DQN-Based Chiller Energy Consumption Optimization in IoT-Enabled Data Center

Yanyi Wang*, Yu Sun*, Bo Cheng*, Gaoxiang Jiang*, and Haibo Zhou*
*School of Electronic Science and Engineering, Nanjing University, Nanjing, China, 210023.
Emails:wyynju@foxmail.com, {yusun, bocheng}@smail.nju.edu.cn, jgxwww@sina.cn, haibozhou@nju.edu.cn.

2023 IEEE 23rd International Conference on Communication Technology (ICCT) | 979-8-3503-2595-9/23/$31.00 ©2023 IEEE | DOI: 10.1109/ICCT59356.2023.10419683

***Abstract*—The swift growth of data center has given rise to substantial energy usage and raised notable environmental apprehensions. Reducing cooling energy consumption to decrease power usage effectiveness (PUE) of data center represents a key focal point in optimizing data center energy consumption. Existing work typically involves energy consumption modeling based on physical models and employs mathematical methods for optimization. However, these methods suffer from limited generalization capability and insufficient stability. To this end, this paper focuses on optimizing the energy consumption of chillers in Internet of Things (IoT)-enabled data center based on Deep Q-Network (DQN). We design a PUE optimization framework for IoT-enabled data center and model the chiller energy consumption leveraging machine learning. To achieve energy saving while ensuring the safe operation of IoT-enabled data center, we transform the process of chiller energy consumption optimization into a Markov Decision Process and propose a DQN-based algorithm for the problem. Through simulations, we demonstrate that the proposed algorithm can significantly reduce chiller energy consumption by more than 50%, surpassing the performance of random action algorithm. Moreover, in comparison to classical optimization algorithms, our algorithm shows promising performance and exhibits potential for future applications in more complex optimization scenarios.**

***Index Terms*—Internet of Things, data center, machine learning, chiller energy consumption optimization, Deep Q-Network.**

I. Introduction

Data center is undergoing rapid development as a pivotal element of the Fourth Industrial Revolution. Simultaneously, the significant energy consumption of data center has gained momentum due to economic, social, and environmental factors. Projections indicate that by 2030, the data center industry is anticipated to be responsible for 8% of the total global carbon emissions [1]. Therefore, the optimization of data center energy consumption is of utmost urgency. The Power Usage Effectiveness (PUE) is a measure that expresses the relationship between the overall energy usage of data center and the energy utilized by its IT equipment. It is noteworthy that nearly half of the energy costs in traditional data center is associated with cooling [2]. Consequently, optimizing cooling energy consumption becomes a critical focal point for reducing PUE in data center [3].

To optimize energy consumption in cooling systems, existing research has devised optimization strategies utilizing energy consumption models. These models are typically constructed based on thermodynamic principles and empirical data, with mathematical techniques employed to adjust inputs for energy optimization. For example, Yu et al. [4] created models for thermodynamic analysis of chillers and cooling towers. These models were utilized to evaluate the impact of various control strategies for cooling towers and condenser water pumps on the performance of cooling systems. Additionally, deriving simplified models from fundamental mass and energy balance principles, Wang et al. [5] characterized intricate mass and energy flows in data center employing raised floors for air circulation and cooling and introduced a model predictive controller for energy consumption optimization. However, these methods exhibit limited generalization capability and face significant challenges in terms of stability and timeliness. The recent development of the Internet of Things (IoT), particularly in the realm of wireless sensor networks, has revolutionized the way we collect data from genuine data center. This cutting-edge technology empowers us to swiftly amass vast amounts of data, creating an unprecedented opportunity for data-driven modeling and optimization strategies tailored specifically to the unique needs and challenges of these facilities. This enhanced capability not only enables us to gain deeper insights into the intricate operations of data center but also paves the way for the development of more efficient, adaptive, and finely tuned solutions that can significantly enhance its overall performance and resource utilization [6]. Moreover, the rapid advancements in Machine Learning (ML) and Reinforcement Learning (RL) have brought forth significant progress, particularly in the form of Deep Reinforcement Learning (DRL), across various domains [7]. Deep Q-Network (DQN), as a DRL algorithm, is anticipated to address the challenges posed by traditional optimization algorithms by combining DL with the Q-learning technique [8].

