Research Article

Matching Learning-Based Relay Selection for Substation Power Internet of Things

Wei Wang ^{1,2}, Ruiqiuyu Wang ^{1,2}, Hao Zhang ^{1,2}, Zhenyu Zhou ^{1,2} and Yanhua He ³

¹State Key Laboratory of Alternate Electrical Power System with Renewable Energy Sources
(North China Electric Power University), Beijing, China

²National Mobile Communications Research Laboratory, Southeast University, Nanjing, Jiangsu, China

³State Grid Zhejiang Electric Power Company Ningbo Bureau, Hangzhou, Zhejiang, China

Correspondence should be addressed to Hao Zhang; haozhang@ncepu.edu.cn

Received 30 November 2021; Revised 12 January 2022; Accepted 27 January 2022; Published 21 February 2022

Academic Editor: Xingwang Li

Copyright © 2022 Wei Wang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Wireless sensor network (WSN) can effectively solve the problems of weak coverage, blind coverage, and low survivability of smart substation communication networks by deploying multiple relay nodes and adopting multihop transmission. However, there are still some challenges in the traditional relay selection strategy of WSN in substation, including incomplete information and the selection conflicts among multisource nodes. In this paper, we propose a matching learning-based relay selection mechanism for WSN-based substation power Internet of things (SPIoT). Firstly, considering the electromagnetic interference caused by the operation of high-voltage equipment, a multihop transmission model of SPIoT is built. Furthermore, based on the upper confidence bound (UCB) algorithm and matching theory, a matching learning-based relay selection (MLRS) algorithm is proposed to minimize the energy consumption of SPIoT devices. Simulation results demonstrate that MLRS outperforms existing algorithms in terms of energy consumption and optimal selection probability.

1. Introduction

Substation has great significance to ensure long-distance power transmission and stable power supply [1]. To achieve 24-hour uninterrupted monitoring of the substation, a large number of substation Internet of things (SPIoT) devices are deployed to collect various types of information including temperature, humidity, and smoke [2, 3]. However, traditional fiber-optical communication cannot meet the on-demand coverage and data transmission requirements of SPIoT due to poor scalability and high construction cost [4, 5]. Wireless sensor network (WSN) has the advantages of flexible networking and low deployment cost. WSN adopts multihop transmission to realize coverage enhancement and increase fault tolerance for on-demand monitoring of substations [6, 7]. It is complementary to fiber-optical communication and acts as an effective enabler for SPIoT.

In multihop transmission of WSN-enabled SPIoT, the source node selects relays for data transmission to shorten the transmission distance and enhance coverage under strong electromagnetic interference [8, 9]. To fully utilize the spectrum and energy resources, relay selection needs to be optimized dynamically according to the network state and service requirements. However, relay selection optimization in SPIoT still faces several critical challenges as below [10].

First, the relay selection strategies are coupled across different devices when multiple devices compete for the same relay. Each device faces an adversarial relay selection problem in which its strategy is affected by the strategies of other devices. Second, the global state information (GSI), including channel state information (CSI) and electromagnetic interference [11], is uncertain due to the limitations of network sensing capability and signaling overhead. The devices are forced to optimize relay selection under incomplete state information. Finally, SPIoT devices impose strict requirements on energy consumption due to limited

battery capacity. It is necessary to take the optimization of energy consumption into consideration, thereby improving the sustainability of SPIoT networks.

There exist some works that have addressed relay selection problems in IoT. In [12], Muller and Speidel proposed optimal relay selection strategies aiming at either maximizing the mean mutual information or minimizing the outage probability. In [13], Mousavi et al. proposed a relay subset selection method for two-hop WSN. However, these works ignore the optimization of energy consumption, and it is infeasible for SPIoT devices with limited battery. In [14], an analysis model for relay selection under the constraint of energy consumption was developed. In [15], Bakhsh et al. proposed a low-energy distributed relay selection algorithm to achieve reliable communication. However, the above studies have neglected the decision coupling among multiple devices.

Matching theory provides a method to solve combination problem and is widely used in relay selection optimization. In [16, 17], Wang et al. and Baidas et al. proposed a relay selection method based on matching theory, but the establishment of matching preference list requires complete GSI. In substations with dynamic network environment and complex electromagnetic interference, the preference list cannot be constructed without complete GSI, thereby making the traditional matching theory-based relay selection approaches unsuitable.

Reinforcement learning (RL) provides a powerful tool to deal with multistage decision-making problems under incomplete information [18, 19]. In [20], Su et al. proposed a deep RL-based relay selection scheme to achieve lower outage probability and higher utility. In [21], Liang et al. modeled the resource allocation problem in vehicular networks as a semi-Markov decision process and utilized DL algorithms to solve the problem. However, when dealing with problems with high-dimensionality spaces, RL invokes the dimensionality curse and has inferior performances in optimality and convergence [22, 23]. The relay selection strategies obtained by RL are unstable, and the influence of electromagnetic interference is ignored.

