Hindawi
Wireless Communications and Mobile Computing
Volume 2022, Article ID 7424854, 9 pages
https://doi.org/10.1155/2022/7424854

Research Article

UCB-Based Route and Power Selection Optimization for SDN-Enabled Industrial IoT in Smart Grid

Xiaoyue Li, Chaoqun Zhou, Zilong Liang, Qiang Yu, Xiankai Chen, and Zhiyuan He

State Grid Qingdao Power Supply Company, Qingdao 266000, China

Correspondence should be addressed to Xiaoyue Li; lxy-qdpower@outlook.com

Received 28 November 2021; Revised 18 December 2021; Accepted 24 December 2021; Published 20 March 2022

Academic Editor: Xingwang Li

As an essential building block for smart grid, the industrial internet of things (IIoT) plays a significant role in providing powerful sensing capability and ubiquitous connectivity for differentiated power services. The rapid development of smart grid imposes higher data monitoring and transmission requirements in terms of delay and energy efficiency. However, due to the severe electromagnetic interference (EMI) caused by massive electrical equipment, the transmission performance of IIoT becomes inferior. The traditional single-hop transmission mode evolves towards a multihop cooperation mode to satisfy differentiated quality of service (QoS) requirements. In this paper, we propose an upper confidence bound- (UCB-) based joint route and power selection optimization algorithm to support multihop cooperation mode evolution, which adopts a software-defined networking- (SDN-) enabled IIoT network framework to simplify network configuration and management. Compared with existing local-side-information-based route selection (LSI-RS) and random route selection (RRS) algorithms, simulation results demonstrate that the proposed algorithm has superior performances in total delay, energy efficiency, and utility.

1. Introduction

The industrial internet of things (IIoT) is an essential building block for smart grid, which has powerful sensing capability and ubiquitous connectivity. With the development of smart grid, a large number of IIoT devices need to be deployed to collect information such as voltage, current, power, temperature, and humidity and transmit the information back for real-time analysis. IIoT has strict requirements on transmission delay, energy efficiency, and network coverage [1]. However, electrical equipment in smart grid emits electromagnetic interference (EMI), which affects the transmission performance of IIoT. Therefore, the traditional single-hop transmission mode needs to evolve towards a multihop cooperation mode to satisfy the quality of service (QoS) requirements [2]. IIoT utilizes massive devices laid in different routes to form a mesh network for multihop transmission. In multihop transmission, dynamic route selection can avoid worse routes with long distance and low quality and enable IIoT to reduce transmission delay and enhance energy efficiency.

Routing selection needs to be optimized according to the dynamic network environment. However, the traditional network architecture with tight coupling between control and data planes cannot adapt to complex IIoT application scenarios. Software-defined networking (SDN) provides a solution by separating the control plane from the data plane [3]. SDN can manage and control IIoT networks through a standard and open programmable interface, which supports more efficient and flexible route selection solutions [4]. However, the research on route selection optimization for SDN-enabled IIoT in smart grid still faces many challenges, which are summarized below.

First, considering the highly time-varying channel states and complex EMI, the global state information (GSI) is incomplete [5]. Traditional GSI-based route selection optimization cannot be applied. Second, IIoT devices based on battery have strict requirements on energy efficiency. Improving energy efficiency via dynamic power selection not only makes route selection more complicated but also possibly leads to larger transmission delay. Therefore, how to meet differentiated

QoS requirements through joint route and power selection optimization is also a challenge. Finally, electric equipment such as inverters and insulation switches emits EMI [6, 7], which greatly reduces QoS performance and brings severe challenges for joint route and power selection optimization.

Route selection of IoT has always been a research hotspot. In [8], Desuo et al. employed an improved Dijkstra algorithm to find the shortest path between two consecutive points for IoT networks. In [9], He et al. proposed an energy-aware route selection algorithm for simultaneous information and power transfer to decrease energy consumption. However, these works do not consider SDN architecture and only consider a single QoS metric. In [10], Saha et al. proposed a traffic-aware QoS route selection scheme by exploiting the flow-based nature of SDN and obtained the optimal route based on Yen's K-shortest path algorithm. In [11], Li et al. proposed an SDN-enabled IoT adaptive transmission architecture for different delay flow situations. However, these works assume that perfect GSI is available, which is not applicable for smart grid with incomplete information.

Upper confidence bound (UCB) as a reinforcement learning algorithm has emerged as a powerful solution to address problems without perfect GSI [12]. In [13], Sun et al. designed an energy-aware mobility management (EMM) scheme based on UCB to optimize energy consumption. In [14], Maghsudi and Stanczak proposed two joint power and channel selection strategies based on UCB to maximize energy efficiency. However, these works only consider energy consumption optimization, which ignore delay and other QoS requirements. In [15], Zhao et al. proposed a delay minimization algorithm based on UCB, but neglected the joint optimization of energy consumption and delay. In [16], Bae et al. proposed a downlink network routing algorithm based on UCB to jointly optimize throughput and delay, but ignored the influence of complex EMI and service priority. Moreover, all the abovementioned works do not consider the impact of complex EMI and service priority of smart grid on the joint optimization of route and power selection.

To address the abovementioned challenges, we propose a UCB-based joint route and power selection optimization algorithm. Firstly, considering the influence of EMI, we construct an SDN-based multihop IIoT framework and formulate the joint route and power selection optimization problem. The objective is to maximize the overall network utility function under the threshold constraints of signal-to-interference-plus-noise ratio (SINR) and energy efficiency. Second, we model the joint optimization problem as a multiarmed bandit (MAB) problem, where the options of route and power are combined to form an arm. Finally, we utilize UCB to learn the optimal route and power combination based on local and historical information. The main contributions of this work are summarized as follows:

(i) We propose an SDN-enabled multihop IIoT framework for smart grid, which greatly simplifies network management through separating control and data planes. In addition, the control plane also supports the configuration of intelligent route and power selection algorithms.