This paper primarily focuses on the optimization of PUE while ensuring safety of real-world data center operations. The cooling system plays a crucial role in determining the PUE of IoT-enabled data center, with chillers being the primary energy-consuming equipment. As a result, this paper specifically targets chillers and employs real data to predict and optimize their energy consumption incorporating the advantages of IoT technology. The main contributions of this research can be summarized as follows:

- We provide an overall description of PUE optimization framework in IoT-enabled data center, encompassing the complete process from data collection to the implementation of optimization algorithms in practical systems.

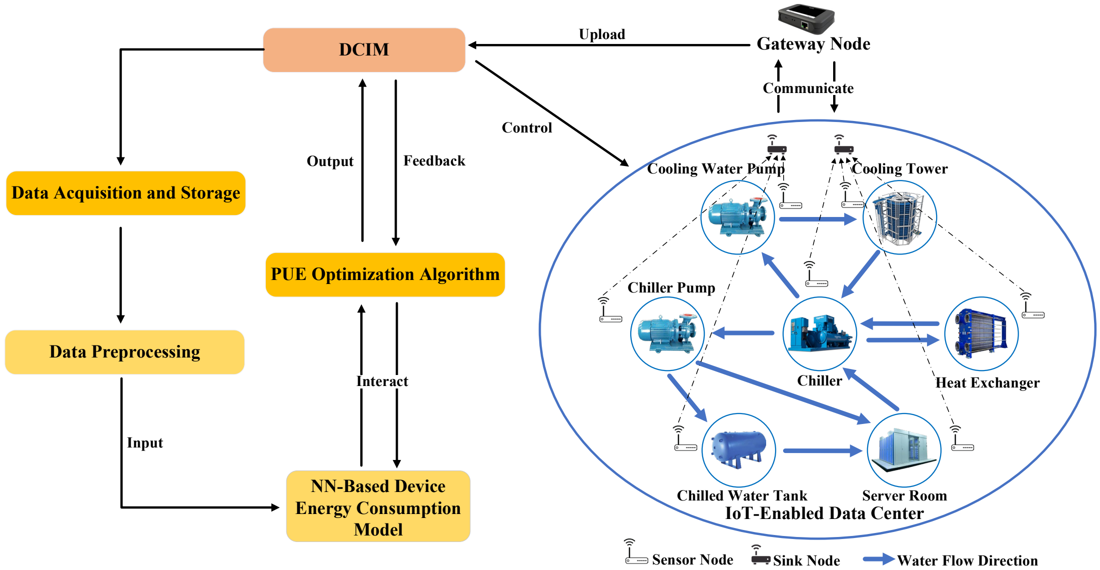

Fig. 1. The PUE optimization framework for IoT-enabled data center.

- We carefully select 16 parameters closely linked to chiller energy consumption and construct a neural network (NN) model that achieves satisfactory performance.
- We propose a DQN-based energy consumption optimization algorithm for the chiller. The algorithm's feasibility and superiority are demonstrated through performance comparison with classical optimization algorithms.

The rest of this paper is organized as follows. Section II introduces the PUE optimization framework and the chiller energy consumption model. Section III presents a DQN-based chiller energy consumption optimization algorithm. The simulation results are shown in Section IV, and Section V is the conclusion of this paper.

II. System Model

In this section, we begin by designing a system framework for PUE optimization. Subsequently, we proceed to construct an NN-based model to predict the energy consumption of the chiller.

