IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received March 18, 2020, accepted April 1, 2020, date of publication April 14, 2020, date of current version May 1, 2020.
Digital Object Identifier 10.1109/ACCESS.2020.2987861

Energy Efficient Two-Tier Data Dissemination Based on Q-Learning for Wireless Sensor Networks

NENG-CHUNG WANG, (Member, IEEE), AND WEI-JUNG HSU
Department of Computer Science and Information Engineering, National United University, Miaoli 36063, Taiwan
Corresponding author: Neng-Chung Wang (ncwang@nuu.edu.tw)

This work was supported in part by the Ministry of Science and Technology of China under Grant MOST-107-2221-E-239-031, and in part by the National United University under Grant NUU-108ch001 and Grant 109-NUUPRJ-03.

ABSTRACT Green communication for different kinds of wireless networks has begun to receive significant research attention recently. Green communication focuses mainly on the issue of improving energy efficiency substantially. A wireless sensor network (WSN) consists of a large number of randomly and widely deployed sensor nodes, and these nodes themselves have the ability to wireless communicate, detect and process data. Sensor nodes can thus detect their surrounding environment, and transmit related data to a sink via wireless communication. This study proposes two two-tier data dissemination schemes based on Q-learning for wireless sensor networks. In the proposed schemes, a source node uses Q-learning to find the most energy efficient data dissemination path from the source node to the sink. The first scheme is called TTDD-QL, and the second scheme is called TTDD-QL-A which is an advanced version of TTDD-QL. In TTDD-QL, the reward is determined by the distance between the current dissemination node and the sink. In each iteration, the proposed scheme will update the Q values. After multiple learning iterations, the Q values are converged, and the data dissemination path is found according to the Q values. In TTDD-QL-A, the reward is determined not only by the distance between the current dissemination node and the sink but also by the remaining energy of the current dissemination node. Simulation results show that TTDD-QL and TTDD-QL-A can reduce sensor node energy consumption and extend the lifetime of the WSN.

INDEX TERMS Data dissemination, grid, Q-learning, sink, wireless sensor network.

I. INTRODUCTION

The Internet of Things (IoT) can connect almost all smart devices with the Internet in order to realize intelligent identification and applications through sensing devices [1]. The IoT has many applications in any fields, such as transportation, medicine, logistics management or even smart homes. In addition, fifth generation mobile communication technology (5G) [2] is a new generation of cellular mobile communication technology. The performance goals of 5G include high transmission rates, low transmission latency, power savings, costs reduction, and system scalability. 5G thus offers high speed, low latency, and multi-device connection. Each of these features is a technical bottleneck that the IoT urgently needs to address. In any scenario involving a large number of connected devices, large-scale terminal-to-terminal communication affects network speed, stability, and transmission delay. Thus, routing and security solutions in 5G and IoT are very important challenges [3]–[5]. Due to the rapid progress of wireless network and mobile communication technologies, a rapidly growing number of applications related to the IoT are becoming available, and related infrastructure technologies, such as wireless sensor networks (WSNs), mobile ad hoc networks (MANETs), and software defined opportunistic networks (SDONs), have become increasingly important [6]–[8].

The WSN is one of the key technologies for the development of the IoT. In WSNs, each sensor has the ability to wirelessly communicate, detect and process data. In addition to the sensing function, other features of the sensor

The associate editor coordinating the review of this manuscript and approving it for publication was Ilsun You.

design include its small capacity, low power use, and low price [9]–[11]. Due to sensors' small size, the application of wireless network technologies is increasing in areas like environmental observation, health monitoring, and building monitoring [12]–[14]. In terms of environmental observation, WSNs can be used to detect changes in the environment, as in air pollution monitoring, ecological environment monitoring or forest fire detection. Health monitoring can be achieved by implanting sensors in the human body to measure individual physical conditions and changes. Building monitoring uses WSNs to identify potential problem with, for example, a building's structural integrity.

A WSN is a network formed by many sensors deployed within a sensing area, which have the ability to sense, wirelessly communicate, and process information. Sensors are not only able to sense and detect environmental targets and changes, but can also process collected data and send that processed data to a sink by wireless transmission [15]–[17]. In the sensing area, when a sensor node detects an event, it is referred to as a source node. A receiving node that collects sensing data in a WSN is referred to as a sink (base station). Sensors usually have limited battery power, and their batteries are quite difficult to replace. Therefore, designing an energy efficient data transmission scheme to reduce the overall energy consumption of a WSN is a very important issue.

Machine learning (ML), a branch of artificial intelligence (AI), is a method of classifying collected data, and then learning and training data by collecting large amounts of raw data for subsequent decision-making. The schemes proposed in this paper applies the concept of machine learning to the two-layer dissemination protocol of WSNs, so that learning agents can choose energy efficient data dissemination paths in the current environment.

The contributions of the proposed schemes are as follows:

1. Two two-tier data dissemination schemes based on Q-learning for WSNs are proposed, called TTDD-QL and TTDD-QL-A.
2. The source node uses Q-learning to find the most energy efficient data dissemination path from the source node to the sink. Two reward update strategies are used in the Q-learning process to improve the data dissemination performance.
3. The proposed TTDD-QL and TTDD-QL-A schemes can effectively reduce the energy consumption of sensor nodes compared with TTDD.

