Engineering Applications of Artificial Intelligence 74 (2018) 303–311

Contents lists available at ScienceDirect

Engineering Applications of Artificial Intelligence

journal homepage: www.elsevier.com/locate/engappai

ThermalNet: A deep reinforcement learning-based combustion optimization system for coal-fired boiler

Yin Cheng [a], Yuexin Huang [a], Bo Pang [a], Weidong Zhang [a,b,*]

[a] Department of Automation, Shanghai Jiaotong University, Shanghai 200240, PR China
[b] School of Computer Engineering and Science, Shanghai University, Shanghai 200444, PR China

ARTICLE INFO

Keywords:
Combustion optimization
DQN
LSTM
Reinforcement learning
DCS

ABSTRACT

This paper presents a combustion optimization system for coal-fired boilers that includes a trade-off between emissions control and boiler efficiency. Designing an optimizer for this nonlinear, multiple-input multiple-output problem is challenging. This paper describes the development of an integrated combustion optimization system called ThermalNet, which is based on a deep Q-network (DQN) and a long short-term memory (LSTM) module. ThermalNet is a highly automated system consisting of an LSTM–ConvNet predictor and a DQN optimizer. The LSTM–ConvNet extracts the features of boiler behavior from the distributed control system (DCS) operational data of a supercritical thermal plant. The DQN reinforcement learning optimizer contributes to the online development of policies based on static and dynamic states. ThermalNet establishes a sequence of control actions that both reduce emissions and simultaneously enhance fuel utilization. The internal structure of the DQN optimizer demonstrates a greater representation capacity than does the shallow multilayer optimizer. The presented experiments indicate the effectiveness of the proposed optimization system.

1. Introduction

Large coal-fired power plants are major contributors to total pollutant emissions; consequently, they offer the possibility of reducing emissions through increased thermal efficiency. It is difficult to understand a complex mechanism such as NOx combustion and emissions with only a limited knowledge of combustion theory and chemical kinetics; however, physical experimentation may be expensive (Janakiraman et al., 2016). Furthermore, to control the pollutant discharge without impairing power generation efficiency, the power system needs to dynamically regulate numerous control variables. In this paper, a data-based deep neural network model is established to overcome these challenges. Time-varying relationships and combustion process mechanisms can be obtained using the proposed model without requiring expert knowledge. The generality and extensibility of deep neural networks could further enable a wide range of data processing applications that execute faster and with lower costs.

Zheng combined a support vector regression (SVR) model with ant colony optimization (ACO) to reduce NOx emissions (Zheng et al., 2008). Using this model requires conducting prior parametric field experiments to determine the relationship between the operating parameters and the combustion characteristics. Zebian simultaneous multivariable gradient-based optimization was performed on a pressurized oxy-coal combustion process (Zebian et al., 2012). This study also required prior knowledge of thermal dynamics and fluid mechanics to conduct simulation experiments. Lv proposed the least squares support vector machine (LSSVM)-based ensemble learning paradigm to predict the NOx emissions of coal-fired boilers (Lv et al., 2013). Based on the simulated annealing genetic algorithm (SAGA), a support vector regression (SVR) model was presented to predict the NOx emission concentration (Wei et al., 2013). The SVR model was also used to further optimize the operating parameters to achieve low NOx emissions for coal-fired boilers. However, the enormous space complexity required by SVR models makes them difficult to scale. An enhanced general regression neural network (Enhanced-GRNN) was designed for on-line applications (Song et al., 2016), but the accuracy of the model suffered from changing constraints during operation.

A distributed control system (DCS) plays a crucial role by monitoring coal-fired boilers in real time through big data theory and electronic technology. It also enables us to find the trends of various parameters, including oxygen content, coal consumption, boiler inlet temperature, and boiler efficiency. This paper describes how to adopt a deep neural network to solve the complex time-effective modeling problem of the

* Corresponding author at: Department of Automation, Shanghai Jiaotong University, Shanghai 200240, PR China.
E-mail address: wdzhang@sjtu.edu.cn (W. Zhang).

https://doi.org/10.1016/j.engappai.2018.07.003
Received 31 December 2016; Received in revised form 23 May 2018; Accepted 5 July 2018

Table 1
Structure of ThermalNet: LSTM-3 denotes an LSTM layer with a 3-dimensional output, and FC-3 denotes a fully connected layer with a 3-dimensional output. The output is a 43-dimensional block composed of 4 consecutive time steps. The details can be found in Section 2.

Predictor		Optimizer
Input		
4 × 40	4 × 3	4 × 43
Conv1D 12 × 2	LSTM-3	Conv1D 16 × 2 Conv1D 16 × 2
Conv1D 12 × 2	LSTM-3	FC-50 FC-20
Merge for FC-3		FC-16
Reward function		Q values for 16 actions
Take actions on control variables		

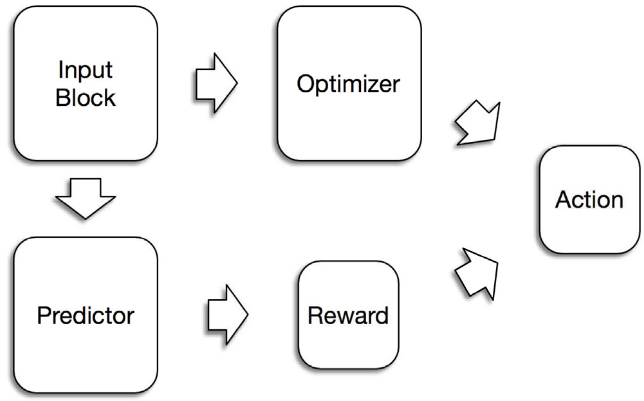

Fig. 1. The architecture of ThermalNet.

coal-fired boiler. Moreover, we apply deep reinforcement learning to make an effective trade-off between emissions reduction and efficiency.