Motivated by the aforementioned challenges, we propose a matching learning-based relay selection (MLRS) algorithm to minimize the energy consumption of SPIoT devices. First, considering the electromagnetic interference of substations, a two-hop transmission model for SPIoT with electromagnetic interference is established. Second, we transform the relay selection problem into a one-to-one matching problem between multisource nodes and multirelay nodes. Based on the upper confidence bound (UCB) algorithm, the SPIoT gateway learns to build the preference lists of source nodes based on the number of selections and empirical performances. Third, the matching conflicts between multisource nodes and multirelay nodes are resolved based on matching with rising price. Finally, we compare MLRS with existing relay selection algorithms to validate its performance. The major contributions are presented as follows:

- (i) **Learning-based matching preference construction without precise GSI:** MLRS utilizes UCB to

learn to construct preference lists based on historical observations of relay node selection times and empirical energy consumption performances. MLRS enables the implementation of iterative matching without precise GSI and achieves an effective compromise between exploration and exploitation.

- (ii) **Stable matching between source and relay nodes:** MLRS utilizes matching theory to avoid selection conflicts between multiple sources and relay nodes, which achieve a stable matching based on the learned preference lists. In MLRS, the conflicts among multiple nodes are resolved by iteratively raising matching prices, and each source node is allocated with the most suitable relay node that ranks highest in its updated preference lists.
- (iii) **Low energy consumption:** MLRS can dynamically learn the relay selection preference, i.e., the historical energy consumption, thus effectively reducing transmission energy consumption based on the optimal relay selection strategy. Extensive simulations are carried out to demonstrate the low energy consumption performances of MLRS compared with existing relay selection algorithms.

The reminder of the work is organized as follows. Section 2 introduces the SPIoT model. Section 3 presents MLRS. Section 4 presents simulation results. Section 5 concludes this article.

2. System Model

A relay selection model of SPIoT network considering electromagnetic interference in complex substation environment is shown in Figure 1. The SPIoT network consists of two parts, i.e., SPIoT devices and gateway. The gateway provides decision-making and computing services for SPIoT devices. The SPIoT network is a many-to-one two-hop transmission network that includes I source nodes, J relay nodes, and one destination node. Each source node collects various types of information including temperature, humidity, smoke, and gas and transmits the collected data to a relay node. The relay node receives the collected data from the source node and forwards the data to the destination node. Specifically, decode-and-forward relay node is adopted to avoid the amplification of noise power. The destination node is the substation interval measurement and control cabinet, which serves as the receiving end of monitoring data. Denote the sets of source nodes and relay nodes as $\mathcal{S} = \{s_1, \dots, s_j, \dots, s_I\}$ and $\mathcal{R} = \{r_1, \dots, r_j, \dots, r_J\}$, respectively. Denote the destination node as u_0 . The total optimization period is divided to T time slots, the set of which is indexed as $\mathcal{T} = \{1, \dots, t, \dots, T\}$. At the beginning of the t th slot, the source node s_i has to transmit $A_i(t)$ amount of data to the destination node. The t th slot ends when all the data have arrived at the destination node. Considering the impact of the dynamic CSI and electromagnetic interference on the transmission delay, the slot duration is unequal. In each slot,

FIGURE 1: Two-hop transmission model for SPIoT with electromagnetic interference.

the gateway learns the optimal relay selection strategy ϕ for each source node based on the historical performance and sends the relay selection strategy to source nodes. The transmission delay and energy consumption models are introduced as follows.

2.1. Transmission Delay Model. In the t th slot, the source node s_i selects a relay node, e.g., r_j , to transmit $A_i(t)$ amount of data to the destination node u_0 based on the relay selection strategy from the gateway. Let $a_{i,j,t} = 1$ represent that s_i selects r_j in the t th slot; otherwise, $a_{i,j,t} = 0$. The transmission process is divided into two hops. The first hop is from the source node to the relay node, and the second hop is from the relay node to the destination node.

The signal-to-noise ratio of the first-hop transmission from s_i to r_j is given by [24, 25]

$$\gamma_{i,j}(t) = \frac{P_{TX,i} G_{i,j}(t)}{\sigma_0 + \varepsilon_{i,j}(t)}, \quad (1)$$

where $P_{TX,i}$ represents the transmission power of s_i , $G_{i,j}(t)$ represents the channel gain between s_i and r_j , σ_0 represents the noise power, and $\varepsilon_{i,j}(t)$ represents the electromagnetic interference power during the first-hop transmission. According to Shannon's formula [26], the first-hop transmission rate is given by

$$R_{i,j}(t) = B \log(1 + \gamma_{i,j}(t)) = B \log\left(1 + \frac{P_{TX,i} G_{i,j}(t)}{\sigma_0 + \varepsilon_{i,j}(t)}\right), \quad (2)$$

where B represents the bandwidth. Therefore, the first-hop transmission delay from s_i to r_j is given by

$$\tau_{i,j,t} = \frac{A_i(t)}{R_{i,j}(t)} = \frac{A_i(t)}{B \log(1 + ((P_{TX,i} G_{i,j}(t)) / (\sigma_0 + \varepsilon_{i,j}(t))))}. \quad (3)$$