(ii) The route and power options are combined to form a set of arms in MAB. The proposed algorithm dynamically learns the optimal combination by interacting with the environment.

(iii) Through dynamically adjusting the values of weight parameters, the proposed algorithm can satisfy differentiated QoS requirements of smart grid by adjusting the tradeoff between delay, energy efficiency, and service priority.

The rest of this paper is organized as follows: Section 2 describes the system model and problem formulation. The proposed joint route and power selection algorithm is introduced in Section 3. Section 4 provides simulation results. Finally, the conclusion is provided in Section 5.

2. System Model

In this section, the system model and the problem formulation are introduced.

2.1. Network Model of SDN-Enabled IIoT for Smart Grid. The SDN-enabled IIoT for smart grid is shown in Figure 1, which consists of two planes, i.e., the data plane and the control plane [17]. The data plane mainly contains IIoT devices which provide data forwarding services. The control plane mainly contains the SDN controller, which locates in the gateway. The SDN controller can obtain IIoT network topology, learn the optimal route and power selection strategy, and send the strategy to the IIoT source device (SD) [18].

The SDN-enabled IIoT network topology is represented by a directed graph $\mathcal{G} = (\mathcal{V}, \mathcal{L})$ [19], where $\mathcal{V}$ denotes I IIoT devices. The set is defined as $\mathcal{V} = \{v_1, \ldots, v_i, \ldots, v_I\}$. v_1 and v_I are the SD and destination device (DD). v_i, $i = 2, 3, \ldots, I-1$, is the relay device. $\mathcal{L}$ denotes physical links, and the set is defined as $\mathcal{L} = \{L(v_i, v_j) | v_i \in \mathcal{V}, v_j \in \mathcal{V}_{N(v_i)}\}$, where $\mathcal{V}_{N(v_i)}$ is the set of devices connected with v_i. There exist M routes between v_1 and v_I, and the set is represented as $\mathcal{F} = \{f_1, \ldots, f_m, \ldots, f_M\}$. Each route consists of K devices, which are SD v_1, DD v_I, and $K-2$ relay devices. The set of devices in f_m is denoted as $\mathcal{A}_m = \{d_m^1, \ldots, d_m^k, \ldots, d_m^k\}$ in the order from SD to DD, where $\mathcal{A}_m \subseteq \mathcal{V}$, $L(d_m^k, d_m^{k+1}) \in \mathcal{L}, 1 \le k \le K$.

In this paper, the set of T time slots is represented as $\mathcal{T} = \{1, \ldots, t, \ldots, T\}$. The slot length depends on the transmission delay from SD to DD [20]. At the beginning of the t-th slot, v_1 generates a data packet of size U_t, $U_{\min} \le U_t \le U_{\max}$, which needs to be transmitted to v_I. Each data packet can only be transmitted in one route [21]. The transmission is unsuccessful if the delay exceeds τ.

2.2. Delay Model. We assume that the data packets are transmitted by wireless channels. We denote P_h as the transmission power, which contains H levels. The set of transmission power levels is given by

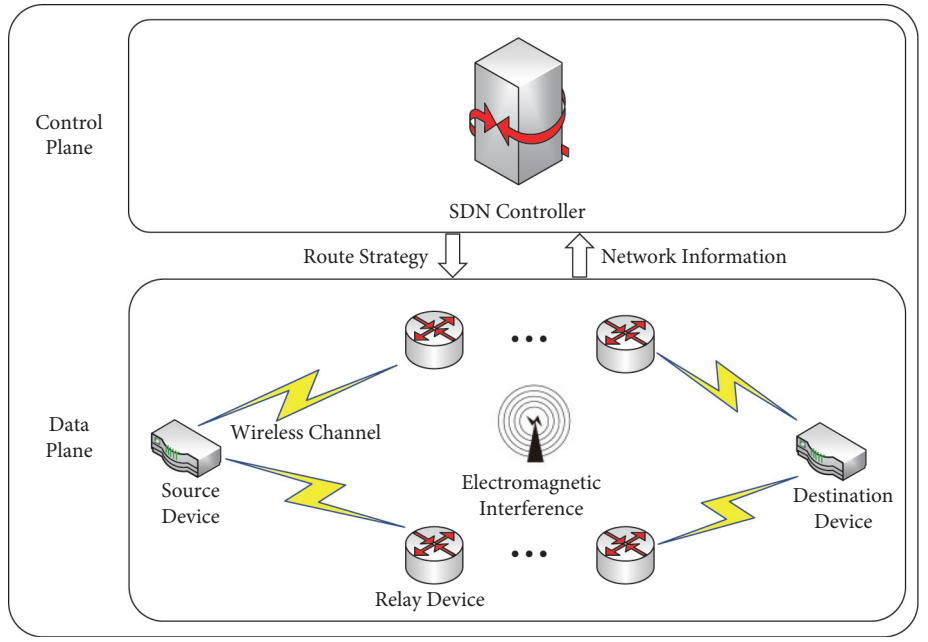

Figure 1: SDN-enabled IIoT framework for smart grid.

$$\mathscr{P} = \left\{ P_{\min}, \ldots, P_{\min} + \frac{(h-1)(P_{\max} - P_{\min})}{H-1}, \ldots, P_{\max} \right\}. \tag{1}$$

The achievable transmission rate from d_m^k to d_m^{k+1} is given by

$$R_t^{d_m^k, d_m^{k+1}}(h) = B(m)\log_2\left(1 + \mathrm{SINR}^{d_m^k, d_m^{k+1}}(h)\right), \tag{2}$$

where $B(m)$ is the transmission bandwidth of route f_m. $\mathrm{SINR}^{d_m^k, d_m^{k+1}}(h)$ is the SINR [22] between d_m^k and d_m^{k+1} and is given by

$$\mathrm{SINR}^{d_m^k, d_m^{k+1}}(h) = \frac{P_h g_t^{d_m^k, d_m^{k+1}}}{\sigma_0 + \lambda_t^{d_m^k, d_m^{k+1}}}, \tag{3}$$

where $g_t^{d_m^k, d_m^{k+1}}$ is the channel gain. σ_0 is the noise power. $\lambda_t^{d_m^k, d_m^{k+1}}$ is the EMI power.