The remainder of this paper is organized as follows. Section 2 introduces previous studies relevant to this research. Section 3 presents the proposed data dissemination schemes in detail. Simulation results are discussed in Section 4. Finally, Section 5 offers conclusions.

II. RELATED WORK

In WSNs, the grid-based data transmission method mainly uses position information to establish a grid structure, and transmits the collected sensing data to the sink through the sensor nodes in the grid.

In TTDD [18], when a node in the sensing area detects an event, that node becomes the source node, and will divide the sensing area into a grid structure containing many cells (grid of cells). When the grid structure is completed, the intersections on the grid are called grid points, and the sensing nodes closest to the grid points are called dissemination nodes, which are responsible for transmitting information to the sink. In PADD [19], only a few dissemination nodes need to know the transmitted information during data dissemination. An appropriate cell size is selected to ensure that the dissemination node can transmitted directly to its eight neighboring dissemination nodes. In addition, when establishing a dissemination path, the method selects a dissemination node with the largest remaining energy to transmit query packets and sensing data in order to evenly distribute the energy consumption of the sensors on the WSN.

CODE [20] divides the sensing area into grids, and each grid has one coordinator that plays the role of an intermediate node to store and transfer data. In the scheme, a data transmission path is established in advance. The source node then sends data to the sink along this path. LEUGB [21] proposes two different types of base station positions, and discusses the effects of different base station placements in a uniformly divided grid for WSNs. The base station is placed at the center and corner positions of the sensing area, and then the network is divided into different numbers of grids to solve the problem of base station placement and grid division. EEUGCR [22] divides the network into grids of different sizes. In each grid, the cluster head is selected based on the remaining power of the sensor node, and the distance between the node and the center point of the grid. The cluster head collects data from the sensor nodes in the grid, aggregates the data, and transmits it to the base station via a multi-hop transmission method.

Reinforcement learning is a machine learning approach [23], [24], which includes enhanced learning as a kind of reinforcement learning algorithm. In the enhanced learning algorithm, the learning agent must try all possible actions [25]. After continuous learning, agents can find an optimal strategy in various states. Enhanced learning mainly defines five terms: agent, environment, reward, state and action. The agent interacts with the environment, and the agent sends an action to the environment, causing the environment to change the current state and return the reward value. Q-learning (QL) is an algorithm for reinforcement learning [26], [27]. In Q-learning, the agent records the values of all possible actions in each state, and records the current state and actions in the Q table. In each state, the agent chooses the action with the highest Q value to perform, and then the Q value is updated. After a number of learning iterations, the Q values reach convergence, and a choice of the best strategy can be found.

The following will introduce the working principle of Q-learning. In Q-learning, the agent records the values of all possible actions in each state, and stores them in the Q table. The Q value is updated by Equation (1):

$$Q(s_t, a_t) \leftarrow Q(s_t, a_t) + \alpha[r_t + \gamma \max_a Q(s_{t+1}, a) - Q(s_t, a_t)] \quad (1)$$

where $Q(s_t, a_t)$ is the expected value obtained by performing action a_t under state s_t, and α is the learning rate. If the previous action is believed, then, usually, α is set to a lower value, otherwise it is increased. If α is 0, it means that learning no longer occurs, and α is usually a value between 0 and 1. γ is the discount rate. After repeated iterations, the learning experience will gradually decrease. If more attention is paid to the future, γ is usually increased, and γ is a value between 0 and 1. r_t is the reward obtained by the learning agent after performing action a_t, and $\max_a Q(s_{t+1}, a)$ is the maximum Q value obtained in all actions a in the state s_{t+1}.

III. DATA DISSEMINATION BASED ON Q-LEARNING

In this study, each sensor node is energy-limited and knows its remaining energy. Each sensing node has position awareness capability, and its position information can be obtained through the global positioning system (GPS) [28], [29]. The proposed schemes each consist four phases: grid construction, selection of dissemination nodes, establishment of data dissemination path, and data transmission.

A. GRID CONSTRUCTION

In a WSN, when a sensor node detects an event, the node becomes the source node, and it divides the sensing area into a grid of cells. The cell size is defined as δ. The source node defines its position as a grid point on the grid structure, and then transmits the data outward to the boundary of the sensing area. Suppose the position of the source node is $L_S = (x_0, y_0)$, and the positions of other grid points are $L_P = (x_i, y_j)$, such that $x_i = x + i\delta$, $y_j = y + j\delta$; $i, j = \pm 1, \pm 2, \ldots$

In the sensing area, the geographic location of the sensing node is represented by (x, y). In that each node may know its own location information, they are equipped with a global positioning system (GPS) device. A logical grid structure is shown in Figure 1.

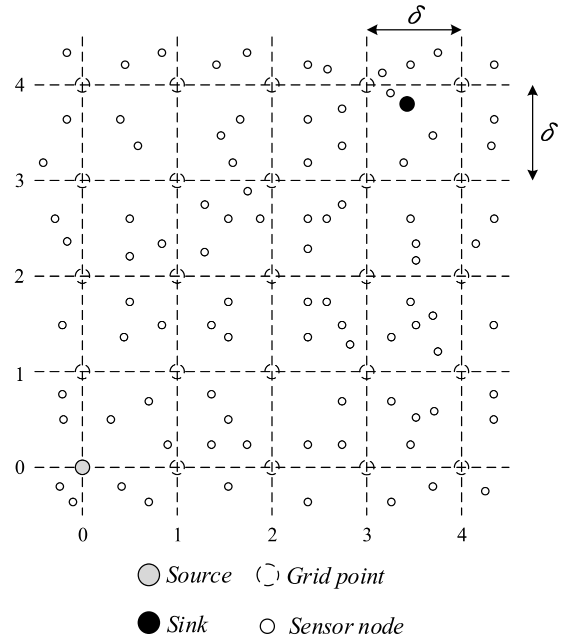


FIGURE 1. A logical grid structure.

