 mm.

The peak pressure gradients between points A and C, equivalent to the maximum $\Delta P/\Delta L$, are shown in Fig. 8. The larger the gradient, the more unstable the plugging process. The pressure gradient varies with the spoiler angle, indicating that the pressure gradient can be regulated by the spoiler angle. Therefore, a relationship model between the plugging states, the spoiler angles, and the pressure gradient is required.

Fig. 8 Peak values of pressure gradient for different IPT models

3 Bi-LSTM neural network predictor for pressure gradient

3.1 Structure of Bi-LSTM neural network

To obtain the pressure gradient at different spoiler angles during the plugging process, a predictor based on Bi-LSTM was established. The structure of the LSTM neural network is shown in Fig.9. The activation function in the network is used to implement short-term memory, and the weight updating is used to implement long-term memory. C_{t-1} and C_t represent the previous and updated cell state, respectively. The variables h_{t-1} and h_t represent the previous and current hidden layer output, respectively. In addition, x_t is the input of the current LSTM unit. LSTM mainly includes cell state, forget gate, input gate and output gate [35]. Among them, the cell state is the core of the entire network and transmits relevant information along the time sequence. The other components are used to update the state.

Fig. 9 Structure of LSTM neural network

The forget gate is mainly responsible for deciding what information should be discarded or retained and can be expressed as

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \quad (7)$$

where f_t is passed to the sigmoid function by the previous hidden state information and the current input information; W_f is the weight term; b_f is the bias term.

The input gate is responsible for selecting new information to record in the cell state and can be expressed as

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \quad (8)$$

$$\tilde{C}_t = \tanh(W_c \cdot [h_{t-1}, x_t] + b_c) \quad (9)$$

$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t \quad (10)$$

where h_{t-1} and x_t are passed to the sigmoid and the tanh activation functions simultaneously, the latter of which obtains the cell state candidate value $\tilde{C}_t$. Together, these functions determine the new cell state C_t .

The output gate is responsible for determining the output of the current hidden layer, and similarly to the input gate, can be expressed as

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \quad (11)$$

$$h_t = o_t * \tanh(C_t) \quad (12)$$

For the prediction of time series problems, past and future information can be used to predict the information at the current time. However, the LSTM neural network can only use the information before the current time to predict the current result. A Bi-LSTM neural network predicts the output information based on the entire time series [36]. It divides the hidden layer into two independent parts, forming two independent hidden layers in the positive and negative directions. Then, it feeds forward to the same output layer, containing both past and future information. The structure of a Bi-LSTM is shown in Fig.10. The first-layer LSTM calculates the order information at the current moment. The second-layer LSTM reads the same time sequence in reverse and adds the reverse order information. The hidden-layer output between them is not only passed to adjacent units but also acts on the input of the next LSTM layer.

Fig. 10 Structure of Bi-LSTM neural network

3.2 Model of pressure gradient with Bi-LSTM predictor

To realize the prediction of the flow field parameters around the IPT, a predictor based on Bi-LSTM was constructed. First, the data collected in the experiment was processed. The pressure data of the entire plugging process (0-100%, with 1% intervals) under different spoiler angles were recorded. A total of 707 datasets were obtained. Next, the pressure gradient $\Delta P/\Delta L$ between the upstream and downstream of the IPT was selected to measure the flow field state around the IPT. This value was used as the output information of the predictor. The plugging states and the spoiler angles were input to the predictor. The experimental data were divided into a training set (90%) and test set (10%), and the model of the predictor is shown in Fig.11.