We denote the power selection variable as $y_t(h) \in \{0, 1\}$. $y_t(h) = 1$ represents that v_1 selects P_h; otherwise, $y_t(h) = 0$. The transmission delay from d_m^k to d_m^{k+1} and the total forwarding delay on the route f_m are given by

$$D_t^{d_m^k, d_m^{k+1}}(h) = \frac{U_t}{R_t^{d_m^k, d_m^{k+1}}(h)},$$
$$D_t(m, h) = \sum_{k=1}^{K-1} D_t^{d_m^k, d_m^{k+1}}(h). \tag{4}$$

We denote the route selection variable $x_t(m) \in \{0, 1\}$. $x_t(m) = 1$ represents that v_1 selects f_m; otherwise, $x_t(m) = 0$ [23]. The total forwarding delay is given by

$$\tau_t = \sum_{m=1}^{M} \sum_{h=1}^{H} y_t(h) x_t(m) D_t(m, h). \tag{5}$$

2.3. Energy Efficiency Model. The energy consumption for data packet transmission from d_m^k to d_m^{k+1} and the total energy consumption on route f_m are given by

$$E_t^{d_m^k, d_m^{k+1}}(h) = D_t^{d_m^k, d_m^{k+1}}(h) P_h,$$
$$E_t(m, h) = \sum_{k=1}^{K-1} E_t^{d_m^k, d_m^{k+1}}(h). \tag{6}$$

We define $\gamma_t(m, h)$ as the energy efficiency of data packet transmission on route f_m with power $P_{\min} + ((h-1)(P_{\max} - P_{\min})/(H-1))$ in the t-th time slot, which is given by

$$\gamma_t(m, h) = \frac{U_t}{B(m) E_t(m, h)}. \tag{7}$$

Therefore, the total energy efficiency is given by

$$\alpha_t = \sum_{m=1}^{M} \sum_{h=1}^{H} x_t(m) y_t(h) \gamma_t(m, h). \tag{8}$$

2.4. Problem Formulation. Since the data packets have different QoS requirements, the service priority needs to be taken into account. We use η_t to represent the priorities of different data packets. We define the overall network utility function related to the total forwarding delay, service priority, and total energy efficiency as

$$\Phi = \frac{1}{T} \sum_{t=1}^{T} \left(\alpha_t + V \frac{\eta_t}{\tau_t} \right), \tag{9}$$

where V is the weight used to balance the order of magnitude.

Therefore, the objective is to maximize Φ by optimizing the route and power selection strategies. The optimization problem is formulated as

$$\begin{aligned}
&\mathbf{P1}: \max_{\{x_t(m), y_t(h)\}} \Phi, \\
&\text{s.t. } C_1: \sum_{m=1}^{M} x_t(m) = 1, \quad \forall t \in \mathcal{T}, \\
&C_2: \sum_{h=1}^{H} y_t(h) = 1, \quad \forall t \in \mathcal{T}, \\
&C_3: P_h \in \mathcal{P}, \quad h = 1, \ldots, H, \\
&C_4: \mathrm{SINR}^{d_m^k, d_m^{k+1}}(h) \geq \mathrm{SINR}_{\min}, \quad \forall d_m^k, d_m^{k+1} \in \mathcal{A}_m, \\
&C_5: \alpha_t \geq \alpha_{\min}, \quad \forall t \in \mathcal{T},
\end{aligned} \tag{10}$$

where $\mathrm{SINR}_{\min}$ and $\alpha_{\min}$ represent the thresholds of SINR and energy efficiency, respectively. C_1 is the route selection constraint; i.e., each data packet can only select one route. C_2 is the power selection constraint; i.e., each data packet can only select one power level. C_3 is the transmission power constraint. C_4 is the SINR constraint. C_5 is the energy efficiency constraint.

3. UCB-Based Route and Power Selection Optimization for SDN-Enabled Industrial IoT in Smart Grid

It is impractical to obtain the perfect GSI due to the dynamic network topology and complex EMI, and IIoT devices should optimize route and power selection based on the local-side information. MAB is an effective solution to solve decision-making problems with incomplete information [24]. In each slot, the decision maker pulls an arm. Then, the pulled arm generates a reward. The decision maker's goal is to maximize the cumulative reward.

We transform **P1** into an MAB problem. The decision maker, arm, and reward are modeled as follows:

(i) Decision maker: the decision maker generates the decision. In this paper, we define the SDN controller as the decision maker.

(ii) Arm: we define $\mathscr{C} = \{c_{1,1}, \ldots, c_{m,h}, \ldots, c_{M,H}\}$ as the set of arms which satisfy $|\mathscr{C}| = M \times H$, where $|\mathscr{C}|$ represents the number of elements in $\mathscr{C}$. The arm $c_{m,h}$ represents the route f_m and power P_h.

(iii) Reward: we define a reward function $\theta_t(c_{m,h})$ to represent the reward obtained by selecting $c_{m,h}$, which is given by

$$\theta_t(c_{m,h}) = \begin{cases} \gamma_t(m,h) + V\dfrac{\eta_t}{D_t(m,h)}, & \text{satisfy } C_3 \text{ and } C_4, \\ 0, & \text{otherwise.} \end{cases} \tag{11}$$

If $\mathrm{SINR}^{d_m^k, d_m^{k+1}}(h) \geq \mathrm{SINR}_{\min}$ and $\alpha_t \geq \alpha_{\min}$, the reward is $\gamma_t(m,h) + V(\eta_t / D_t(m,h))$. Otherwise, the reward is zero.