B. SELECTION OF DISSEMINATION NODES

Once the grid structure is completed, each grid point in the grid infrastructure selects a sensor node to serve as a dissemination node, which is responsible for storing and transferring data. The sensor nodes closest to the grid points are selected for data dissemination, as shown in Figure 2.

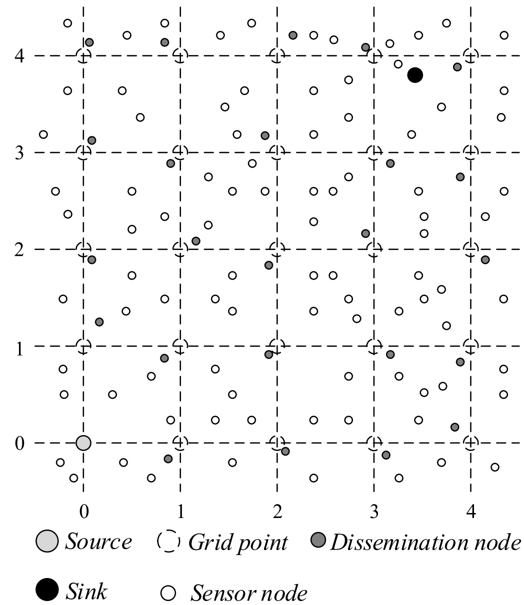


FIGURE 2. Selection of dissemination nodes.

C. ESTABLISHMENT OF DATA DISSEMINATION PATH

When an event is detected, the source node sends an announcement message to its four neighboring dissemination nodes. The announcement message contains two parameters: the location of source node and the cell size. When the dissemination nodes receive the announcement message, they store their own ID and the two parameters. They then forwards the announcement message to their four neighboring dissemination nodes, with the exception of the upstream node from which they received the message. The announcement message is thus recursively transmitted to all dissemination nodes in the sensor field. When the sink receives an announcement message from a source, the sink sends a query message to its neighboring immediate dissemination node, and the neighboring immediate dissemination node forwards the query message upstream from the sink until the query message is transmitted to the source node.

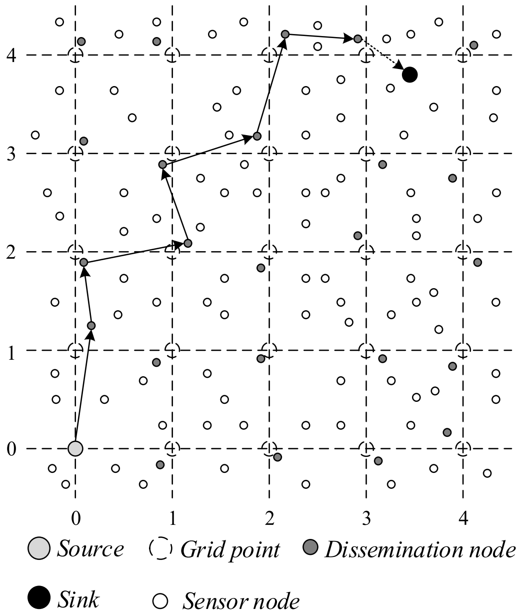


FIGURE 3. An example of a two-tier data dissemination path.

When the source node receives the query message, it then starts to establish a two-tier data dissemination path for data dissemination. The first-tier data dissemination path is from "the source node" to "the dissemination node closest to the sink". The second-tier data dissemination path is from "the dissemination node closest to the sink" to "the sink". An example of a two-layer data dissemination path is shown in Figure 3.

In the proposed schemes, this study includes a Q-learning mechanism in the grid-based data dissemination scheme in order to find an energy efficient data transmission path. The agent thus records the Q values of all possible actions in each state. First, some relevant terminologies must be defined:

(1) Agent: Learning agent.
(2) State: Each dissemination node represents a state.
(3) Action: The sensor node transmits the data to the next sensor node, which is called an action.
(4) Q table: The learning agent computes the Q values and records them in the Q table, from which the data dissemination path can be obtained. In the Q table, the initial Q values are all zero.
(5) Reward: The reward value that the learning agent receives when performing an action.

In the proposed schemes, the agent records each of the dissemination nodes and all possible transmission paths, and computes the state transition reward values in order to update the Q values. Once the update process is completed, the updated Q values are stored in the Q table. The Q table can then be used to obtain the data dissemination path. The Q value is updated according to Equation (2):

$$Q(S_t, A_t) \leftarrow Q(S_t, A_t) + \alpha \left[R_t + \gamma \max_{A} Q(S_{t+1}, A) - Q(S_t, A_t)\right] \quad (2)$$

where $Q(S_t, A_t)$ is the expected Q value obtained by the learning agent performing the action A_t in the current state S_t. α is the learning rate, usually a value between 0 and 1. R_t is the reward obtained by the learning agent performing the action A_t. γ is the discount rate, usually a value between 0 and 1. $\max_{A} Q(S_{t+1}, A)$ is the maximum Q value in the state S_{t+1} for all actions A. Thus, through Q-learning, an energy efficient data transmission path can be obtained from the Q table. The TTDD-QL algorithm is shown as Algorithm 1:

Algorithm 1 Data Dissemination Based on TTDD-QL

```
Construct the grid infrastructure of the sensing area
 with m × n cells
Decide the grid points GP[i, j], where 1 ≤ i ≤ m
 and 1 ≤ j ≤ n, the grid points are the intersections
 on the grid
Decide the dissemination nodes DN[i, j], where
 1 ≤ i ≤ m and 1 ≤ j ≤ n, the dissemination node
 DN[i, j] is the sensor node closest to the grid point
 GP[i, j]
Initialize Q(s, a) // start Q-learning
s ← DN[i, j];
while Q is not converged {
    Initialize Qrouting, Rear
    s ← Source;
    while s is Sink {
        Select a from s using the policy derived
        from exploration strategy of Q
        (e.g. ϵ-greedy)
        reward ← R(s, a); // reward update
        s′ ← T(s, a); // state transition
        Q(s, a) ← Q(s, a) × α ×
        (reward +γ × max_a Q(s′, a′) − Q(s, a));
        s ← s′;
        Qrouting[Rear] ← s′;
        Rear ← Rear + 1;
    }
 }
 return Qrouting;
```

The following proposes two data dissemination schemes using two different reward strategies. The first scheme is a two-tier data dissemination scheme based on Q-learning, called TTDD-QL. In TTDD-QL, the reward is determined by the distance between the current dissemination node and the sink. The second scheme is an advanced version of the two-tier data dissemination scheme based on Q-learning, called TTDD-QL-A. In TTDD-QL-A, the reward is determined not only by the distance between the current dissemination node and the sink, but also by the remaining energy of the current dissemination node.

First, two reward factors r_1 and r_2 are defined, and then the definitions of the rewards in TTDD-QL and TTDD-QL-A are given. r_1 is the distance reward factor, determined according to d_{iB}, where d_{iB} represents the distance between the current (ith) dissemination node and the sink. d_{SB} represents the distance between the source node and the sink. r_2 is the energy reward factor, determined according to E_i, where E_i represents the remaining energy of the current (ith) dissemination node, and E_{max} represents the initial power of the current dissemination node. The definition of r_1 is shown in Equation (3), and the definition of r_2 is shown in Equation (4). *Reward*(TTDD-QL) is shown in Equation (5), and *Reward*(TTDD-QL-A) is shown in Equation (6):

$$r_1 = \begin{cases} 1, & \frac{4}{5}d_{SB} < d_{iB} \le d_{SB} \\ 2, & \frac{3}{5}d_{SB} < d_{iB} \le \frac{4}{5}d_{SB} \\ 3, & \frac{2}{5}d_{SB} < d_{iB} \le \frac{3}{5}d_{SB} \\ 4, & \frac{1}{5}d_{SB} < d_{iB} \le \frac{2}{5}d_{SB} \\ 5, & 0 < d_{iB} \le \frac{1}{5}d_{SB} \end{cases} \tag{3}$$

$$r_2 = \begin{cases} 1, & 0 < E_i \le \frac{1}{5}E_{max} \\ 2, & \frac{1}{5}E_{max} < E_i \le \frac{2}{5}E_{max} \\ 3, & \frac{2}{5}E_{max} < E_i \le \frac{3}{5}E_{max} \\ 4, & \frac{3}{5}E_{max} < E_i \le \frac{4}{5}E_{max} \\ 5, & \frac{4}{5}E_{max} < E_i \le E_{max} \end{cases} \tag{4}$$

$$Reward(\text{TTDD-QL}) = r_1 \tag{5}$$

$$Reward(\text{TTDD-QL-A}) = r_1 + r_2 \tag{6}$$

D. DATA TRANSMISSION

Once the construction of the data transmission path is complete, data transmission begins. In the data transmission process, the source node transmits data along the decided data dissemination path until the data reaches the sink, and the data dissemination of this round is complete. The selection of the dissemination nodes and the data dissemination path based on Q-learning is re-established at the beginning of each data dissemination round. The proposed data dissemination schemes can evenly distribute energy consumption of sensor nodes, thereby extending the lifetime of the WSN.

The key advantages of the proposed schemes are given below:

1. At the beginning of each data dissemination round, the dissemination nodes are re-selected and the data dissemination path with Q-learning is re-established, so that the energy consumption of the sensor nodes can be distributed efficiently.
2. The proposed schemes use Q-learning to find an energy efficient data dissemination path for data transmission, which can effectively shorten the data dissemination path and extend the network lifetime.

IV. SIMULATION RESULTS

In the simulations conducted in this research, the first order radio model [30], [31] was adopted to evaluate the power consumption of sensor nodes. E_{elec} was the power consumption of the circuit itself, where $E_{elec} = 50$ nJ/bit. E_{amp} was the power consumed by the amplifier when transmitting packets, where $E_{amp} = 100$ pJ/bit/m^2. A transmission amplifier at the sender node further consumed $E_{amp}d^2$, where d was the distance between nodes. Thus, the power consumed to transmit a k-bit message a distance d was:

$$E_{Tx}(k, d) = E_{elec} \times k + E_{amp} \times k \times d^2 \tag{7}$$

and the power consumed to receive this message was:

$$E_{Rx}(k) = E_{elec} \times k \tag{8}$$

This study developed a simulator with MATLAB for performance evaluation. Simulation experiments were conducted for three data dissemination schemes: TTDD, TTDD-QL, and TTDD-QL-A. The simulation network area was 400 m × 400 m, the number of nodes ranged from 100 to 500, and the node transmission range was set to 150 m. The initial power of the sensors was set to 0.25 J and 0.5 J, respectively. The cell size was set to 100 m, and the packet size was 2000 bits. The simulation parameters are listed in Table 1.