In the t th slot, the signal-to-noise ratio of the second-hop transmission from r_j to the destination node u_0 is given by

$$\gamma_{j,0}(t) = \frac{P_{TX,j} G_{j,0}(t)}{\sigma_0 + \varepsilon_{j,0}(t)}, \quad (4)$$

where $P_{TX,j}$ represents the transmission power of r_j , $G_{j,0}(t)$ represents the channel gain between r_j and u_0 , and $\varepsilon_{j,0}(t)$ represents the electromagnetic interference power during the second-hop transmission. The transmission rate from r_j to u_0 in the second hop is given by

$$R_{j,0}(t) = B \log(1 + \gamma_{j,0}(t)) = B \log\left(1 + \frac{P_{TX,j} G_{j,0}(t)}{\sigma_0 + \varepsilon_{j,0}(t)}\right). \quad (5)$$

Therefore, in the t th slot, the second-hop transmission delay from r_j to u_0 is given by

$$\tau_{j,0,t} = \frac{A_i(t)}{R_{j,0}(t)} = \frac{A_i(t)}{B \log(1 + ((P_{TX,j} G_{j,0}(t)) / (\sigma_0 + \varepsilon_{j,0}(t))))}. \quad (6)$$

Therefore, the total transmission delay from s_i to u_0 through r_j in the t th slot, i.e., the length of slot t , is the sum of the first-hop transmission delay and the second-hop transmission delay, which is given by

$$\tau_t = \tau_{i,j,t} + \tau_{j,0,t} = A_i(t) \left(\frac{1}{B \log(1 + ((P_{TX,i} G_{i,j}(t))/(\sigma_0 + \varepsilon_{i,j}(t))))} + \frac{1}{B \log(1 + ((P_{TX,j} G_{j,0}(t))/(\sigma_0 + \varepsilon_{j,0}(t))))} \right). \quad (7)$$

2.2. Energy Consumption Model. The energy consumption of the two-hop data transmission includes transmission energy consumption and reception energy consumption. The transmission energy consumption includes data transmission energy consumption and transmitting circuit energy consumption. The reception energy consumption is receiving circuit energy consumption. The first-hop data transmission energy consumption from s_i to r_j is given by [27]

$$E_{i,j,t}^1 = P_{TX,i} \tau_{i,j,t}. \quad (8)$$

The second-hop data transmission energy consumption from r_j to u_0 is given by

$$E_{i,j,t}^2 = P_{TX,j} \tau_{j,0,t}. \quad (9)$$

The transmission circuit energy consumption of s_i is given by

$$E_{i,t}^{tra} = A_i(t) E_{tra}, \quad (10)$$

where E_{tra} is energy consumption coefficient of transmission circuit.

The transmission circuit and reception circuit energy consumption of r_j is given by

$$E_{i,t}^{trrec} = A_i(t) E_{trrec}. \quad (11)$$

where E_{trrec} is energy consumption coefficient of transmission circuit and reception circuit.

The reception circuit energy consumption of u_0 is given by

$$E_{i,t}^{rec} = A_i(t) E_{rec}. \quad (12)$$

where E_{rec} is energy consumption coefficient of reception circuit.

The total energy consumption is the sum of transmission energy consumption and reception energy consumption, which is given by

$$E_{i,j,t} = E_{i,t}^{tra} + E_{i,j,t}^1 + E_{i,t}^{trrec} + E_{i,j,t}^2 + E_{i,t}^{rec}, \quad (13)$$

where the first and second terms represent the energy consumption of the first-hop transmission from s_i to r_j , the third term represents the forwarding energy consumption of r_j , and the fourth and fifth terms represent the energy consumption of the second-hop transmission from r_j to u_0 .

2.3. Problem Formulation. We address the relay selection problem for SPIoT under dynamic CSI and electromagnetic interference. The relay selection problem is formulated as

$$\begin{aligned} \text{P1: } & \min_{\{a_{i,j,t}\}} \sum_{t=1}^T \sum_{i=1}^I \sum_{j=1}^J a_{i,j,t} E_{i,j,t}, \\ \text{s.t. } & C_1: a_{i,j,t} \in \{0, 1\}, \quad \forall s_i \in \mathcal{S}, \forall r_j \in \mathcal{R}, \forall t \in \mathcal{T}, \\ & C_2: \sum_{r_j \in \mathcal{R}} a_{i,j,t} \leq 1, \quad \forall s_i \in \mathcal{S}, \forall t \in \mathcal{T}, \\ & C_3: \sum_{s_i \in \mathcal{S}} a_{i,j,t} \leq 1, \quad \forall r_j \in \mathcal{R}, \forall t \in \mathcal{T}, \end{aligned} \quad (14)$$

where C_1 defines the value of the selected indicator variable and C_2 and C_3 ensure that there is no selection conflict in each time slot. In other words, a source node selects at most one relay node for data forwarding in each time slot, and a relay node is selected by one source node at most.

3. Matching Learning-Based Relay Selection for SPIoT

In this section, the implementation process of the proposed MLRS algorithm for two-hop data transmission in SPIoT is elaborated.