We propose a UCB-based joint route and power selection algorithm for SDN-enabled IIoT in smart grid to address the MAB problem. UCB is a low-complexity learning-based algorithm to balance exploitation and exploration [25]. The proposed algorithm enables the SDN controller to take action based on local state information such as delay. Afterwards, the obtained reward and updated state information is perceived by the SDN controller for the next selection. The implementation of the proposed algorithm is shown in Figure 2.

The proposed algorithm consists of three phases, which is summarized in Algorithm 1.

(i) Phase I: $x_t(m)$, $y_t(h)$, $\theta_t(c_{m,h})$, and $\hat{n}_t(c_{m,h})$ are initialized as zero. When $t \leq |\mathscr{C}|$, the controller sequentially selects each arm and obtains the initial value.

(ii) Phase II: based on (12), the preference of the SD towards arm $c_{m,h}$ in the t-th slot is given by

$$\tilde{\theta}_t(c_{m,h}) = \overline{\theta}_{t-1}(c_{m,h}) + \omega\sqrt{\frac{\ln t}{\hat{n}_{t-1}(c_{m,h})}}, \tag{12}$$

where $\overline{\theta}_{t-1}(c_{m,h})$ is the average reward of $c_{m,h}$ up to the $(t-1)$-th time slot. $\hat{n}_{t-1}(c_{m,h})$ is the number of times to select $c_{m,h}$. ω is the weight of exploration. The second term allows the controller to explore arms with selections to improve estimation and to focus on the exploitation when arms have been estimated enough. After obtaining $\tilde{\theta}_t(c_{m,h})$, the selected arm is given by

$$c_{m^*,h^*} = \mathrm{argmax}_{\mathscr{C}}\{\tilde{\theta}_t(c_{m,h})\}. \tag{13}$$

c_{m^*,h^*} represents SD selects f_{m^*} and P_{h^*}, which is given by

$$c_{m^*,h^*} \Rightarrow \{x_t(m^*) = 1, y_t(h^*) = 1\}. \tag{14}$$

(iii) Phase III: the controller observes delay and energy efficiency performances as well as service priority. Then, $\theta_t(c_{m^*,h^*})$ is updated as (11). Accordingly, $\overline{\theta}_t(c_{m,h})$ and $\hat{n}_t(c_{m,h})$ are updated as

$$\overline{\theta}_t(c_{m,h}) = \frac{\overline{\theta}_{t-1}(c_{m,h})\hat{n}_{t-1}(c_{m,h}) + \theta_t(c_{m,h})x_t(m)y_t(h)}{\hat{n}_{t-1}(c_{m,h}) + x_t(m)y_t(h)}, \tag{15}$$

$$\hat{n}_t(c_{m,h}) = \hat{n}_{t-1}(c_{m,h}) + x_t(m)y_t(h). \tag{16}$$

Finally, the algorithm terminates until $t > T$.

4. Simulation Results

In this section, we firstly introduce the simulation parameter setting. Then, the simulation analysis is described.

4.1. Simulation Parameter Setting. In this section, we evaluate the proposed algorithm through simulations. The

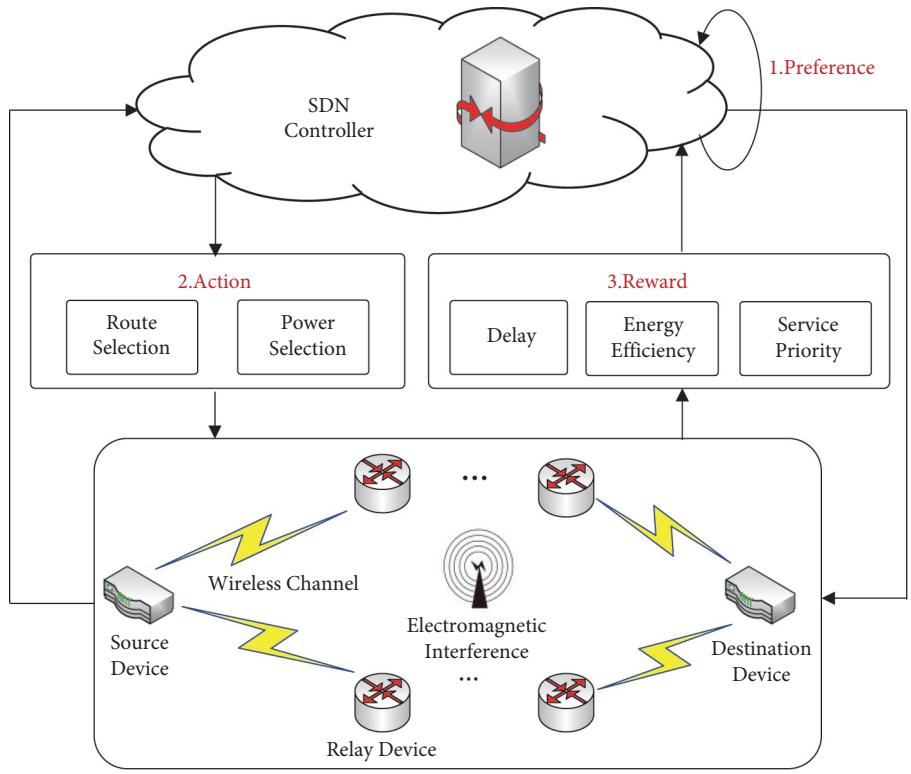

Figure 2: The implementation of the proposed algorithm.