Fig. 11 Model of pressure gradient with Bi-LSTM predictor

To improve the convergence speed and accuracy of the predictor, the dataset needed to be normalized to a range of 0-1. The Bi-LSTM neural network contained a number of freely definable hyperparameters, such as the number of network layers, number of neurons in each layer, activation function, and optimization algorithm, etc. These hyperparameters had a significant impact on the prediction accuracy of the neural network, so the mean square error of the test set was selected as the objective function. A random search algorithm was used to find the optimal combination of hyperparameters, and the train set was trained. To prevent over-fitting, a dropout layer was added between the Bi-LSTM and the fully connected layers. The purpose of the dropout layer was to discard the output value of the hidden layer with a certain probability, ensuring that the neurons will not affect the forward propagation during the training process. The mean square error (MSE) was used as the loss function, and the number of iterations was 800. The prediction result was evaluated through the test set. The hyperparameters obtained after optimization are shown in Table 1.

Table 1 Hyperparameters of the predictor

Hyperparameters of the predictor	Values of hyperparameters
The number of network layers	4
The number of neurons in each layer	50
Activation function	tanh
Optimization algorithm	Adam
Dropout rate	0.1

To verify the advantages of the Bi-LSTM neural network method, other neural network methods were selected for comparison. The hyperparameters of all neural network methods were obtained by a random search algorithm. The results are shown in Table 2. The mean absolute percentage error (MAPE) and root mean squared error (RMSE) were

used as performance indicators, and the formulas are shown in Eq. (13)-(14). Table 2 shows that the prediction performance of Bi-LSTM is significantly better than other models. Compared with the one-way LSTM model, the MAPEs of the train and test sets are close, but the RMSE is reduced by 13.79% and 15.84%, respectively, which indicates an improved prediction accuracy.

$$L_{RMSE} = \sqrt{\frac{\sum_{i=1}^N (\hat{y}_i - y_i)^2}{N}} \quad (13)$$

$$L_{MAE} = \frac{1}{N} \sum_{i=1}^N |\hat{y}_i - y_i| \quad (14)$$

where $\hat{y}_i$ and y_i are the predicted and actual values of the i -th predicted point in the test respectively; N is the number of test samples.

Table 2 Prediction effect of different neural network methods

Neural network methods	MAPE (Train set)	RMSE (Train set)	MAPE (Test set)	RMSE (Test set)
LSTM	13%	0.58	1%	1.01
RNN	19%	5.63	4%	6.00
GRU	16%	0.60	1%	0.95
BP	18%	5.90	6%	7.38
Bi-LSTM	12%	0.50	1%	0.85

3.3 Analysis of prediction results

The training loss of the Bi-LSTM predictor is shown in Fig.12. As the number of iterations increases, the training loss shows a downward trend and eventually stabilizes. Despite certain fluctuations, it is finally controlled at the order of 10⁻⁵.

Fig. 12 Training loss of predictor

After the training was completed, the test set was used to evaluate the predictor. The prediction effects of the train and test sets improved, as shown in Fig.13. The maximum error between the predicted and actual values in the train set was 0.69 kPa/m, accounting for 8.77% of the actual value. The

average error and average error rate were 0.43 kPa/m and 0.94%, respectively. The maximum error between the predicted and actual values in the test set was 1.49 kPa/m, accounting for 0.82% of the actual value. The average error and average error rate were 0.70 kPa/m and 0.61%, respectively. The prediction errors of the train and test sets were controlled within 9%, meeting the requirements of prediction accuracy. Therefore, the established Bi-LSTM predictor could predict the pressure gradient accurately, which is useful for optimizing the flow field parameters and the realization of the vibration reduction of the IPT.

Fig.13 Prediction effect of the train set and the test set

4 Design of ADP controller for IPT

4.1 Structure of ADP controller

The ADP controller was designed to optimize the pressure gradient for the plugging process. Dynamic programming is crucial for optimal control. This method is based on the Bellman optimality principle, which divides the entire process into several intervals. The optimal solution is solved in each interval, such that the global problem is achieved [37]. Assume a discrete-time nonlinear system,

$$x(k+1) = f[x(k), u(k), k] \quad (15)$$

where x is the state variable of the system; u is the control variable; k ($k=0,1,\dots,n-1$) is the stage variable.