3.1. Problem Transformation. P1 is intractable because the relay selection strategies of all source nodes are coupled. Based on matching theory [28, 29], the selection problem between multisource node and multirelay node can be transformed into a one-to-one matching problem between I source nodes and J relay nodes. To solve the problem of incomplete global information in preference lists' construction of matching theory, MLRS utilizes UCB to enable the implementation of iterative matching without precise GSI and utilizes matching theory to avoid selection conflicts, which achieves a stable matching based on the learned preference lists. Since the relay node is unconditionally selected by the source node for data forwarding, the unilateral matching theory is adopted [30]. The definition of matching is given below.

Definition 1. (matching). A matching ϕ is a one-to-one correspondence of the set $\mathcal{S} \cup \mathcal{R}$ to itself. $\phi(s_i) = r_j$ means $a_{i,j,t} = 1$, i.e., s_i and r_j are matched. Otherwise, $\phi(s_i) = s_i$ means s_i does not match with any r_j .

However, the traditional matching theory needs to construct a preference list for each source node based on global information, which is not suitable to real-world SPIoT scenario, due to the dynamic changes of electromagnetic interference and the uncertain global information. Reinforcement learning can solve decision-making problems with incomplete information through continuous interaction with the environment, thereby enabling a high-precision and low-complexity preference list construction. It has the advantages of fast convergence and strong adaptability.

3.2. The Proposed MLRS Algorithm. Based on the multi-armed bandit (MAB) framework in reinforcement learning [31, 32], the source nodes and the relay nodes are modeled as players and arms, respectively. In the t th slot, s_i selects r_j to transmit data, and the performance of r_j can only be observed afterwards. The UCB algorithm is an effective method to solve the MAB problem. It is employed by MLRS to construct the preference list. MLRS is mainly divided into two steps: UCB-based preference list construction and iterative matching with rising price.

3.2.1. UCB-Based Preference List Construction. Denote the historical energy consumption of $\phi(s_i) = r_j$ up to slot t as $\bar{E}_{i,j,t-1}$, i.e., the empirical estimation of the two-hop transmission energy consumption $E_{i,j,t-1}$. Denote the total number of times that s_i selects r_j up to slot t as $k_{i,j,t-1}$, which is given by

$$k_{i,j,t-1} = \sum_{m=1}^{m=t-1} a_{i,j,m}. \quad (15)$$

Based on the UCB algorithm and the optimization objective P1, the matching preference value of $s_i \in \mathcal{S}$ towards $r_j \in \mathcal{R}$ is constructed to minimize the total energy consumption of the two-hop transmission, which is given by

$$\tilde{Q}_{i,j,t} = \frac{1}{\bar{E}_{i,j,t-1}} + \sqrt{\frac{\omega \ln t}{k_{i,j,t-1}}} - H_j, \quad (16)$$

where the first item represents the empirical performance of r_j up to slot t . The matching preference value of s_i towards r_j decreases as the historical energy consumption increases. The second item is the confidence bound, which represents the estimation uncertainty. It decreases as $k_{i,j,t-1}$ increases, indicating that the estimated performance gradually approaches the actual expected value. $\omega > 0$ represents the weight for exploration. H_j represents the matching cost of r_j . The initial value of H_j is set as zero.

Denote the preference list of source nodes $s_i \in \mathcal{S}$ toward all the relay nodes $r_j \in \mathcal{R}$ as $\mathcal{F}_i$. Sort all $\tilde{Q}_{i,j,t}, \forall r_j \in \mathcal{R}$, in the descending order to construct the preference list $\mathcal{F}_i$ of the source nodes.

3.2.2. Iterative Matching with Rising Price. The SPIoT gateway makes matching decisions based on the following steps. The implementation procedure is provided in Algorithm 1, which includes three steps.

Step 1: initialization.

Set $\phi(s_i) = \emptyset, \forall s_i \in \mathcal{S}, \Omega = \emptyset, H_j = 0$, and $\forall r_j \in \mathcal{R}$, where Ω is denoted as the set of relay nodes that have received more than one matching request.

Step 2: iterative matching.

Repeat

$\exists \phi(s_i) = \emptyset$, each $s_i \in \mathcal{S}$ makes a request to its most preferred r_j according to $\mathcal{F}_i$.

For any $r_j \in \mathcal{R}$, it is matched with the source node which sends the initial matching request. If r_j has been requested by more than one source node, add r_j into Ω . If $\Omega \neq \emptyset$ • Each $r_j \in \Omega$ raises its price as

$$H_j = H_j + \Delta H_j, \quad (17)$$

where ΔH_j represents the rising step.

Each s_i update $\mathcal{F}_i$ based on H_j and renew its request. Repeat price rising process until r_j is not requested by more than one source node. Remove r_j from Ω .

For $s_i \in \mathcal{S}$ that $\phi(s_i) = s_i$, the data transmission is suspended and waits for the next iteration.

Until $\phi(s_i) \neq \emptyset$.

Finally, the gateway sends the result ϕ to source nodes. Each source node $s_i \in \mathcal{S}$ selects $\phi(s_i)$ and transmits data.

Step 3: learning.