A deep neural network (DNN) is essentially data-driven. Due to the nature of inherent nonlinearity and universal functional approximation (Adhikari and Agrawal, 2013), DNNs offer an ideal approach for modeling data-intensive applications in practical situations. DNNs have been successfully applied to signal feature extraction (Xiong et al., 2016; Qiu et al., 2014) and complex systems analysis (Jiang et al., 2016). The proposed LSTM–ConvNet predictor combines a long short-term memory (LSTM) module with a convolutional neural network (ConvNet). The LSTM has the advantage of a better memory structure (an architecture with three types of gates), which enables processing historical information and, thus, dynamically extracting the inner relation with the time sequence (Chen et al., 2015). ConvNet is highly effective at extracting features from original DCS signals. Features processed by ConvNet are transmitted to the LSTM module to establish the dynamic model for coal-fired boilers. As the monitoring data increases, the system needs to adapt dynamically to address the ever-changing operating conditions. The modeling accuracy depends on factors such as equipment age, measurement noise, and signal fidelity. These features are related to the time sequence of the entire process. Therefore, for coal-fired boiler modeling, the LSTM–ConvNet predictor is advantageous because it can easily handle the relation with the time sequence and – simultaneously – avoid a lack of fidelity, making the LSTM–ConvNet predictor a desirable choice for coal-fired boiler modeling.

Based on the previous LSTM–ConvNet predictor, a Deep Q-Networks (DQN) optimizer is adopted to control the coal-fired boiler. A DQN is a general-purpose reinforcement learning framework for decision-making that has been successfully applied to many challenging problems in stochastic and deterministic situations (Gu et al., 2016). Recently, it has proven possible to obtain a stable and scalable general reinforcement learning system by using deep networks to represent value function and policy (Gu et al., 2017). A DQN can achieve satisfactory control performance even with high-dimensional states. DQN networks, which merely require pixels and game scores, can learn practicable policies that outperform other linear algorithms (Mnih et al., 2015). Moreover, they are also applicable to situations with enormous search spaces and considerable complexity. AlphaGo is a perfect example: this DQN managed to beat Lee Sedol, one of the top Go players in the world (Silver et al., 2016). Applications of reinforcement learning in industrial control are still in their early stages (Lewis and Vrabie, 2009) due to the difficulties in finding good approximation of value functions for control objects. The architecture of the deep Q-network enables researchers to evaluate actions based on abstract features extracted by ConvNet. The DQN is introduced into the control of coal-fired boilers by using the real data recorded by the DCS.

In this paper, we present ThermalNet (see Fig. 1, Table 1), a framework to integrate a combustion optimization architecture with deep reinforcement learning. First, we combine ConvNet and LSTM to develop reliable dynamic characteristics for coal-fired boilers. Second, a DQN optimizer is used to acquire control actions (e.g., to reduce emissions and improve boiler power efficiency). The overall design of ThermalNet involves an LSTM–ConvNet predictor and a DQN optimizer in the training phase. The predictor consists of the convolutional feature extractor and the LSTM time-step connector. In the training phase, a DCS monitoring dataset from a real thermal plant is utilized in both the optimizer and the predictor. The training process starts with LSTM–ConvNet, where the 4-time step input state is split into an observation part and a performance part, which are respectively handled by the convolutional layers and LSTM modules. After training the predictor, the DQN optimizer is trained. Specifically, we produce a reward function to quantify how our optimizer interacts with an environment; the optimizer is trained to take the action that maximizes its reward. The DQN takes the same input states, assigning the resulting Q-values to the different actions. Furthermore, epsilon-greedy policies are used to enable the optimizer to choose more effective actions. Under the supervision of the predictor, the DQN regulates its inherent parameters to obtain the ability to choose the most valuable action (i.e., the action that results in the highest reward value).

2. Distributed control system and data set

Boiler performance directly determines the overall behavior of the thermal plant. To achieve high boiler efficiency, regulating the fuel and air appropriately is crucial. In a practical combustion process, it is difficult to avoid incomplete combustion; therefore, reducing combustion emissions while simultaneously satisfying energy demand are the top priorities for a thermal plant.

2.1. Coal-fired boiler control variables and performance variables

Based on operational experience, the control variables are air volume A_v, fuel content F_c, oxygen content O_c, and feedwater flow F_w. These independent variables are directly related to the combustion process; it is convenient to regulate them in practice. Other input variables, such as the boiler inlet air temperature T_b and the fresh-steam turbine pressure $P_{fs.t}$, are used as reference variables Ref_t in the model (Fig. 2). Furthermore, a series of preprocessing steps, including filtering, scaling, and normalization, are performed in the training phase. Together, the control variables and observational variables are

$$\begin{aligned} \mathrm{Con}_t &= \left[A_v^t, F_c^t, O_c^t, F_w^t\right]^T \in \mathbb{R}^4 \\ \mathrm{Obs}_t &= \left[\mathrm{Ref}_t \quad \mathrm{Con}_t\right]^T \in \mathbb{R}^{40} \end{aligned} \tag{1}$$

In this paper, all the variables (see Table 7) are assumed to be continuous. In the training process, the ranges of all variables are normalized