(1) **Input**: ω, $\{\eta_t\}$, $\{D_t(m,h)\}$, $\{\gamma_t(m,h)\}$.
(2) **Phase 1**:
(3) Set $x_t(m) = 0$, $y_t(h) = 0$, $\theta_t(c_{m,h}) = 0$, and $\hat{n}_t(c_{m,h}) = 0$, $\forall t \in \mathcal{T}, \forall f_m \in \mathcal{F}, \forall P_h \in \mathcal{P}$.
(4) **for** $t = 1$ to $|\mathcal{C}|$ **do**
(5) Select arms sequentially, and obtain the initial values.
(6) **end for**
(7) **for** $t = |\mathcal{C}| + 1$ to T **do**
(8) **Phase 2**:
(9) Calculate the preference of the SD towards arm $c_{m,h}$ as (12).
(10) Select c_{m^*,h^*} based on (13).
(11) **Phase 3**:
(12) Observe delay and energy efficiency performances.
(13) Calculate $\theta_t(c_{m^*,h^*})$ based on (11).
(14) Update $\bar{\theta}_t(c_{m,h})$ and $\hat{n}_t(c_{m,h})$ based on (15) and (16).
(15) **end for**

Algorithm 1: UCB-based route and power selection optimization for SDN-enabled industrial IoT in smart grid.

considered IIoT route topology is shown in Figure 3, which includes 9 IIoT devices and 6 routes. v_1 and v_9 are the SD and DD, respectively. The distances of adjacent devices on each route are shown in Table 1. In the case of large-scale fading, the channel gain is calculated according to $g_{i,j}(t) = 127 + 30\log(r_{i,j})$ [26], where $r_{i,j}$ is the distance between v_i and v_j. The EMI varies from 28 dBm to 30 dBm. The service priority η_t is set as $[0.1, 0.2, 0.3, 0.4, 0.5]$. The setting of simulation parameters is summarized in Table 2 [27, 28]. We consider two existing algorithms for comparison. The first one is the UCB-based route selection algorithm named UCB-RS [29]. The other one is the shortest route selection algorithm named SRS [30]. Both UCB-RS and SRS neglect the optimization of power selection.

4.2. Simulation Analysis. Figure 4 shows the average utility versus time slot. Compared with UCB-RS and SRS, the simulation result demonstrates that the proposed algorithm improves the performance of utility by27.28% and 37.31%, respectively. The reason is that the proposed algorithm jointly optimizes the route and power selection. In contrast,

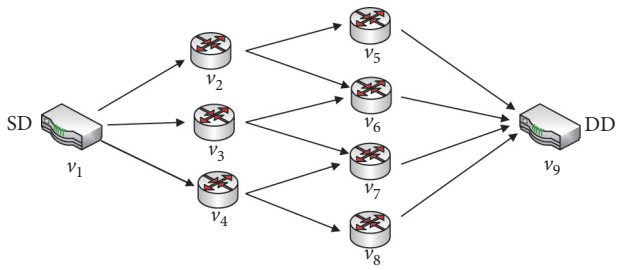

Figure 3: IIoT route topology.

Table 1: Distance between devices.

Path	Distance (m)
$v_1 \longrightarrow v_2$	50
$v_1 \longrightarrow v_4$	70
$v_2 \longrightarrow v_6$	50
$v_3 \longrightarrow v_7$	50
$v_4 \longrightarrow v_8$	70
$v_6 \longrightarrow v_9$	30
$v_8 \longrightarrow v_9$	70
$v_1 \longrightarrow v_3$	30
$v_2 \longrightarrow v_5$	50
$v_3 \longrightarrow v_6$	30
$v_4 \longrightarrow v_7$	50
$v_5 \longrightarrow v_9$	70
$v_7 \longrightarrow v_9$	70

Table 2: Simulation parameter settings.

Simulation parameters	Value
The number of time slots T	1000
The number of IIoT devices M	9
Transmission power P_h (W)	$[0.1, 0.2, 0.3, 0.4, 0.5]$
Packet size U_t (Mbits)	$[1, 2]$
The number of transmission power levels H	5
Transmission bandwidth B (MHz)	$[2, 3]$
Noise power σ_0 (dBm)	-114
EMI λ_t (dBm)	$[28, 30]$
Service priority η_t	$[0.1, 0.2, 0.3, 0.4, 0.5]$
Exploration weight ω	2
The weight V	12

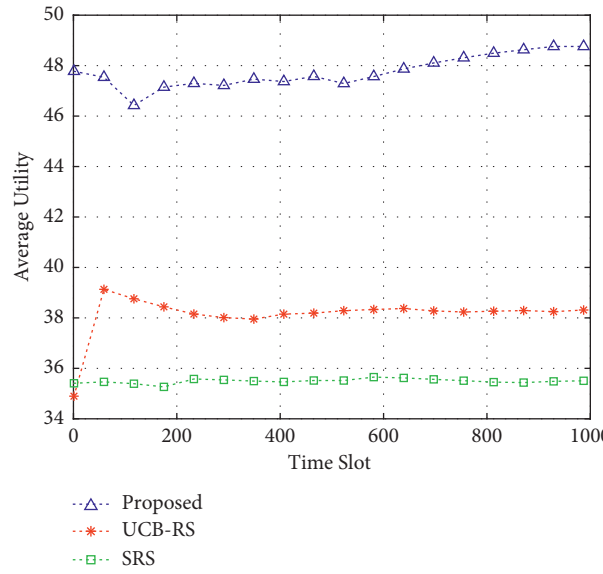

Figure 4: **Q2-2:** average utility versus time slot.

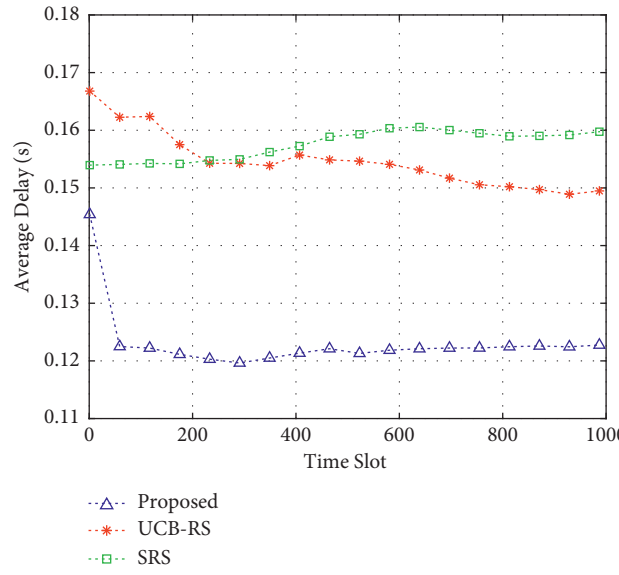

Figure 5: Average delay versus time slot.