Observe the transmission performance and get the energy consumption $E_{i,j,t}$. Update $\bar{E}_{i,j,t}$ and $k_{i,j,t}$ as

$$\bar{E}_{i,j,t} = \frac{\bar{E}_{i,j,t-1}k_{i,j,t-1} + E_{i,j,t}a_{i,j,t}}{k_{i,j,t-1} + a_{i,j,t}}, \quad (18)$$

$$k_{i,j,t} = k_{i,j,t-1} + a_{i,j,t}. \quad (19)$$

3.3. Performance Analysis

Theorem 1 (stability). *For any $s_i \in \mathcal{S}$, there is no situation, where s_i prefers r_j than $\phi(s_i)$ and ϕ is stable.*

The detailed proof is in [33, 34].

Theorem 2 (convergence). *Due to the stability derived in Theorem 1, MLRS is convergent.*

Proof. Based on proof by contradiction, we assume MLRS is not convergent. Hence, there exist s_i preferring r'_j when $\phi(s_i) = r_j$ and $\phi(s'_i) = r'_j$. Due to the higher preference value of s_i towards r'_j , r'_j should be requested by s_i before r_j is requested by s_i . However, r'_j is matched with s'_i , i.e., r'_j refused s_i , which is in contradiction with the assumption. Therefore, MLRS is convergent.

Complexity of MLRS: the computational complexity of MLRS consists of three parts. In the first step, the computational complexity is $\mathcal{O}(J)$. In the second step, we assume that MLRS takes ζ iterations to resolve matching conflicts. When the number of source nodes is greater than or equal to the number of relay nodes, i.e., $I \geq J$, the computational complexity is $\mathcal{O}(J) + \mathcal{O}(J \log J) + \mathcal{O}(I\zeta)$, while the computational complexity of the enumeration method is $\mathcal{O}(J^I I!)$. In the third step, the computational complexity is $\mathcal{O}(J)$. Therefore, the overall computational complexity of MLRS is $3\mathcal{O}(J) + \mathcal{O}(J \log J) + \mathcal{O}(I\zeta)$. $\square$

```

(1) Input:  $\omega, T, s, \mathcal{R}$ .
(2) Output:  $\phi$ .
(3) In the first  $I \times J$  time slots,  $\forall s_i \in \mathcal{S}$  traverse to select  $\forall r_j \in \mathcal{R}$  once to observe the transmission performance.
(4) while  $t \leq T$  do
(5)   All source nodes  $s_i \in \mathcal{S}$  construct a preference list  $\mathcal{F}_i$  as (16).
(6)   Step 1: Initialization
(7)   Set  $\phi(s_i) = \emptyset, \forall s_i \in \mathcal{S}, \Omega = \emptyset, H_j = 0, \forall r_j \in \mathcal{R}$ .
(8)   Step 2: Iterative Matching
(9)   while  $\exists \phi(s_i) = \emptyset$  do
(10)    for  $s_i \in \mathcal{S}$  do
(11)      $\forall s_i \in \mathcal{S}$  makes a request to its most preferred  $r_j$  according to  $\mathcal{F}_i$ .
(12)    end for
(13)    Add the relay nodes selected by more than one source node into  $\Omega$ 
(14)    if  $\Omega \neq \emptyset$  then
(15)     for  $r_j \in \Omega$  do
(16)       $r_j$  raises its price as (17).
(17)       $\forall s_i$  selecting  $r_j$ , update  $\mathcal{F}_i$  and renew its request.
(18)      Repeat price rising process until  $r_j$  is not requested by more than one source node. Remove  $r_j$  from  $\Omega$ .
(19)     end for
(20)    end if
(21)  end while
(22)  The gateway sends the result  $\phi$  to source nodes.  $\forall s_i \in \mathcal{S}$  selects  $\phi(s_i)$  and transmits data.
(23)  Step 3: Learning
(24)  Observe transmission performance and get energy consumption  $E_{i,j,t}$ .
(25)  Update  $\bar{E}_{i,j,t}$  and  $k_{i,j,t}$  as (18) and (19).
(26) end while

```

ALGORITHM 1: MLRS.

4. Simulation Results

In this section, we validate the performance of MLRS by simulations. We consider a scenario with 10 source nodes, 15 relay nodes, and 1 destination node. The number of time slots is 500. The transmission power $P_{TX,i}$ and $P_{TX,j}$ are set as $[0.1, 1]$ W. The channel gains from the source node to the relay node and from the relay node to the destination node all satisfy the normal distribution $\mathcal{N}(127 + 30 \log(l), 1)$ [35], where l is the distance. The simulation parameters are summarized in Table 1 [36–38]. We consider two state-of-the-art algorithms for comparison. The first one is the traditional UCB algorithm [39], where each source node calculates its preference value for each relay node based on historical performance and sends a data transmission request to the favorite relay node. Considering the case of matching conflicts, the source nodes with conflicts' preferences towards a relay node will be randomly matched with the remaining unmatched relay nodes. The second one is the energy efficiency performance selection (EEPS) algorithm [40], where each source node sends a transmission request to the relay node with the best historical performance, and the rest process is the same as the traditional UCB algorithm.