A. PUE Optimization Framework

As depicted in Fig. 1, in IoT-enabled data center, a variety of sensors are deployed on devices, allowing them to connect to a sensor network. For instance, the chillers are equipped with temperature, humidity, pressure sensors, and more. Sensor nodes have the responsibility of perceiving and collecting various parameters from the physical environment. Subsequently, the collected data is aggregated and processed at the sink nodes before transmission to the gateway node [9], [10]. Afterwards, the data undergoes additional processing, gets uploaded to the Data Center Infrastructure Management (DCIM) system, and is stored in the database following predefined rules. The collected data undergoes preprocessing, including data cleaning, augmentation, feature extraction, and normalization. Next, the device energy consumption model is trained using the processed dataset and deployed in the PUE prediction algorithm. Based on the feedback from the device energy consumption model, the device energy consumption optimization algorithm is designed, taking into consideration the operational safety boundaries of the devices. The PUE optimization algorithm generates control instructions that are dispatched to the DCIM system. The DCIM system adjusts the tunable parameters of the devices in the real system accordingly. Simultaneously, the real system continuously transmits data through the device sensors.

Our primary focus is on the construction of an NN-based chiller energy consumption model and the optimization of chiller energy consumption. We firstly construct the chiller energy consumption model.

B. Chiller Energy Consumption Model

Chillers play a central role in the water cooling system of our designated configuration [11]. Notably, all chillers integrated into our real-world system conform to identical specifications, facilitating the selection of a single chiller for our modeling efforts. Nevertheless, developing an energy consumption model based solely on a physical framework presents significant challenges, primarily due to the lack of universally accepted equations tailored to this specific chiller model. In light of this, we pursue an alternative approach by harnessing the potential of IoT technology and utilizing the abundant data meticulously collected from real-world chillers. By adopting a data-driven approach [12], we overcome the constraints imposed by the scarcity of customized formulas. This strategy, founded on the utilization of abundant real-world data, emerges as a viable solution. To construct a highly accurate energy consumption model for chillers, we seamlessly integrate this data-driven paradigm with the capabilities of a Multilayer Perceptron (MLP). The inherent capacity of the MLP architecture to identify intricate patterns within complex datasets complements our pursuit, thereby enhancing the precision of our predictive model [13].

For the energy consumption model of the chiller, we carefully chose 16 device measurement points as input parameters, including cooling water outlet temperature (CWOT), evaporator outlet water temperature(EOWT), chilled water supply pressure (CWSP), cooling water outlet pressure (CWOP) and so on. The energy consumption of the chiller serves as the output parameter for the MLP. The architecture of the neural network used in this model is illustrated in Fig. 2.

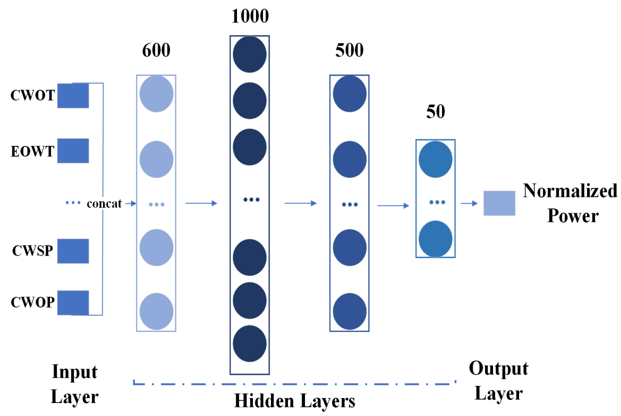

Fig. 2. The MLP in the chiller energy consumption model.

III. Chiller Energy Consumption Optimization Algorithm

In this section, we firstly convert the process of optimizing chiller energy consumption into a Markov decision process (MDP). Subsequently, we design an algorithm based on DQN to achieve energy consumption optimization.

A. Chiller Energy Consumption Optimization Problem

We transform the original stochastic optimization problem into an MDP problem. Given the observed state s_t, the agent conducts operation a_t. Then, the environment feeds back the corresponding reward r_t, and the state is transferred into s_{t+1}. The state, action, state transition, state termination and reward are described as follows:

State: We construct a $16{*}1$ dimensional array containing the values of 16 input parameters of the chiller energy consumption model at a specific time point t_0. This array represents the state of the chiller at time t_0.

Action: Considering the actual conditions of chillers, we choose CWOT as a tunable parameter. We assume the state space of CWOT to be $[T_1^l, T_1^u]$. The action space is assumed to be $[A_1^l, A_1^u]$ with a minimum step size of A_1^m. If, based on the policy, the chosen action in the current state would cause CWOT to exceed its state space, we set the next state of CWOT to be the CWOT critical value closest to the exceeding value (either T_1^l or T_1^u).