UCB-RS neglects the power selection. SRS always selects the shortest route, which cannot overcome the adverse impact caused by the dynamic change of channel state, thereby performing the worst.

Figure 5 shows the average delay versus time slot. The simulation result shows that the proposed algorithm outperforms UCB-RS and SRS by 17.93% and 23.17% in delay performance, respectively. Both UCB-RS and SRS do not take the optimization of power selection into consideration, which result in worse delay performance.

Figure 6 shows the average energy efficiency versus time slot. Compared with UCB-RS and SRS, the proposed algorithm improves the performance of energy efficiency by 28.61% and 51.15%, respectively. The proposed algorithm can select suitable power to optimize energy efficiency.

Figure 7 shows the ratio of optimal route selection versus time slot. SRS performs the worst. The reason is that the proposed algorithm and UCB-RS can dynamically adjust the route selection strategy. However, SRS always selects the shortest route fixedly and cannot get rid of the adverse impact of EMI.

Figures 8–10 show the average energy efficiency, average delay, and average utility versus $\mathrm{SINR}_{\min}$. With the increase of $\mathrm{SINR}_{\min}$, the energy efficiency and delay of the proposed algorithm decrease, while the utility increases

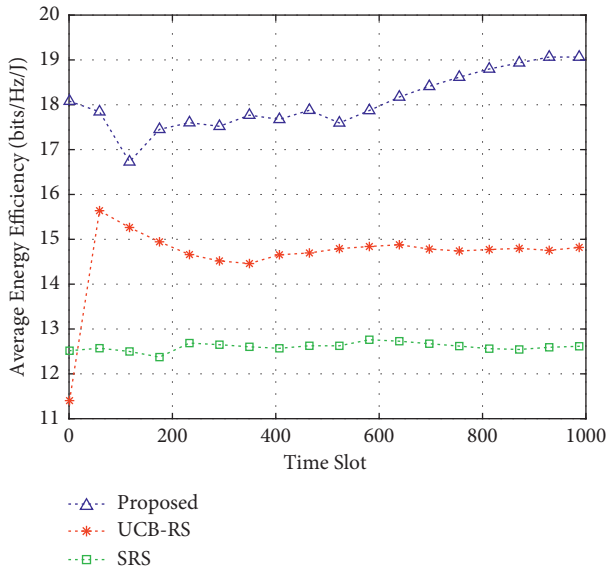

Figure 6: Average energy efficiency versus time slot.

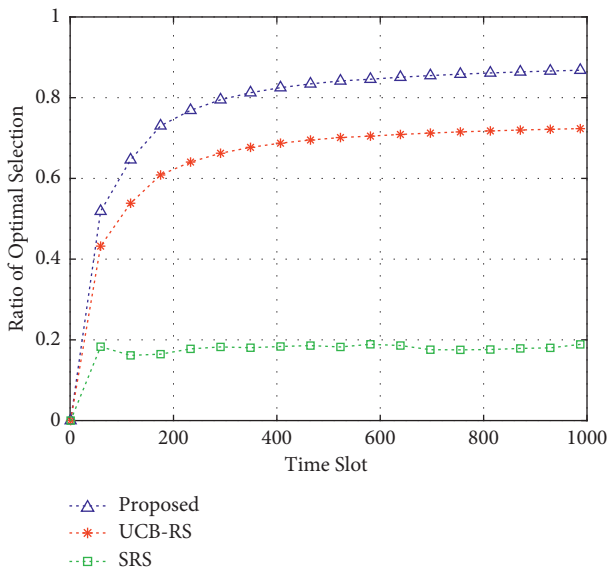

Figure 7: Ratio of optimal selection versus time slot.

first and then decreases. When $SINR_{min} = 14$ dB, the utility reaches the maximum value. The performance of UCB-RS fluctuates. The reason is that the proposed algorithm can learn to optimize power selection to meet more stringent SINR constraint. UCB-RS neglects power selection, which makes it difficult to adapt to different SINR constraints.

Figure 11 shows the impact of V on the delay and energy efficiency. As V increases, the proposed algorithm lays more emphasis on energy efficiency rather than delay. The proposed algorithm can dynamically balance the tradeoff between energy efficiency and delay. Moreover, the simulation results provide a reference for the setting of the weight V.

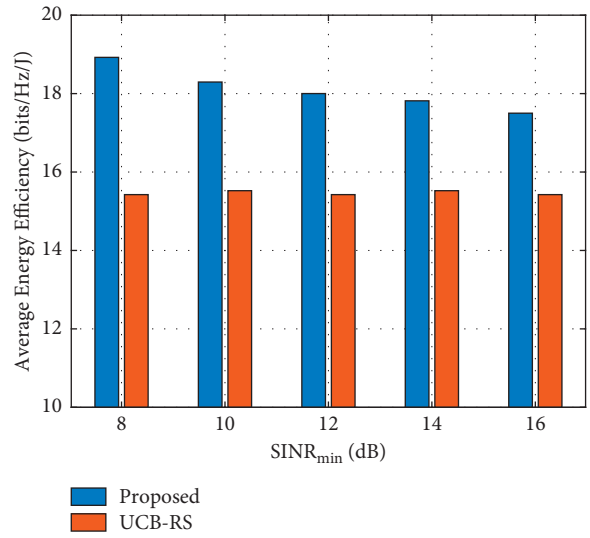

Figure 8: Average energy efficiency versus $SINR_{min}$.