TABLE 1. Parameters used in the simulations.

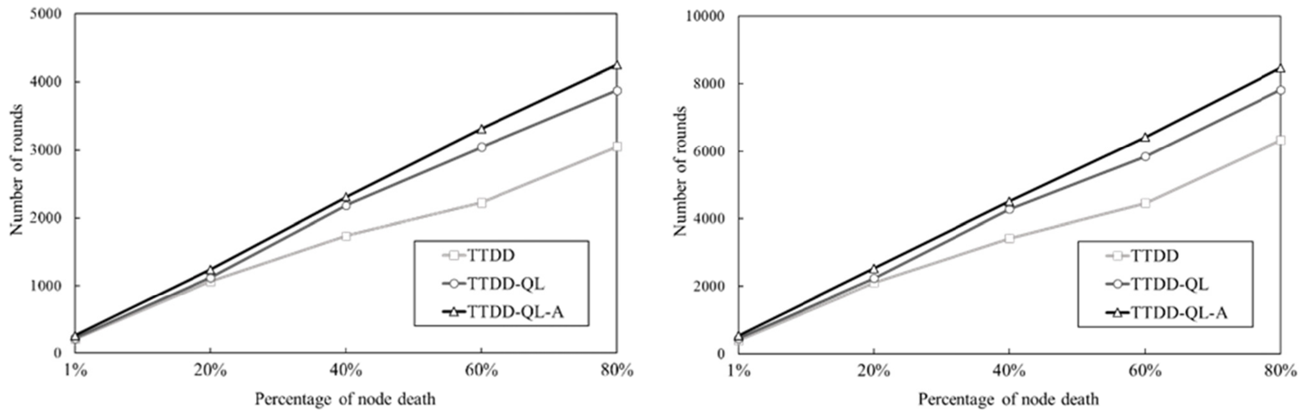
Parameter	Value
Network size	400 m × 400 m
Number of nodes	100 - 500
Initial power	0.25 J/node, 0.5 J/node
Cell size	100 m
Transmission range	150 m
Packet size	2000 bits
Learning rate	0.9
Discount rate	0.8

In our simulations, the cell size is determined based on the following two criteria:

1. The cell size was smaller than the transmission range, so that the data could be transmitted directly by one-hop.
2. The cell size was as large as possible in order to achieve better transmission performance.

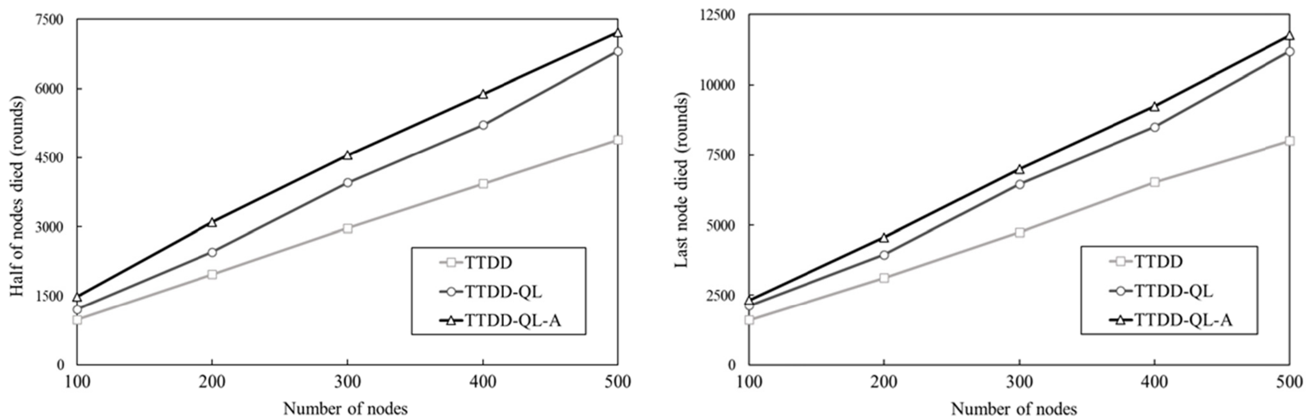
A. IMPACT OF INITIAL ENERGY

Firstly, the number of rounds executed for the different percentages of node death were examined. 200 nodes were used with cell size 100 m, and the initial power was 0.25 J and 0.5 J. Three schemes were simulated: TTDD-QL, TTDD-QL-A and TTDD, respectively. When the percentage of node deaths for each scheme reached 1%, 20%, 40%, 60%, and 80%, the number of rounds that each simulated scheme could perform was noted. In Figure 4, TTDD-QL-A was able to perform a great number of rounds than TTDD-QL and TTDD.



(a) (b)

FIGURE 4. Impact of initial energy: (a) 0.25 J and (b) 0.5 J.



(a) (b)

FIGURE 5. Impact of number of nodes: (a) when half of the nodes died and (b) when last node died.