State Transition and Termination: Among the various parameters in the chiller energy consumption model, EOWT is significantly influenced by CWOT and serves as an important reference parameter for safe chiller operation. We construct an EOWT prediction model using an MLP, with the labels being the input parameters of the chiller energy consumption model excluding EOWT and the features being the normalized values of EOWT. After each adjustment of CWOT based on the policy, the value of EOWT is updated using the EOWT prediction model. As for the termination condition, we limit the maximum exploration attempts of the agent to 20 per round. Through algorithm training, it is typically feasible to find optimization solutions that meet our expectations within 20 exploration attempts.

Reward: We define the reward r_t as a combination of two components:

$$r_t = r_{dec,t} + r_{penal,t}, \tag{1}$$

where $r_{dec,t}$ is the reward for the reduction in energy consumption achieved by transitioning from the current state s_t to the next state s_{t+1} after taking action a_t, and $r_{penal,t}$ represents a penalty for EOWT exceeding the safety margin in the next state s_{t+1}.

We denote the normalized energy consumption value as P which can be obtained by invoking the chiller energy consumption model N_1:

$$P = N_1(s). \tag{2}$$

$r_{dec,t}$ can be further expressed as:

$$r_{dec,t} = \begin{cases} P_t - P_{t+1}, & \text{if } P_t - P_{t+1} > 0; \\ \lambda_1(P_t - P_{t+1}), & \text{if } P_t - P_{t+1} \leq 0, \end{cases} \tag{3}$$

where λ_1 is a cost pricing factor greater than 1, thus the costs of ineffective actions significantly outweigh the benefits of effective actions. This discrepancy ensures that the agent's exploration is more focused and directed.

Considering the critical requirements of practical chillers, we denote the lower and upper bounds of EOWT as T_2^l and T_2^u. We create a $15 * 1$ dimensional array s_p to represent the input of the EOWT prediction model. Based on the EOWT prediction model N_2 and the boundary conditions, the penalty is defined as follows:

$$r_{penal,t} = \begin{cases} 0, & \text{if } N_2(s_{p,t+1}) \in [T_2^l, T_2^u]; \\ -\lambda_2 \cdot e^{\left(T_2^l - N_2(s_{p,t+1})\right)}, & \text{if } N_2(s_{p,t+1}) < T_2^l; \\ -\lambda_2 \cdot e^{\left(N_2(s_{p,t+1}) - T_2^u\right)}, & \text{if } N_2(s_{p,t+1}) > T_2^u, \end{cases} \tag{4}$$

where λ_2 denotes the penalty pricing factor.

B. DQN-based Chiller Energy Consumption Optimization Algorithm

In a Markov process, the cumulative reward at a given time G_t is composed of the immediate reward R_t and the discounted future rewards. It can be represented as follows:

$$G_t = R_t + \gamma \cdot R_{t+1} + \gamma^2 \cdot R_{t+2} + \gamma^3 \cdot R_{t+3} + \cdots \quad (5)$$

We can derive an iterative formula for G:

$$G_t = R_t + \gamma \cdot G_{t+1}. \quad (6)$$

We denote the expected value of taking action a in state s as $Q(s,a)$, and there exists the following relationship between Q and G:

$$Q(s_t, a_t) = E\left[G_t \mid s = s_t, a = a_t\right]. \quad (7)$$

Based on (5-7) and utilizing Monte Carlo approximation, we derive the update rule for the Q-values in the Temporal Difference (TD) method:

$$Q(s_t,a_t) \leftarrow Q(s_t,a_t) + \alpha\left[R_t + \gamma \max_a Q(s_{t+1},a) - Q(s_t,a_t)\right]. \quad (8)$$

In DQN, we introduce ω as the parameters of the neural network:

$$f(s,a;\omega) = Q(s,a). \quad (9)$$

Based on equations (8-9), we derive the value function update mechanism for DQN as follows:

$$f(s_t,a_t;\omega) \leftarrow f(s_t,a_t) + \alpha\left[f^{Target} - f(s_t,a_t;\omega)\right], \quad (10)$$

where $f^{Target} = R_t + \gamma \max_a f(s_{t+1},a;\omega)$ is TD target.