The α -stable distribution is employed to describe the electromagnetic interference. The characteristic function of the electromagnetic interference power from s_i to r_j is given by

$$\log \varepsilon_{i,j}(t) = \begin{cases} -\sigma^\alpha |t|^\alpha \left\{ 1 - j\beta \text{sign}(t) \tan\left(\frac{\pi\alpha}{2}\right) \right\} + j\mu t, & \alpha \neq 1, \\ -\sigma |t| \left\{ 1 + j\beta \text{sign}(t) \frac{2}{\pi} \log|t| \right\} + j\mu t, & \alpha = 1, \end{cases} \quad (20)$$

where $\alpha \in (0, 2]$ is the characteristic exponent, $\beta \in [-1, 1]$ is the skew parameter, $\beta = 0$ indicates a symmetric α -stable distribution, $\mu \in R$ is the location parameter, and σ is the scale parameter [41].

Figure 2 shows the average energy consumption versus time slot. When $t = 500$, compared with UCB and EEPS, the average energy consumption of MLRS is reduced by 17.49% and 24.22%, respectively. The reason is that the random selection mechanism in UCB and EEPS leads to unmatched source nodes and suspended data transmission when multiple source nodes compete for the same relay node. Moreover, UCB and EEPS randomly assign unmatched source nodes to relay nodes. As a result, these source nodes may be matched with some relay nodes with poor historical performance, leading to higher energy consumption. MLRS can effectively resolve matching conflicts between multi-source nodes and multirelay nodes based on dynamically learned relay selection preference.

Figure 3 shows the optimal relay node selection probability versus time slot. In the initial stage for all the algorithms, the optimal relay node selection probability is low. The probability increases gradually with the number of

TABLE 1: Simulation parameters.

Parameters	Value
The number of source nodes	10
The number of relay nodes	15
The number of time slots	500
The transmission power, $P_{TX,i}$ and $P_{TX,j}$	$[0.1, 1]$ W
The channel gains from s_i to r_j and from r_j to u_0	$\mathcal{N}(127 + 30 \log(l), 1)$
The amount of task data at the source node, $A_i(t)$	$[30, 40]$ kbit
Channel bandwidth, B	0.1 MHz
The noise power, n_0	-114 dBm
The energy consumption coefficient, E_{tra} and E_{rec}	50 nJ/bit
The energy consumption coefficient, E_{trec}	100 nJ/bit
The weight of exploration, ω	0.3
The characteristic exponent, α	1.4
The skew parameter, β	-0.791
The location parameter, μ	$-4 * 10^{-14}$
The scale parameter, σ	10^{-15}

FIGURE 2: The average energy consumption versus time slot.

selections and finally converges as to 68%, 54%, and 34% for MLRS, UCB, and EEPS, respectively. MLRS always outperforms UCB and EEPS. The reason is that MLRS adopts a tendency exploration scheme instead of random selection, which can achieve an effective compromise between exploration and exploitation.

Figure 4 shows the optimal relay node selection probability versus the number of relay nodes. The optimal relay node selection probability of MLRS converges around 67% and outperforms UCB and EEPS by 18.08% and 18.20%, respectively. When the number of relay nodes increases, i.e., the network topology becomes more complex, MLRS can better adapt to complex network topology with more relay nodes and learn the most appropriate relay selection strategy by combining with learning-based matching preference lists' construction and matching-based conflict resolution.

FIGURE 3: The optimal relay node selection probability versus time slot.

FIGURE 4: The optimal relay node selection probability versus the number of relay nodes.

However, with the increase of the number of relay nodes, the problem of matching conflict becomes prominent in UCB and EEPS, which leads to sharp performance decrease.

Figure 5 shows the average energy consumption versus ω . As ω increases, the performance of MLRS improves. When $\omega = 0.3$, the algorithm performance is optimal. A too small ω leads to biased preference towards exploitation and inability to explore potential better options, while a too large ω leads to inability to exploit existing optimal option.

Figure 6 shows the average energy consumption versus the electromagnetic interference intensity. We divide the electromagnetic interference into five levels by changing the

FIGURE 5: The average energy consumption versus the exploration weight ω in MLRS.

FIGURE 6: The average energy consumption versus the electromagnetic interference intensity.

TABLE 2: Electromagnetic interference levels.

Level	I	II	III	IV	V
μ	$4 * 10^{-16} \text{ W}$	$4 * 10^{-15} \text{ W}$	$4 * 10^{-14} \text{ W}$	$4 * 10^{-13} \text{ W}$	$4 * 10^{-12} \text{ W}$

location parameter μ of the α -stable distribution, which are summarized in Table 2. As the electromagnetic interference intensity level increases, compared with UCB and EEPS, MLRS always has the lowest average energy consumption. The reason is that MLRS can learn the optimal relay selection strategy and resolve the matching conflict regardless of the electromagnetic interference level, which verifies that MLRS shows adaptability to various wireless environment.

5. Conclusion

In this paper, a novel relay selection algorithm named MLRS was proposed for SPIoT. MLRS can minimize the energy consumption of SPIoT devices under complex electromagnetic interference through relay selection optimization without GSI. Simulation results indicate that, compared with UCB and EEPS, MLRS reduces the energy consumption by

17.49% and 24.22%, respectively. In the future, we will focus on the joint optimization of multiple quality of service (QoS) performance metrics including delay and throughput, considering the differentiated QoS requirements of SPIoT.