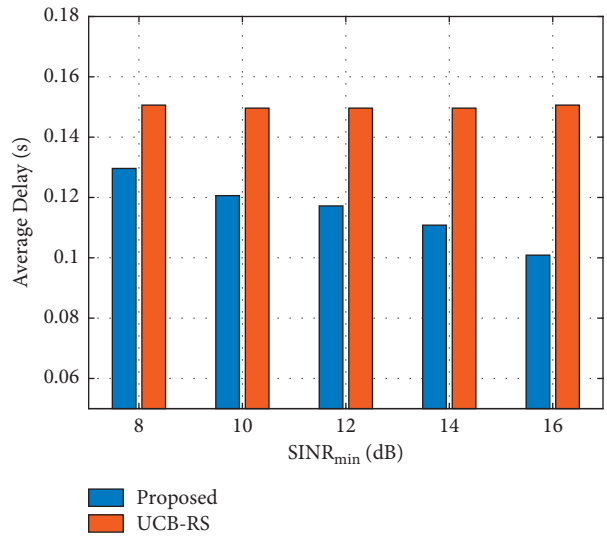

Figure 9: Average delay versus $SINR_{min}$.

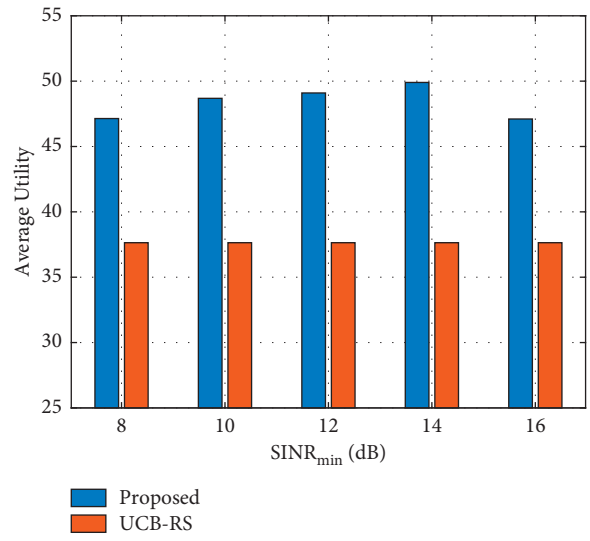

Figure 10: Average utility versus $SINR_{min}$.

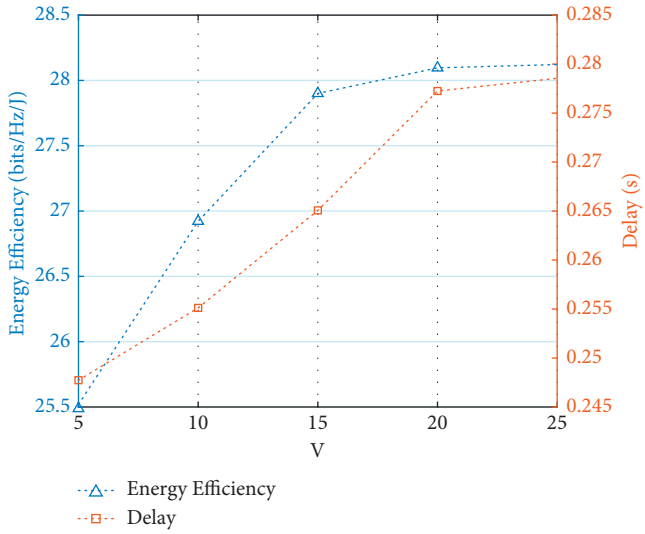

FIGURE 11: The impact of V on the delay and energy efficiency.

5. Conclusions

In this paper, we proposed an UCB-based joint route and power selection optimization algorithm for SDN-enabled IIoT. The proposed algorithm can effectively optimize route and power selection strategies based only on local information and historical observations. It can provide a low-complexity route and power selection strategy while maximizing the overall network utility. Simulation results show that the proposed algorithm has superior performances in delay, energy efficiency, and utility. Compared with existing LSI-RS and RRS algorithms, the proposed algorithm reduces the delay by 17.93% and 23.17%, improves the utility by 27.28% and 37.31%, and improves the energy efficiency by 28.61% and 51.15%. In the future, we will use deep reinforcement learning to optimize the multidimensional resource allocation in SDN-enabled IIoT.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare no conflicts of interest.

Acknowledgments

This work was supported by the Science and Technology Project of the State Grid Shandong Power Supply Company under grant no. 5206021900VV.

References

[1] L. You, Y. Huang, D. Zhang, Z. Chang, W. Wang, and X. Gao, "Energy efficiency optimization for multi-cell massive MIMO: centralized and distributed power allocation algorithms," *IEEE Transactions on Communications*, vol. 69, no. 8, pp. 5228–5242, 2021.

[2] G. Chen, J. P. Coon, and S. E. Tajbakhsh, "Secure routing for multihop Ad Hoc networks with inhomogeneous eavesdropper clusters," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 11, pp. 10660–10670, 2018.

[3] X. Li, Y. Zheng, M. D. Alshehri et al., "Cognitive AmBC-NOMA IoV-MTS networks with IQI: reliability and security analysis," *IEEE Transactions on Intelligent Transportation Systems*, pp. 1–12, 2021.

[4] Z. Zhou, J. Gong, Y. He, and S. E. Tajbakhsh, "Software defined machine-to-machine communication for smart energy management," *IEEE Communications Magazine*, vol. 55, no. 10, pp. 52–60, 2017.

[5] X. Li, J. Li, Y. Liu, Z. Ding, and A. Nallanathan, "Residual transceiver hardware impairments on cooperative NOMA networks," *IEEE Transactions on Wireless Communications*, vol. 19, no. 1, pp. 680–695, 2020.

[6] X. Li, M. Zhao, M. Zeng et al., "Hardware impaired ambient backscatter NOMA system: reliability and security," *IEEE Transactions on Communications*, vol. 69, no. 4, pp. 2723–2736, 2021.

[7] B. Wang, Y. Sun, Q. Cao, S. Li, and Z. Sun, "Bandwidth slicing for socially-aware D2D caching in SDN-enabled networks," *IEEE Access*, vol. 6, no. 99, pp. 50910–50926, 2018.