We train a network, referred to the Q-Net, to approximate the value function, with the TD target being the feature values. The architecture of the Q-Net is shown in Fig. 3.

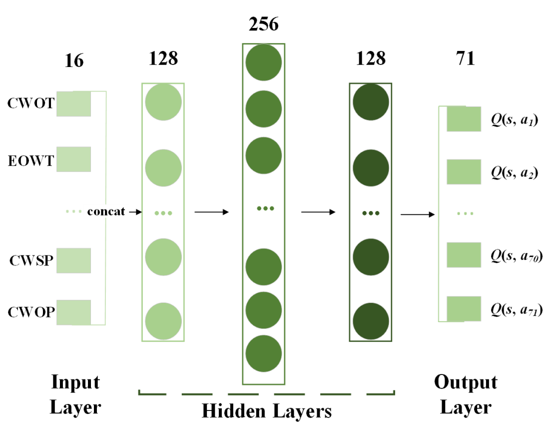

Fig. 3. The Q-Net in the chiller energy consumption optimization algorithm.

It is important to note that including the value function in the TD target implies that the feature values incorporate the output of the Q-Net. However, updating the feature values synchronously with the neural network during training can lead to instability. To address this, we introduce two networks: the training network, denoted as Q-Net(ω), and the target network, denoted as Q-Net(ω^-). Q-Net(ω) is responsible for computing the term $f(s_t,a_t;\omega)$ in (10) and updating its parameters using standard gradient descent, while Q-Net(ω^-) is used to compute the term $R_t + \gamma \max_a f(s_{t+1},a;\omega)$ in (10) and utilizes an older set of neural network parameters compared to Q-Net(ω). The loss function for training the Q-Net is defined as follows:

$$L(\omega) = E\left[\left(R_t + \gamma \max_a f(s_{t+1},a;\omega^-) - f(s_t,a_t;\omega)\right)^2\right]. \quad (11)$$

We employ the off-policy experience replay technique, which includes saving all experiences in a designated storage unit called an experience replay buffer. When conducting training, a group of experiences is chosen at random from this buffer to be used for the training process. This approach ensures that the samples satisfy the assumption of independent distribution while maximizing the utilization of samples.

Based on (11), we update the parameters of the Q-Net using the gradient descent method:

$$\omega = \omega - \alpha \cdot L(\omega) \cdot \frac{\partial f(s_t,a_t;\omega)}{\partial \omega}, \quad (12)$$

where α denotes the learning rate of the Q-Net.

The whole training procedure is shown in **Algorithm 1**.

Algorithm 1 DQN-based Chiller Energy Consumption Optimization Algorithm

Input: Initial state s_0.
Output: Terminal state s_{term}.
Initialization: Initialize Q-Net(ω) and Q-Net(ω^-) with the same randomly selected network parameters. Initialize the experience replay buffer R.
for episode $= 1 : N$ **do**
  Reset environment and obtain initial state s_0.
  for time step $t = 1 : T$ **do**
    Select an action using an ε-greedy policy (with CWOT adjustment) based on the current Q-Net(ω).
    Obtain the corresponding reward r_t and transition to the next state s_{t+1}.
    Store the tuple (s_t, a_t, r_t, s_{t+1}) in the experience replay buffer R.
    When the data volume in R reaches the minimal size, randomly sample M tuples from R.
    For the selected M tuples, calculate the target loss by (11), and update ω by (13).
    Utilize Q-Net(ω) to update Q-Net(ω^-).
  end for
end for

IV. Simulation Results

A. Neural Network Prediction

The hyperparameter settings for the MLP of the chiller energy consumption model include the following values: activation function (ReLU), optimization algorithm (Adam), learning rate ($3e$-4) and batch size (512). The changes in train loss and test accuracy throughout the training process are shown in

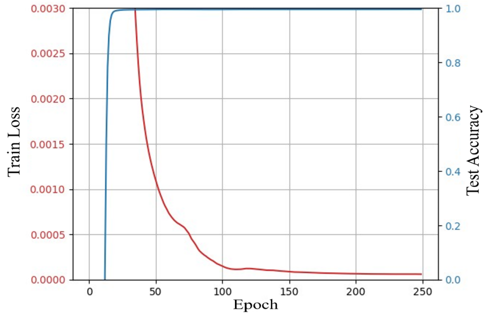

Fig. 4. Training process of the chiller energy consumption model.