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 that they have no conflicts of interest.

Acknowledgments

This work was partially supported by the State Key Laboratory of Alternate Electrical Power System with Renewable Energy Sources, under Grant LAPS202125, and supported by the Open Research Fund of National Mobile Communications Research Laboratory, Southeast University, under Grant no. 2021D12.

References

- [1] L. Zhao, I. Matsuo, Y. Zhou, and W. J. Lee, "Design of an industrial IoT-based monitoring system for power substations," *IEEE Transactions on Industry Applications*, vol. 55, no. 4, pp. 5666–5674, 2019.
- [2] 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.
- [3] 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.
- [4] A. Amari, O. A. Dobre, R. Venkatesan, O. S. S. Kumar, P. Ciblat, and Y. Jaouën, "A survey on fiber nonlinearity compensation for 400 gb/s and beyond optical communication systems," *IEEE Communications Surveys & Tutorials*, vol. 19, no. 4, pp. 3097–3113, 2017.
- [5] L. Zhao, X. Li, and B. Gu, "Vehicular communications: standardization and open issues," *IEEE Commun. Standards Mag.*, vol. 2, no. 4, pp. 74–80, 2020.
- [6] J. Cheng, P. Yang, K. Navaie, Q. Ni, and H. Yang, "A low-latency interference coordinated routing for wireless multi-hop networks," *IEEE Sensors Journal*, vol. 21, no. 6, pp. 8679–8690, 2021.
- [7] G. Sheng, H. Wang, F.-Q. Wen, and X. Wang, "Fast angle estimation and sensor self-calibration in bistatic MIMO radar with gain-phase errors and spatially colored noise," *IEEE Access*, vol. 8, pp. 123701–123710, 2020.
- [8] X. Li, M. Zhao, M. Zeng et al., "Hardware impaired ambient backscatter NOMA systems: reliability and security," *IEEE Transactions on Communications*, vol. 69, no. 4, pp. 2723–2736, 2021.
- [9] Z. Zhou, H. Liao, X. Zhao, B. Ai, and M. Guizani, "Reliable task offloading for vehicular fog computing under information asymmetry and information uncertainty," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 9, pp. 8322–8335, 2019.
- [10] C. Wang, X. Gaimu, C. Li, H. Zou, and W. Wang, "Smart mobile crowdsensing with urban vehicles: a deep reinforcement learning perspective," *IEEE Access*, vol. 7, pp. 37334–37341, 2019.
- [11] Z. Tang and T. Jian, "Research on influences factors of electromagnetic interferences for partial discharge detection in substations," in *Proceedings of the IEEE Electrical Insulation Conference (EIC)*, pp. 42–45, Baltimore, MD, USA, August 2017.
- [12] A. Muller and J. Speidel, "Relay selection in dual-hop transmission systems: selection strategies and performance results," in *Proceedings of the IEEE International Conference on Communication (ICC)*, pp. 4998–5003, Beijing, China, May 2008.
- [13] S. H. Mousavi, J. Haghighat, and W. Hamouda, "A relay subset selection scheme for wireless sensor networks based on channel state information," in *Proceedings of the IEEE International Conference on Communication (ICC)*, pp. 1–6, Kuala Lumpur, Malaysia, May 2016.
- [14] M. Adil, R. Khan, J. Ali, B.-H. Roh, Q. T. H. Ta, and M. A. Almaiah, "An energy proficient load balancing routing scheme for wireless sensor networks to maximize their life-span in an operational environment," *IEEE Access*, vol. 8, pp. 163209–163224, 2020.
- [15] S. T. Bakhsh, "Energy-efficient distributed relay selection in wireless sensor network for internet of things," in *Proceedings of the 13th International Wireless Communication on Mobile Computers Conference (IWCMC)*, pp. 1802–1807, Valencia, Spain, July 2017.
- [16] B. Wang, Y. Sun, H. M. Nguyen, and T. Q. Duong, "A novel socially stable matching model for secure relay selection in D2D communications," *IEEE Wireless Communications Letters*, vol. 9, no. 2, pp. 162–165, 2020.
- [17] M. W. Baidas, M. M. Afghah, and F. Afghah, "A matching-theoretic approach to distributed swipt in ad-hoc wireless networks," in *Proceedings of the International Symposium on Network, Computer Communication (ISNCC)*, pp. 1–6, Istanbul, Turkey, June 2019.
- [18] H. Liao, Z. Zhou, X. Zhao et al., "Learning-based context-aware resource allocation for edge-computing-empowered industrial IoT," *IEEE Internet of Things Journal*, vol. 7, no. 5, pp. 4260–4277, 2020.
- [19] L. Zhao, K. Yang, and Z. Tan, "A novel cost optimization strategy for SDN-enabled UAV-assisted vehicular computation offloading," *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 6, pp. 3664–3674, 2020.
- [20] Y. Su, X. Lu, Y. Zhao, L. Huang, and X. Du, "Cooperative communications with relay selection based on deep reinforcement learning in wireless sensor networks," *IEEE Sensors Journal*, vol. 19, no. 20, pp. 9561–9569, 2019.
- [21] H. Liang, X. Zhang, X. Hong et al., "Reinforcement learning enabled dynamic resource allocation in the internet of vehicles," *IEEE Transactions on Industrial Informatics*, vol. 17, no. 7, pp. 4957–4967, 2021.
- [22] Z. Zhou, J. Feng, B. Gu et al., "When mobile crowd sensing meets UAV: energy-efficient task assignment and route planning," *IEEE Transactions on Communications*, vol. 66, no. 11, pp. 5526–5538, 2018.
- [23] X. Li, Y. Zheng, M. D. Alshehri et al., "Cognitive AmBC-NOMA IoV-MTS networks with IQI: reliability and security analysis," in *Proceedings of the IEEE Transactions on Intelligent Transportation Systems*, pp. 1–12, September 2021.
- [24] J. Wang, B. Li, M. Liu, and J. Li, "SNR estimation of time-frequency overlapped signals for underlay cognitive radio,"