The cost-to-go function J of this nonlinear system, shown in Eq. (16), is related to time k and state $x(k)$. The purpose of dynamic programming is to solve the optimal control $u(k)$ and minimize the cost-to-go function J [38].

$$J[x(k), k] = \sum_{i=k}^{\infty} \gamma^{i-k} U[x(i), u(i), i] \quad (16)$$

where γ is the discount factor, and $0 < \gamma \leq 1$; U is the utility function; k is the stage variable.

The traditional dynamic programming method has a large amount of calculation when dealing with problems, making

it prone to "dimensional disaster". It has a significant impact on the data storage and calculation speed, limiting the dynamic programming algorithm. ADP is a non-model strategy optimization method, which integrates dynamic programming, reinforcement learning, and neural network. The neural network approximation method is used for the optimization calculation [39], which can effectively solve the shortcomings of traditional dynamic programming. This approach is suitable for the high-dimensional complicated non-linear system.

In this paper, based on the "Actor-Critic" in reinforcement learning, an ADP controller was designed, composed of a critic network and an action network. The critic network was used to approximate the cost-to-go function J , and the action network was used to approximate the optimal control signal $u(k)$. Its structure is shown in Fig.14. The plugging state

was selected as the input of the action network. The control signal $u(k)$ of the pneumatic spoiler control system was used as the output of the action network. The plugging state was obtained by detecting the displacement of the left pneumatic cylinder in Fig.3. After the control signal $u(k)$ was input to the pneumatic spoiler control system, the spoiler angles changed. According to the established Bi-LSTM neural network predictor, the pressure gradient at this time could be obtained, which corresponded to the plugging state and the control signal $u(k)$. The pressure gradient was used as the cost-to-go function of the controller. The plugging state and the control signal $u(k)$ were combined as the input of the critic network. The pressure gradient was used as the output of the critic network. A multi-layer LSTM neural network was used in both the critic network and action network.

Fig. 14 Structure of ADP controller for IPT

4.2 Critic network

In the ADP controller, the critic network was used to approximate the cost-to-go function J . Here, a three-layer LSTM neural network was designed as the critic network, which contained 100 input-layer neurons, 150 hidden-layer neurons and one output-layer neuron. Its structure is shown in Fig.15. The input signal $x(k)$ was the plugging state, obtained by detecting the displacement of the left pneumatic cylinder of the pneumatic control system. The input control signal $u(k)$ was obtained from the output signal of the action

network. The activation function of the critic network was the tanh function, and the weight updating algorithm was the Adam algorithm.

Fig. 15 Structure of critic network

The cost-to-go function of the critic network's output was the objective function of the ADP controller, as shown in Eq. (17). In the optimization process, according to the control signals input to the pneumatic spoiler control system, the corresponding spoiler angles can be obtained. Then, the pressure gradient could be calculated by the Bi-LSTM neural network predictor, which was used as the objective function. Through continuous training of the critic network, the corresponding objective function values could be output according to the plugging states and the control signals.

$$J[x(k), u(k)] = \frac{\Delta P}{\Delta L}(k) \quad (17)$$

where: $x(k)$ is the plugging state of stage k ; $u(k)$ is the control signal of stage k ; $\frac{\Delta P}{\Delta L}(k)$ is the pressure gradient of stage k ; k is the stage variable.

In the training process, the MSE between the objective function output by the critic network and the pressure gradient output by the Bi-LSTM predictor was used as the loss function of the critic network. It was trained 100 times, and the loss during the training process is shown in Fig.16. The training loss of the critic network dropped to the order of 10^{-4} , and finally stabilized, which indicates that the critic network could achieve the approximation of the objective function.

Fig. 16 Training loss of critic network

4.3 Action network

The action network of the ADP controller was used to obtain the control signal $u(k)$. The action network and critic network had the same structure. The input signal $x(k)$ was the plugging state, the output signal was the control signal $u(k)$ of the pneumatic spoiler control system. The tanh function was used as the activation function, and the Adam algorithm was used as the weight updating algorithm. Its structure is shown in Fig.17.