TTDD-QL-A was able to perform about 1.1 times more rounds than TTDD-QL, and 1.4 times more than TTDD. In TTDD-QL-A, Q-learning was used to find an energy efficient data dissemination path. The Q-learning considered the remaining energy of the dissemination nodes and the distance between the dissemination nodes and the sink. In each transmission round, the maximum Q value was selected for transmission, which effectively shared the energy consumption of the dissemination nodes. Simulation results show that the more initial power a node has, the more rounds it can perform.

B. IMPACT OF NUMBER OF NODES

Next, the impact of different numbers of sensor nodes was examined by observing the number of rounds that each scheme was able to execute when half of the nodes died and when the last node died. The initial power was set to 0.25 J, and the cell size was set to 100 m. Each simulation was conducted first with 100 nodes, and repeated with an additional 100 more nodes, up to 500 nodes. Figure 5 shows the number of execution rounds that each network was able to perform under different conditions. In Figure 5(a), when half of the nodes died, the lifetime of TTDD-QL-A was better than that of TTDD-QL and TTDD. In Figure 5(b), when the last node died, the lifetime of TTDD-QL-A was much higher than the other two schemes. Simulation results show that the higher the number of nodes, the higher the number of rounds that can be executed. As the number of nodes increased, more nodes were able to share the energy consumed during data transmission from the source node to the sink, extending the lifetime of the WSN.

C. IMPACT OF CELL SIZE

This study also explored the effect of different cell sizes on the lifetime of the WSN by observing the number of rounds different schemes were able to perform with an initial power of 0.25 J. 200 nodes were used, with an initial cell

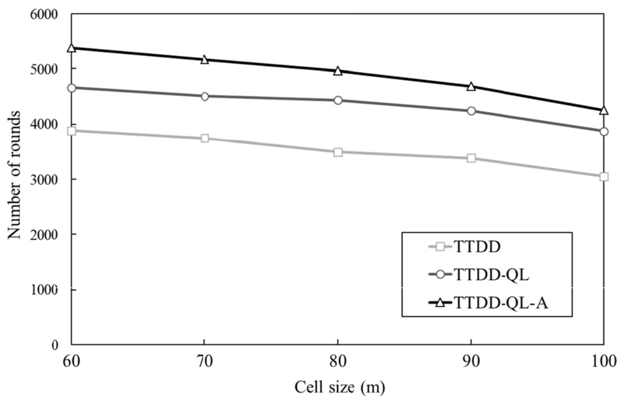


FIGURE 6. Number of rounds vs. cell size.

size of 60 m, which was increased by 10 m each simulation, up to 100 m. Figure 6 shows the number of execution rounds each network was able to perform with various cell sizes. In Figure 6, TTDD-QL-A achieved the best performance. Simulation results show that as the cell size increased, the number of execution rounds each scheme was able to perform decreased. As cell size decreased, so too did the transmission distance between the dissemination nodes during the data transmission process from the source node to the sink. When the transmission distance was reduced, so too was the energy consumed, thus the smaller the cell, the higher the number of execution rounds.

D. NETWORK LIFETIME

The lifetime of a WSN is defined as the number of transmission rounds performed from the time that a node begins transmitting data until the time at which the node can no longer transmit data. In order to compare the lifetime achieved using various schemes, the initial node power was set to 0.25 J, the cell size was set to 100 m, and 200 nodes were used. As shown in Figure 7, the network using TTDD-QL-A completed 4768 transmission rounds, the network using TTDD-QL completed 4511, and the network using TTDD completed 3289. Therefore, TTDD-QL-A can effectively extend the lifetime of the WSN.

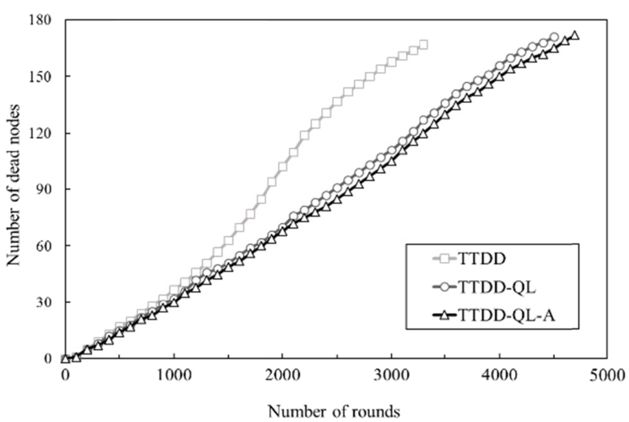


FIGURE 7. Number of dead nodes vs. number of rounds.

E. TOTAL ENERGY CONSUMPTION

Finally, this study examined the performances of the three schemes in terms of total energy consumption. The total energy consumption includes the data transmission cost of the transmitter and the receiver. In the simulations conducted, 200 nodes with an initial power of 0.25 J were used, so it can be calculated that their initial total energy was set to 50 J and the cell size was set to 100 m. The total energy consumption corresponding to the number of execution rounds is shown in Figure 8. In the three schemes, as the number of rounds performed increased, so too did the total energy consumption. In the same number of rounds, TTDD-QL-A exhibited the lowest total energy consumption. This is because TTDD-QL-A uses Q learning to find the most energy efficient transmission path from the source node to the sink, which effectively reduce energy consumption.