[8] L. Desuo, M. Bessani, R. Fanucchi, T. Gross, and C. D. Maciel, "A multi-objective swarm intelligence approach for field crews patrol optimization in power distribution systems restoration," *IEEE Latin America Transactions*, vol. 17, no. 2, pp. 338–346, 2019.

[9] S. He, K. Xie, W. Chen, D. Zhang, and J. Wen, "Energy-aware routing for SWIPT in multi-hop energy-constrained wireless network," *IEEE Access*, vol. 6, no. 99, pp. 17996–18008, 2018.

[10] N. Saha, S. Bera, and S. Misra, "Sway: traffic-aware QoS routing in software-defined IoT," *IEEE Transactions on Emerging Topics in Computing*, vol. 9, no. 1, pp. 390–401, 2021.

[11] X. Li, D. Li, J. Wan, C. Liu, and M. Imran, "Adaptive transmission optimization in SDN-based industrial internet of things with edge computing," *IEEE Internet of Things Journal*, vol. 5, no. 3, pp. 1351–1360, 2018.

[12] M. Tariq, M. Ali, F. Naeem, and H. V. Poor, "Vulnerability assessment of 6G-enabled smart grid cyber–physical systems," *IEEE Internet of Things Journal*, vol. 8, no. 7, pp. 5468–5475, 2021.

[13] Y. Sun, S. Zhou, and J. Xu, "EMM: energy-aware mobility management for mobile edge computing in ultra dense networks," *IEEE Journal on Selected Areas in Communications*, vol. 35, no. 11, pp. 2637–2646, 2017.

[14] S. Maghsudi and S. Stanczak, "Joint channel selection and power control in infrastructureless wireless networks: a multi-player multi-armed bandit framework," *IEEE Transactions on Vehicular Technology*, vol. 64, no. 10, pp. 4565–4578, 2014.

[15] Y. Zhao, J. Lee, and W. Chen, "Q-greedy UCB: a new exploration policy to learn resource-efficient scheduling," *China Communications*, vol. 18, no. 6, pp. 12–23, 2021.

[16] J. Bae, J. Lee, and S. Chon, "Learning to schedule network resources throughput and delay optimally using Q+-learning," *IEEE/ACM Transactions on Networking*, vol. 29, no. 2, pp. 750–763, 2021.

[17] A. Singh, G. Aujla, S. Garg, S. Kaddoum, and G. Singh, "Deep-learning-based SDN model for internet of things: an incremental tensor train approach," *IEEE Internet of Things Journal*, vol. 6, no. 7, pp. 6302–6311, 2020.

[18] F. Naeem, M. Tariq, and H. V. Poor, "SDN-enabled energy-efficient routing optimization framework for industrial

internet of things," *IEEE Transactions on Industrial Informatics*, vol. 17, no. 8, pp. 5660–5667, 2021.

[19] X. Lai, X. Ji, X. Zhou, and L. Chen, "Energy efficient link-delay aware routing in wireless sensor networks," *IEEE Sensors Journal*, vol. 18, no. 2, pp. 837–848, 2018.

[20] Z. Zhou, Y. Guo, Y. He, X. Zhao, and W. M. Bazzi, "Access control and resource allocation for M2M communications in industrial automation," *IEEE Transactions on Industrial Informatics*, vol. 15, no. 5, pp. 3093–3103, 2019.

[21] Y. Li and Y. Liang, "Compressed sensing in multi-hop large-scale wireless sensor networks based on routing topology tomography," *IEEE Access*, vol. 6, no. 99, pp. 27637–27650, 2018.

[22] S. Shi, S. Pang, Y. Li, F. Wang, H. Gacanin, and D. Zhang, "Buffer-aided relaying network with hybrid BNC for the internet of things: protocol and performance analysis," *IEEE Access*, vol. 8, no. 99, pp. 19646–19656, 2020.

[23] Y. Wang, R. Zhao, Y. Huang et al., "Asymptotic performance analysis of massive MIMO relay systems with multi-pair devices over correlated fading channels," *IEEE Access*, vol. 7, no. 99, pp. 27565–27578, 2019.

[24] F. Naeem, S. Seifollahi, Z. Zhou, and M. Tariq, "A generative adversarial network enabled deep distributional reinforcement learning for transmission scheduling in internet of vehicle," *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 7, pp. 4550–4559, 2021.

[25] R. Sutton and A. Barto, *Reinforcement Learning: An Introduction*, MIT Press, Cambridge, MA, USA, 2018.

[26] H. Yu, Z. Zhou, Z. Jia, X. Zhao, L. Zhang, and X. Wang, "Multi-timescale multi-dimension resource allocation for NOMA-edge computing-based power IoT with massive connectivity," *IEEE Transactions on Green Communications and Networking*, vol. 5, no. 3, pp. 1101–1113, 2021.

[27] Z. Wang, Z. Jia, and H. Liao, "Energy-aware and URLLC-aware task offloading for internet of health things," in *Proceedings of the GLOBECOM 2020-2020 IEEE Global Communications Conference*, pp. 1–6, Taipei, Taiwan, February 2021.

[28] Z. Jia, H. Liao, and Z. Zhou, "Multi-dimension resource allocation for NOMA-edge computing-based 6G power IoT," in *Proceedings of the 2021 IEEE International Conference on Communications Workshops (ICC Workshops)*, pp. 1–6, Montreal, QC, Canada, June 2021.

[29] X. He, H. Jiang, Y. Song, C. He, and H. Xiao, "Routing selection with reinforcement learning for energy harvesting multi-hop CRN," *IEEE Access*, vol. 7, no. 99, pp. 54435–54448, 2019.

[30] H. Liao, Z. Zhou, W. Kong et al., "Learning-based intent-aware task offloading for air-ground integrated vehicular edge computing," *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 8, pp. 5127–5139, 2021.