Fig. 17 Structure of action network

The action network was mainly used to approximate the control signals with the optimal objective function. The optimal control signals were input to the pneumatic spoiler control system according to the current plugging states. For the training process of the action network, the MSE between the output of the action network and the optimal control signal was used as the loss function. The goal was to minimize the pressure gradient. The action network was trained 800 times. The training loss is shown in Fig.18. The training loss of the action network fluctuated greatly, but generally maintained at values of the order of 10^{-3} at the end. The error met the requirements, indicating that the action network could achieve the approximation of optimal control signals.

Fig. 18 The training loss of action network

5 Results and discussion

In this paper, we employed the interface between Python and Matlab/Simulink for simulation. Python was responsible for the realization of the Bi-LSTM neural network predictor and ADP controller. Matlab/Simulink was used to establish the pneumatic spoiler control system. The neural network models were built using Keras software. The experimental data was imported, and the predictor was trained. According to the plugging states and the pressure gradient output by the predictor, the control signals were optimized. In Matlab/Simulink, we first performed the modeling and linearization of the pneumatic spoiler control system. Then, the optimal control signals obtained in Python were input to optimize the spoiler angles. Finally, the optimal angle sequence and random angle sequence obtained from the entire control process were substituted into the predictor. The pressure gradient was compared to verify the optimization effect of the ADP controller.

The optimal angle sequence and random angle sequence obtained by the ADP controller are shown in Fig. 19. The optimal angle did not change significantly within the range of 0-90% of the plugging states, maintaining stability. The optimal angle suddenly increases when the plugging was 90% complete, and then tended to restabilize. If the pressure gradient was optimized through the random angle sequence, it would cause the pneumatic system to be regulated too frequently, causing oscillation. Therefore, the optimal angle sequence could help achieve adequate control performance of the pneumatic spoiler control system.

Fig. 19 Optimal angle sequence and random angle sequence

The pressure gradient under the optimal angle sequence and the random angle sequence is shown in Fig.20. Compared to the random sequence, the pressure gradient of the optimal control sequence remained lower. The pressure gradient of the entire plugging process dropped by 29.72 kPa/m at the maximum and 11.77 kPa/m on average. The pressure gradient of the optimal control sequence decreased by an average of 25.94% compared with the random sequence. Therefore, the designed ADP controller could effectively optimize the pressure gradient during the plugging process, diminishing the pressure pulsation in the plugging process [40]. The vortex and backflow phenomenon in the flow field could be reduced [41]. Additionally, the impact of the fluid on the IPT could be alleviated, which allows for a safe and stable plugging operation since the vibration of the IPT would lessen.

Fig. 20 Optimization effect of pressure gradient

6 Conclusions

- (1) The flow field around the IPT was optimized by regulating the spoiler angles at the tail of the IPT. A mathematical model of the pneumatic spoiler control system was established and linearized. We designed and conducted a disturbing flow experiment of in-pipe intelligent plugging, recording the experimental data of

different spoiler models in the plugging process. The pressure gradient between the upstream and downstream of the IPT was selected as the standard to measure the flow field.

- (2) According to the experimental data, we established the predictor based on Bi-LSTM. The spoiler angles and plugging states were the input of the predictor, and the pressure gradient was the output of the predictor. A random search algorithm optimized the hyperparameters of the neural network for maximum predictor performance. The errors of the train and test sets were controlled within 9%, indicating that the established predictor could accurately estimate the pressure gradient.
- (3) We designed an ADP controller composed of the critic network and action network. The critic network approximated the objective function to be optimized. The action network obtained the optimal control signals of the pneumatic spoiler control system. During the plugging process, the pressure gradient was adjusted to obtain the optimal control sequence. Through interactive simulation between Python and Matlab/Simulink, the results demonstrated that the ADP controller could reduce the pressure gradient during the entire plugging process by an average of 25.94%. This approach lessened the impact of fluid on the IPT, reducing its vibration, and achieving stability of plugging operation.