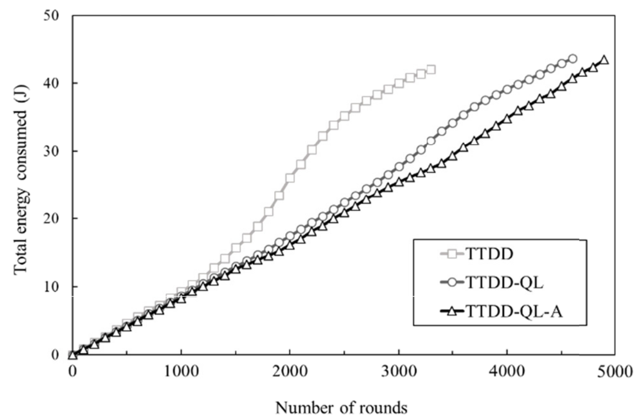


FIGURE 8. Total energy consumed vs. number of rounds.

V. CONCLUSION

This paper proposed two two-tier data dissemination schemes based on Q-learning for WSNs, called TTDD-QL and TTDD-QL-A. When a node in the sensing area detects an event, that node becomes the source node. The source node will then divide the sensing area into a grid structure consisting of cells. The intersections on the grid are called grid points, and the sensor node closest to a grid point is called a dissemination node, and is responsible for transferring data to the sink. In the proposed schemes, the source node uses Q-learning to find the most energy efficient data dissemination path from the source node to the sink. In TTDD-QL-A, since the remaining energy of the dissemination nodes and the distance between the sink and the dissemination nodes are considered in order to update the reward values, a better dissemination path can be found by the Q-learning process. Simulation results show that the proposed TTDD-QL and TTDD-QL-A schemes can effectively reduce the energy consumption of the sensor nodes compared with TTDD. Of the two proposed schemes, TTDD-QL-A achieves better performance in extending the lifetime of the WSN.

REFERENCES

[1] R. K. Lenka, A. K. Rath, and S. Sharma, "Building reliable routing infrastructure for green IoT network," *IEEE Access*, vol. 7, pp. 129892–129909, Sep. 2019.

[2] A. Ghosh, A. Maeder, M. Baker, and D. Chandramouli, "5G evolution: A view on 5G cellular technology beyond 3GPP release 15," *IEEE Access*, vol. 7, pp. 127639–127651, Sep. 2019.

[3] K. N. E. A. Siddiquee, K. Andersson, F. J. M. Arrebola, Z. Abedin, and M. S. Hossain, "Estimation of signal coverage and localization in Wi-Fi networks with AODV and OLSR," *J. Wireless Mobile Netw., Ubiquitous Comput., Dependable Appl.*, vol. 9, no. 3, pp. 11–24, Sep. 2018.

[4] G. Choudhary, J. Kim, and V. Sharma, "Security of 5G-mobile backhaul networks: A survey," *J. Wireless Mobility Netw., Ubiquitous Comput. Dependable Appl.*, vol. 9, no. 4, pp. 41–70, Dec. 2018.

[5] C.-S. Shih, J.-J. Chou, and K.-J. Lin, "WuKong: Secure run-time environment and data-driven IoT applications for smart cities and smart buildings," *J. Internet Services Inf. Secur.*, vol. 8, no. 2, pp. 1–17, May 2018.

[6] D. Estrin, R. Govindan, J. Heidemann, and S. Kumar, "Next century challenges: Scalable coordination in sensor networks," in *Proc. 5th Annu. ACM/IEEE Int. Conf. Mobile Comput. Netw. (MobiCom)*, Aug. 1999, pp. 263–270.

[7] J. P. Macker and M. S. Corson, "Mobile Ad Hoc Networking and the IETF," *ACM SIGMOBILE Mobile Comput. Commun. Rev.*, vol. 2, no. 2, pp. 9–14, Jan. 1998.

[8] H. Li, K. Ota, M. Dong, and M. Guo, "Mobile crowdsensing in software defined opportunistic networks," *IEEE Commun. Mag.*, vol. 55, no. 6, pp. 140–145, Jun. 2017.

[9] R. Min, M. Bhardwaj, S. H. Cho, E. Shih, A. Sinha, A. Wang, and A. Chandrakasan, "Low-power wireless sensor networks," in *Proc. Int. Conf. VLSI Design*, Jan. 2001, pp. 205–210.

[10] Q. Jiang and D. Manivannan, "Routing protocols for sensor networks," in *Proc. Consum. Commun. Netw. Conf.*, Jan. 2004, pp. 93–98.

[11] J. N. Al-Karaki and A. E. Kamal, "Routing techniques in wireless sensor networks: A survey," *IEEE Wireless Commun.*, vol. 11, no. 6, pp. 6–28, Dec. 2004.

[12] E. Shih, S.-H. Cho, N. Ickes, R. Min, A. Sinha, A. Wang, and A. Chandrakasan, "Physical layer driven protocol and algorithm design for energy-efficient wireless sensor networks," in *Proc. 7th Annu. Int. Conf. Mobile Comput. Netw. (MobiCom)*, Jul. 2001, pp. 272–286.

[13] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, "A survey on sensor networks," *IEEE Commun. Mag.*, vol. 40, no. 8, pp. 102–114, Aug. 2002.

[14] I. Akyildiz, W. Su, Y. Sankarasubramanian, and E. Cayirci, "Wireless sensor networks: A survey," *Comput. Netw.*, vol. 38, pp. 393–422, Dec. 2002.