Fig. 4. Additionally, Fig. 5 shows a comparison between the actual and predicted normalized energy consumption values.

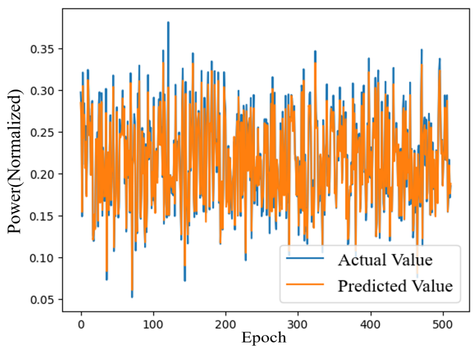

Fig. 5. The comparison between actual and predicted values.

R^2 (R-squared) is a commonly used metric to assess the goodness-of-fit of a model. It is calculated by comparing the squared differences between predicted and actual values with the total sum of squares. In real-world scenarios, a greater R^2 value signifies a stronger alignment between the model and the data. Testing the chiller energy consumption model on randomly selected test sets, we consistently observe an R^2 exceeding 0.97, and the loss remains at the order of 10^{-5}. Meanwhile, the EOWT prediction model, using the same hyperparameters as the chiller energy consumption model, also shows good performance with a loss below 0.007. These results demonstrate that our models serve as a reliable and secure foundation for subsequent optimization algorithms for chiller energy consumption.

B. Performance Evaluation of the Chiller Energy Consumption Optimization Algorithm

We conduct simulations to evaluate the effectiveness of the proposed optimization algorithm. We denote the CWOT of the initial state s_0 as CWOT_0, and similarly, other variables of s_0 follow this convention. The main simulation parameters are summarized in Table I.

We train our proposed DQN-based chiller energy consumption optimization algorithm and selected an algorithm based on completely random actions as the control. The hyperparameter settings for the Q-Net include the following values: optimization algorithm (Adam), activation function (ReLU), learning rate ($2e$-3), batch size (128) and discount factor γ (0.95). We

TABLE I
MAIN SIMULATION PARAMETERS

Parameter	Value	Parameter	Value
CWOT_0	33.14 °C	A_1^u	0.35 °C
EOWT_0	17.60 °C	T_2^l	17.1 °C
T_1^l	28.60 °C	T_2^u	17.7 °C
T_1^u	36.40 °C	λ_1	10
A_1^m	0.01 °C	λ_2	5
A_1^l	−0.35 °C		

compare the iterative processes of cumulative rewards (returns) obtained by the agent using two different algorithms, which are depicted in Fig. 6. It can be observed that our designed

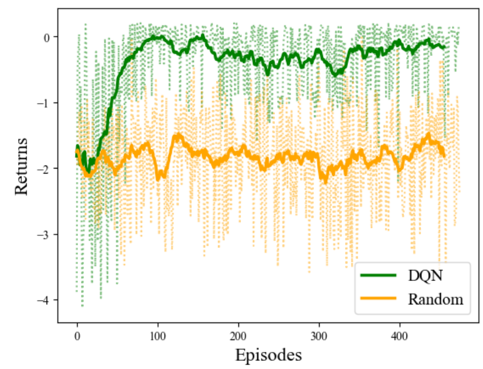

Fig. 6. The comparison of returns across iterations.