7 Declaration

Acknowledgements

The authors sincerely thanks to Professor Hong Zhao of China University of Petroleum (Beijing) for her critical discussion and reading during manuscript preparation.

Funding

Supported by National Natural Science Foundation of China (No. 51575528), Science Foundation of China University of Petroleum, Beijing (No.2462020XKJS01).

Availability of data and materials

The datasets supporting the conclusions of this article are included within the article.

Authors' contributions

The author' contributions are as follows: Xing-yuan Miao wrote the manuscript and conducted the laboratory analysis; Hong Zhao was in charge of the whole trial and proposed the method.

Competing interests

The authors declare no competing financial interests.

Consent for publication

Not applicable

Ethics approval and consent to participate

Not applicable

References

- [1] Mohammed S.E, Ahmed S, Tarek Z, et al. Artificial neural network models for predicting condition of offshore oil and gas pipelines[J]. *Automation in Construction*, 2014, 45:50-65.
- [2] Malihe M, Mansour R. Dynamic analysis and simulation of long pig in gas pipeline [J]. *J Nat Gas Sci Eng*, 2015, 23:294-303.
- [3] Sun H F. Technology of Non-stop Transport with Pressure Hole Sealing Applied on Gas Pipeline [J]. *Pipeline Technique and Equipment*, 2016, 4(05):30-33.
- [4] Zhang K, Zhu X X, Zhang S M, et al. Experimental research on frictional resistance of smart isolation tool during maintenance operation in offshore pipeline [J]. *Ind Lubr Tribol*, 2018, 70(07): 1145-1150.
- [5] Bowie. Low Pressure Hot Tap Pipeline Isolation [P]. Google Patents.2015, USA.
- [6] Li Z B, You J, Cheng H H, et al. Structural design and plugging performance of 508 mm single-section double-seal internal plugging device[J]. *Oil & Gas Storage And Transportation*, 2019, 38(09): 1065-1071.
- [7] Li Z B, Hu T H, Fu D R, et al. Influence of the profile angle of slip in pipeline plugging device on the anchoring performance [J]. *Oil & Gas Storage And Transportation*, 2018, 37(11):1308-1314.
- [8] Li Z B, Zhang J, Jing R, et al. The effect of the rubber tube length of in-pipe plug on the plugging performance [J]. *Oil & Gas Storage And Transportation*, 2019, 38(07): 810-815.
- [9] Quan Y. Design of $\Phi 508$ Spherical Double Plugging for Oil and Gas Pipeline [D]. China University of Petroleum, Beijing, 2017.
- [10] Shen K. The development of feeding and connecting mechanism of spherical double-plugging plug in $\Phi 711$ mm pipeline [D]. China University of Petroleum, Beijing, 2016.
- [11] Zhu X X, Wang H, Lin F Q, et al. Research on the automatically trajectory control of a spherical isolation plug in subsea pipeline [J]. *Ocean Eng*, 2017, 141:101-107.
- [12] Wang W M, Guo J Q, Fan J C, et al. Research on the proportional-integral-derivative synchronous control method of the marine spherical isolation plug in the rotation process [J]. *Proceedings of the Institution of Mechanical Engineers, Part M: Journal of Engineering for the Maritime Environment*, 2020, 234(04):1-10.
- [13] Zhao H, Hu H R. Flow field analysis on internal plug at different plugging states [J]. *Oil & Gas Storage And Transportation*, 2016, 35(10): 1060-1065.
- [14] Zhao H, Hu H R. Optimal design of a pipe isolation plugging tool using a computational fluid dynamics simulation with response surface methodology and a modified genetic algorithm [J]. *Adv Mech Eng*, 2017, 9(10):1-12.
- [15] Zhao H, Hu H R, He T. Simulation on the distribution laws of flow