- IEEE Communications Letters*, vol. 19, no. 11, pp. 1925–1928, 2015.
- [25] M. A. Raza and A. Hussain, “Maximum likelihood SNR estimation of hyper cubic signals over Gaussian channel,” *IEEE Communications Letters*, vol. 20, no. 1, pp. 45–48, 2016.
  - [26] S.-H. Lee and S.-Y. Chung, “Capacity scaling of wireless ad hoc networks: Shannon meets Maxwell,” *IEEE Transactions on Information Theory*, vol. 58, no. 3, pp. 1702–1715, 2012.
  - [27] C. Wang, D. Deng, L. Xu, W. Wang, and F. Gao, “Joint interference alignment and power control for dense networks via deep reinforcement learning,” *IEEE Wireless Communications Letters*, vol. 10, no. 5, pp. 966–970, 2021.
  - [28] Z. Zhou, C. Zhang, C. Xu, F. Xiong, Y. Zhang, and T. Umer, “Energy-efficient industrial internet of UAVs for power line inspection in smart grid,” *IEEE Transactions on Industrial Informatics*, vol. 14, no. 6, pp. 2705–2714, 2018.
  - [29] S. Jeong, T.-H. Lin, and M. M. Tentzeris, “A real-time range-adaptive impedance matching utilizing a machine learning strategy based on neural networks for wireless power transfer systems,” *IEEE Transactions on Microwave Theory and Techniques*, vol. 67, no. 12, pp. 5340–5347, 2019.
  - [30] Y. Yuan, T. Yang, Y. Hu, H. Feng, and B. Hu, “Two-timescale resource allocation for cooperative D2D communication: a matching game approach,” *IEEE Transactions on Vehicular Technology*, vol. 70, no. 1, pp. 543–557, 2021.
  - [31] N. Papandreou and T. Antonakopoulos, “Resource allocation management for indoor power-line communications systems,” *IEEE Transactions on Power Delivery*, vol. 22, no. 2, pp. 893–903, 2007.
  - [32] L. Zhao, W. Zhao, A. Hawbani et al., “Novel online sequential learning-based adaptive routing for edge software-defined vehicular networks,” *IEEE Transactions on Wireless Communications*, vol. 20, no. 5, pp. 2991–3004, 2021.
  - [33] Z. Zhou, C. Gao, C. Xu, Y. Zhang, S. Mumtaz, and J. Rodriguez, “Social big-data-based content dissemination in internet of vehicles,” *IEEE Transactions on Industrial Informatics*, vol. 14, no. 2, pp. 768–777, 2018.
  - [34] F. Naeem, S. Seifollahi, Z. Zhou, and M. Tariq, “A generative adversarial network enabled deep distributional reinforcement learning for transmission scheduling in internet of vehicles,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 7, pp. 4550–4559, 2021.
  - [35] 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.
  - [36] 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.
  - [37] J. Zhang, J. Tang, and F. Wang, “Cooperative relay selection for load balancing with mobility in hierarchical WSNs: a multi-armed bandit approach,” *IEEE Access*, vol. 8, pp. 18110–18122, 2021.
  - [38] J. Luo, J. Hu, D. Wu, and R. Li, “Opportunistic routing algorithm for relay node selection in wireless sensor networks,” *IEEE Transactions on Industrial Informatics*, vol. 11, no. 1, pp. 112–121, 2015.
  - [39] M. Endo, T. Ohtsuki, T. Fujii, and O. Takyu, “Secure channel selection using multi-armed bandit algorithm in cognitive radio network,” in *Proceedings of the IEEE 85th Vehicle Technology Conference (VTC Spring)*, pp. 1–5, Sydney, NSW, Australia, November 2017.
  - [40] B. Klaiqi, X. Chu, and J. Zhang, “Energy-Efficient and low signaling overhead cooperative relaying with proactive relay subset selection,” *IEEE Transactions on Communications*, vol. 64, no. 3, pp. 1001–1015, 2016.
  - [41] Z. Zhou, X. Yang, and C. Xu, “Performance evaluation of multi-antenna based m2m communications for substation monitoring,” in *Proceedings of the International Conference on Information and Communication Technology Convergence (ICTC)*, pp. 97–102, Jeju, Korea (South), October 2016.