DQN-based optimization algorithm achieves convergence after approximately 100 episodes. Compared to the random action algorithm, the DQN-based optimization algorithm outperforms it in most training iterations. As the algorithm converges, the returns of our algorithm consistently and significantly surpass those of the random action algorithm. Furthermore, we record s_{term} in each episode and measure the corresponding percentage reduction in energy consumption compared to s_0. The changes in energy reduction percentage for both the DQN-based optimization algorithm and the random action algorithm as the training progresses are depicted in Fig. 7. It can be

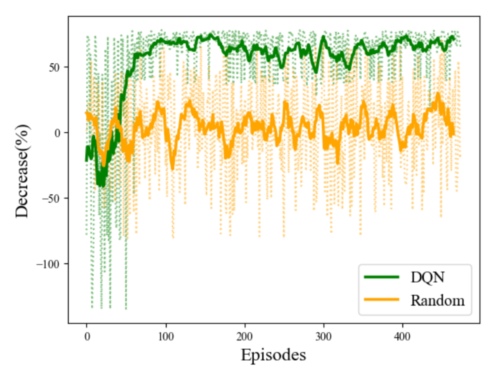

Fig. 7. The comparison of optimization ratios across iterations.

observed that our proposed algorithm guides the agent to stay in states with lower energy consumption for the majority of training episodes. After 100 training iterations, the energy reduction percentage stabilizes at about 77%.

In contrast to the policy-based DRL algorithm, we consider three classic global optimization algorithms as our benchmarks, including simulated annealing [14], Bayesian optimization [15] and genetic algorithm [16], in the given problem scenario. We compare the optimization results and convergence speed between our algorithm and the three optimization algorithms as depicted in Fig. 8-9.

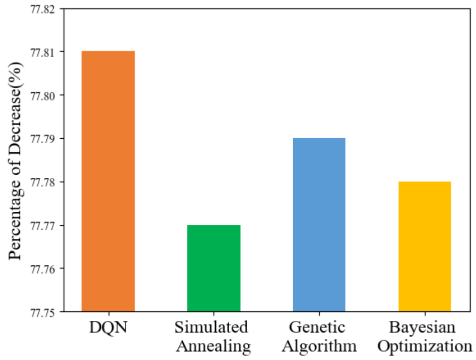

Fig. 8. The comparison of chiller energy consumption optimization percentages achieved by the four optimization algorithms.

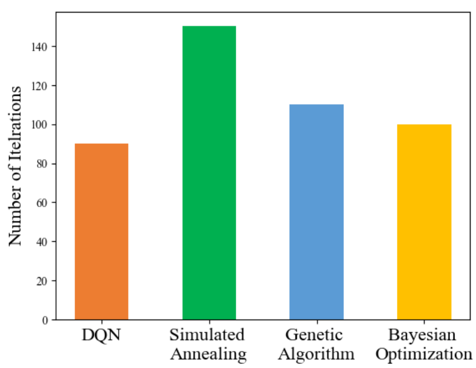

Fig. 9. The comparison of the number of iterations required for convergence among the four optimization algorithms.

It can be observed that our proposed algorithm converges faster and achieves a higher percentage of energy consumption decrease. Furthermore, it's important to underline that our proposed algorithm's potential for exceptional performance extends to intricate optimization scenarios. This optimistic projection is rooted in the heightened generalization prowess of DQN, allowing the algorithm to seamlessly transfer insights across various contexts. Moreover, the concurrent utilization of neural networks furnishes the algorithm with substantial parallel computational capacity, empowering it to efficiently address multifaceted challenges. Taken together, there's a strong foundation to anticipate that our algorithm, while maintaining its superior performance, will exhibit an even more pronounced advancement when applied to diverse and complex data center energy consumption optimization scenarios.

V. Conclusion

In this paper, we have proposed a DQN-based approach to reduce chiller energy consumption and optimize IoT-enabled data center PUE. Specifically, w e h ave d escribed t he overall framework of PUE optimization and developed a chiller energy consumption model based on neural network. We have transformed the chiller energy consumption optimization process into a Markov decision process. Moreover, we have designed a chiller energy consumption optimization algorithm using DQN and validated its effectiveness through simulations. The algorithm has demonstrated superior performance in terms of optimization results and convergence speed. In the future, our objective involves the development of an energy consumption optimization algorithm for IoT-enabled data center, which will encompass an expanded array of adjustable input parameters, allowing for a more comprehensive and refined approach to enhancing energy efficiency.