- field around the in-pipe intelligent plug of different end faces [J]. Oil & Gas Storage And Transportation, 2018, 37(04): 413-420.
- [16] Ma M, Zhao H, He T. DAMS-based Kinetic Research on the Locking Mechanism of Pipeline Plugging Device [J]. China Petroleum Machinery, 2014, 42(04): 101-105.
- [17] Wu T T, Zhao H. An energy-saving and velocity-tracking control design for the pipe isolation tool [J]. Adv Mech Eng, 2019, 11(04):1-13.
- [18] Wu T T, Zhao H, Gao B X, et al. Energy-Saving for a Velocity Control System of a Pipe Isolation Tool Based on a Reinforcement Learning Method [J]. Int J Pr Eng Man-Gt, 2021.
- [19] Zeng H J, Guo J S. Fault Prognostic of Aeroengine Using Bidirectional LSTM Neural Network [J]. JOURNAL OF AIR FORCE ENGINEERING UNIVERSITY (NATURAL SCIENCE EDITION), 2019, 20(04): 26-32.
- [20] Kang S Q, Zhou Y, Wang Y J, et al. RUL Prediction Method of a Rolling Bearing based on Improved SAE and Bi-LSTM [J]. ACTA AUTOMATICA SINICA. 2020, 1-11.
- [21] Zhou Y Q, Xu Y L. Real-time human pose recognition in complex environment based on the bidirectional LSTM [J]. Chinese Journal of Scientific Instrument, 2020, 41(03): 192-201.
- [22] Chen D D, Lin Y C. Online optimizing hot forming parameters for alloy parts based on action-dependent heuristic dynamic programming [J]. The International Journal of Advanced Manufacturing Technology, 2019, 104.
- [23] Huang Z J, Liu X Z, Wen J Y, et al. Adaptive Navigating Control Based on the Parallel Action-Network ADHDP Method for Unmanned Surface Vessel [J]. Adv Mater Sci Eng, 2019, 1-8.
- [24] Ruan G W, Li T S, Yu R H, et al. Online optimal control for ship fin stabilizer system based on action dependent heuristic dynamic programming [J]. Journal of Nanjing University of Information Science & Technology (Natural Science Edition), 2021, 13(01):10-16.
- [25] Liu B, Ge X S, Zhao J G. Adaptive dynamic programming method for spacecraft solar panel stretching process [J]. Journal of Beijing Information Science & Technology University, 2020, 35(03):36-40.
- [26] Miao X Y, Zhao H, Wang C J. Design and Optimization of Vibration Reduction Structure of Intelligent Pluggers inside Pipe [J]. China Petroleum Machinery, 2020, 48(05): 99-106.
- [27] Miao X Y, Zhao H. Active Control of Pneumatic Damping System for Intelligent Pipeline Plug [J]. China Petroleum Machinery, 2021, 49(06): 51-60.
- [28] Xiang Z Y, Zhao H. Research on Neural Network PID Control Method on Plugging Speed of Intelligent Plugging Isolation Tool [J]. Oil Field Equipment, 2020, 49(06): 20-26.
- [29] Miao X Y, Zhao H. Design of Pneumatic Vibration Reduction Control System for In-pipe Intelligent Plugging Device [J]. MACHINE TOOL & HYDRAULICS, 2021, 49(11): 75-81.
- [30] Zhao H, Pinhas, B. Synchronous position control strategy for bi-cylinder electro-pneumatic systems [J]. International Journal of Control, Automation and Systems, 2016, 14(06):1501-1510.
- [31] Adam S, Tamas B, Peter G. Control Design and Validation for Floating Piston Electro-Pneumatic Gearbox Actuator [J]. Applied Sciences, 2020, 10.
- [32] Hu H R. Study on Simulation and Control Optimization in the Plugging Process of SmartPlug [D]. China University of Petroleum, Beijing, 2017
- [33] Miao X Y, Zhao H. Optimal Design of Vibration Structure of Intelligent Plugging Isolation Tool Based on Monte Carlo Algorithm [J]. Oil Field Equipment, 2020,49(6): 46-54.
- [34] Hu H R, Zhao H. Optimization of end face structure of smart plug [J]. China Petroleum Machinery, 2016, 44(05): 71-76.
- [35] Cheng Y, Haung Y X, Pang B, et al. ThermalNet: A deep reinforcement learning-based combustion optimization system for coal-fired boiler [J]. Eng Appl Artif Intel, 2018, 74:303-311.
- [36] Yu Y, Yao H P, Liu Y M. Structural dynamics simulation using a novel physics-guided machine learning method [J]. Eng Appl Artif Intel, 2020, 96:1-14.
- [37] Liu X Y, Sheng X, Huang Y Y. Optimal control for earth pressure balance of shield machine based on action-dependent heuristic dynamic programming [J]. Isa T, 2019, 94:28-35.
- [38] José, Patrícia, João. Online discrete-time LQR controller design with integral action for bulk Bucket Wheel Reclaimer operational processes via Action-Dependent Heuristic Dynamic Programming [J]. Isa T, 2019.
- [39] You Y, Sun D Y, Liu J L, et al. Gearshift Strategy for HMPRT Base on Dynamic Programming [J]. JOURNAL OF MECHANICAL ENGINEERING, 2019, 55(08):106-117.
- [40] Zhao H, Zhao Y X, Ye Z H. Interactions between the fluid and an isolation tool in a pipe: laboratory experiments and numerical simulation [J]. Petrol Sci, 2016, 13: 746-759.
- [41] Zhang H, Cui C, Liao N S, et al. Dynamic Modeling and Analysis of Pipeline Inspection Gauge Passing Girth Welds Considering Fluid-Solid Coupling [J]. JOURNAL OF MECHANICAL ENGINEERING, 2020, 56(23): 129-140.