[15] F. J. Block and C. W. Baum, "Energy-efficient self-organizing communication protocols for wireless sensor networks," in *Proc. Commun. Netw.-Centric Oper., Creating Inf. Force (MILCOM)*, Oct. 2001, pp. 362–366.

[16] D. Braginsky and D. Estrin, "Rumor routing algorithm for sensor networks," in *Proc. ACM Int. Workshop Wireless Sensor Netw. Appl.*, pp. 22–31, Sep. 2002.

[17] W. R. Heinzelman, J. Kulik, and H. Balakrishnan, "Adaptive protocols for information dissemination in wireless sensor networks," in *Proc. 5th Annu. ACM/IEEE Int. Conf. Mobile Comput. Netw. (MobiCom)*, Aug. 1999, pp. 174–185.

[18] F. Ye, H. Luo, J. Cheng, S. Lu, and L. Zhang, "A two-tier data dissemination model for large-scale wireless sensor networks," in *Proc. 8th Annu. Int. Conf. Mobile Comput. Netw. (MobiCom)*, Sep. 2002, pp. 148–159.

[19] N.-C. Wang and Y.-K. Chiang, "Power-aware data dissemination protocol for grid-based wireless sensor networks with mobile sinks," *IET Commun.*, vol. 5, no. 18, pp. 2684–2691, Dec. 2011.

[20] L. Xuan Hung and S. Lee, "A coordination-based data dissemination protocol for wireless sensor networks," in *Proc. Intell. Sensors, Sensor Netw. Inf. Process. Conf.*, Dec. 2004, pp. 13–18.

[21] T. J. Swamy, T. Sandhya, and G. Ramamurthy, "Link evaluation of uniform grid based wireless sensor networks to base station with leveling and clustering," in *Proc. Int. Conf. Comput. Commun. Inform.*, Jan. 2013, pp. 1–5.

[22] B. Prince, P. Kumar, M. P. Singh, and J. P. Singh, "An energy efficient uneven grid clustering based routing protocol for wireless sensor networks," in *Proc. Int. Conf. Wireless Commun., Signal Process. Netw. (WiSPNET)*, Mar. 2016, pp. 1580–1584.

[23] R. S. Sutton and A. G. Barto, *Reinforcement Learning: An Introduction*. Cambridge, MA, USA: MIT Press, Mar. 1998, pp. 1–322.

[24] L. P. Kaelbling, M. L. Littman, and A. W. Moore, "Reinforcement learning: A survey," *J. Artif. Intell. Res.*, vol. 4, no. 1, pp. 237–285, May 1996.

[25] J. Jiang, M. Kamel, and L. Chen, "Reinforcement learning and aggregation," in *Proc. IEEE Int. Conf. Syst., Man Cybern.*, Oct. 2004, pp. 1303–1308.

[26] T. Fukao, T. Sumitomo, N. Ineyama, and N. Adachi, "Q-learning based on regularization theory to treat the continuous states and actions," in *Proc. IEEE Int. Joint Conf. Neural Netw., IEEE World Congr. Comput. Intell.*, vol. 2, May 1998, pp. 1057–1062.

[27] C. J. C. H. Watkins and P. Dayan, "Q-learning," *Mach. Learn.*, vol. 8, nos. 3–4, pp. 279–292, May 1992.

[28] J. Albowicz, A. Chen, and L. Zhang, "Recursive position estimation in sensor networks," in *Proc. 9th Int. Conf. Netw. Protocols (ICNP)*, Nov. 2001, pp. 35–41.

[29] E. D. Kaplan, *Understanding GPS: Principles and Applications*. Norwood, MA, USA: Artech House, 1996.

[30] H. Ö. Tan and I. Körpeoğlu, "Power efficient data gathering and aggregation in wireless sensor networks," *ACM SIGMOD Rec.*, vol. 32, no. 4, pp. 66–71, Dec. 2003.

[31] P. Andreou, A. Pamboris, D. Zeinalipour-Yazti, P. K. Chrysanthis, and G. Samaras, "ETC: Energy-driven tree construction in wireless sensor networks," in *Proc. 10th Int. Conf. Mobile Data Manage., Syst., Services Middleware*, May 2009, pp. 35–40.

NENG-CHUNG WANG (Member, IEEE) received the B.S. degree in information and computer engineering from Chung Yuan Christian University, Taiwan, in June 1990, and the M.S. and Ph.D. degrees in Computer Science and Information Engineering from National Cheng Kung University, Taiwan, in June 1998 and June 2002, respectively. He joined the faculty of the Department of Computer Science and Information Engineering, Chaoyang University of Technology, Taiwan, as an Assistant Professor, in August 2002. From August 2006 to July 2007, he was an Assistant Professor with the Department of Computer Science and Information Engineering, National United University, Taiwan, where he was an Associate Professor with the Department of Computer Science and Information Engineering, from August 2007 to July 2011. Since August 2011, he has been a Full Professor with the Department of Computer Science and Information Engineering, National United University. His research interests include wireless and mobile networks, wireless communications, mobile computing, and the Internet of Things. He is a member of the IEEE Communication Society and the IEEE Computer Society.

WEI-JUNG HSU received the B.S. degree from the Information and Communications Department, Chaoyang University of Technology, Taiwan, in June 2016, and the M.S. degree in Computer Science and Information Engineering from National United University, Taiwan, in June 2018. His research interests include wireless networks and machine learning.