References

[1] A.-H. Fawaz, A. F. Y. Mohammed, L. I. Y. Laku, and R. Alanazi, "PUE or GPUE: a carbon-aware metric for data centers," in *2019 21st International Conference on Advanced Communication Technology (ICACT)*. IEEE, 2019, pp. 38–41.

[2] J. Ni and X. Bai, "A review of air conditioning energy performance in data centers," *Renewable and sustainable energy reviews*, vol. 67, pp. 625–640, 2017.

[3] Z. Cao, X. Zhou, H. Hu, Z. Wang, and Y. Wen, "Towards a systematic survey for carbon neutral data centers," *IEEE Communications Surveys & Tutorials*, 2022.

[4] F. W. Yu and K. Chan, "Optimization of water-cooled chiller system with load-based speed control," *Applied Energy*, vol. 85, no. 10, pp. 931–950, 2008.

[5] R. Zhou, Z. Wang, C. E. Bash, A. McReynolds, C. Hoover, R. Shih, N. Kumari, and R. K. Sharma, "A holistic and optimal approach for data center cooling management," in *Proceedings of the 2011 American Control Conference*. IEEE, 2011, pp. 1346–1351.

[6] M. K. J. Ramphela, P. A. Owolawi, T. Mapayi, and G. Aiyetoro, "Internet of things (iot) integrated data center infrastructure monitoring system," in *2020 International Conference on Artificial Intelligence, Big Data, Computing and Data Communication Systems (icABCD)*. IEEE, 2020, pp. 1–6.

[7] T. P. Lillicrap, J. J. Hunt, A. Pritzel, N. Heess, T. Erez, Y. Tassa, D. Silver, and D. Wierstra, "Continuous control with deep reinforcement learning," *arXiv preprint arXiv:1509.02971*, 2015.

[8] Y. Li, Y. Wen, D. Tao, and K. Guan, "Transforming cooling optimization for green data center via deep reinforcement learning," *IEEE transactions on cybernetics*, vol. 50, no. 5, pp. 2002–2013, 2019.

[9] B. Qian, H. Zhou, T. Ma, K. Yu, Q. Yu, and X. S. Shen, "Heterogeneous multi-operator spectrum sharing architecture for massive iot access with noma," in *2020 IEEE 92nd Vehicular Technology Conference (VTC2020-Fall)*. IEEE, 2020, pp. 1–6.

[10] B. Qian, H. Zhou, T. Ma, K. Yu, Q. Yu, and X. Shen, "Multi-operator spectrum sharing for massive IoT coexisting in 5G/B5G wireless networks," *IEEE Journal on Selected Areas in Communications*, vol. 39, no. 3, pp. 881–895, 2020.

[11] H. D. Vu, K. S. Chai, B. Keating, N. Tursynbek, B. Xu, K. Yang, X. Yang, and Z. Zhang, "Data driven chiller plant energy optimization with domain knowledge," in *Proceedings of the 2017 ACM on Conference on Information and Knowledge Management*, 2017, pp. 1309–1317.

[12] W. Zhang, Y. Wen, Y. W. Wong, K. C. Toh, and C.-H. Chen, "Towards joint optimization over ICT and cooling systems in data centre: A survey," *IEEE Communications Surveys & Tutorials*, vol. 18, no. 3, pp. 1596–1616, 2016.

[13] Y. Xu, F. Li, and A. Asgari, "Prediction and optimization of heating and cooling loads in a residential building based on multi-layer perceptron neural network and different optimization algorithms," *Energy*, vol. 240, p. 122692, 2022.

[14] K. A. Dowsland and J. Thompson, "Simulated annealing," *Handbook of natural computing*, pp. 1623–1655, 2012.

[15] P. I. Frazier, "A tutorial on Bayesian optimization," *arXiv preprint arXiv:1807.02811*, 2018.

[16] S. Mirjalili and S. Mirjalili, "Genetic algorithm," *Evolutionary Algorithms and Neural Networks: Theory and Applications*, pp. 43–55, 2019.