Biographical notes

Xing-Yuan Miao, born in 1996, is currently a PhD student at *College of Mechanical and Transportation Engineering, China University of Petroleum, Beijing, China*. He received his bachelor degree from *China University of Petroleum, Beijing, China*, in 2019. His research interests include in-pipe robotics, deep reinforcement learning, electromechanical system design and automation. Tel: +86-15027650591; E-mail: miaoxingyuan96@126.com

Hong Zhao, born in 1974, is currently a Professor in *China University of Petroleum, Beijing, China*. She received her Ph.D. degree from *Xi'an Jiaotong University, Xi'an, Shanxi Province, China*, in 2003. Her research interests include mechanical and electrical transmission control, virtual instrument measurement and control technology.

E-mail: hzhao_cn@163.com

Appendix

Parameter	Value	Unit	Meaning
x	6	mm	The distance between the bottom of the spoilers and the center of the chassis
l	12	mm	The length of spoilers
l_1	14	mm	The length of link 1
P_1	—	Pa	Cylinder intake chamber pressure
A_1	1.96×10^{-2}	m^2	Cylinder intake chamber area
P_2	—	Pa	Cylinder exhaust chamber pressure
A_2	1.35×10^{-2}	m^2	Cylinder exhaust chamber area
M	2	kg	Load quality
y	—	mm	Cylinder piston displacement
t	—	s	Time
B	50	$\text{N} \cdot \text{s/m}$	Viscous reduction coefficient
F_L	3925	N	External load
F_f	—	N	Cylinder friction
Q_{m1}	—	m^3/s	Cylinder intake chamber flow
Q_{m2}	—	m^3/s	Cylinder exhaust chamber flow
C_q	0.628	—	Flow Coefficient
x_1	—	mm	Proportional valve core displacement
d_1	6	mm	Proportional valve core diameter
$A(x)$	—	m^2	Proportional valve opening area, $A(x) = \pi d_1 x_1$
T	293	K	Temperature
P_s	0.8	bar	Air pressure
k	1.4	—	Adiabatic coefficient of air
R	287.1	$\text{J}/(\text{kg} \cdot \text{K})$	Gas constant
C_t	0.528	—	Critical pressure ratio
V_1	—	m^3	Cylinder intake chamber volume
V_2	—	m^3	Cylinder exhaust chamber volume
F_m	—	N	Cylinder viscous reduction force