IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received April 22, 2019, accepted May 21, 2019, date of publication May 24, 2019, date of current version June 10, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2918928

Reducing Delay and Maximizing Lifetime for Wireless Sensor Networks With Dynamic Traffic Patterns

QIANG WANG[1], WEI LIU[2], TIAN WANG[3], MING ZHAO[1], XIONG LI[4], MANDE XIE[5], MING MA[6], GUOPING ZHANG[7], AND ANFENG LIU[1]
[1]School of Computer Science and Engineering, Central South University, Changsha 410083, China
[2]School of Informatics, Hunan University of Chinese Medicine, Changsha 410208, China
[3]College of Computer Science and Technology, Huaqiao University, Xiamen 361021, China
[4]School of Computer Science and Engineering, Hunan University of Science and Technology, Xiangtan 411201, China
[5]School of Computer Science and Information Engineering, Zhejiang Gongshang University, Hangzhou, Zhejiang 310018, China
[6]Department of Computer Science, Stony Brook University, Stony Brook, NY 11794, USA
[7]Faculty of Informatics and Electronics, Zhejiang SCI-Tech University, Hangzhou 310018, China

Corresponding author: Mande Xie (xiemd@zjgsuedu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 61772554, Grant 61572528, and Grant 6157256, in part by the Open Research Project of the State Key Laboratory of Industrial Control Technology, Zhejiang University, China, under Grant ICT1800391, and in part by the Natural Science Foundation of Zhejiang Province under Grant LY17F020032.

ABSTRACT Fast and energy efficient data collection is the most important function of wireless sensor networks (WSNs) which governs the performance of WSNs. Time division multiple access (TDMA) is an effective data collection strategy in which node only schedules the active slot to save energy when transmitting data. However, many previous studies are designed for static traffic pattern such that each node to generate one data packet in a data collection cycle. Moreover, all nodes in the network use the same transmission radius r, so that the network data collection has a larger delay, a lower lifetime, and energy utilization rate. In view of the shortcomings of previous research, this paper proposes a reducing delay and maximizing lifetime (RDML) data collection strategy for WSNs with dynamic traffic patterns. The main innovation point of our work is the data transmission radius is no longer unified as r, while some nodes with energy surplus use a data transmission radius of $mr (m > 1)$. The following network performance can be achieved by adopting the RDML scheme. First, because the nodes in the near sink mr radius area can send data directly to sink, the amount of data carried by hotspots nodes in the original strategy is greatly reduced, and the lifetime is improved. Second, since some nodes can go mr distance to the sink in data transmission, thus accelerating the data routing and reducing the time needed for data collection. Finally, since the remaining energy of nodes is fully utilized to increase the radius of data transmission, the energy utilization rate can be effectively improved. Through the theoretical analysis and experimental results, the RDML scheme can reduce the maximum delay of data collection by 22.32% and improve lifetime as well as the energy utilization rate by 0.86%−84.40%, and 23.33%−88.00%, respectively, compared with previous strategies.

INDEX TERMS Dynamic data collection, mixed transmission radius, TDMA scheduling, delay, lifetime.

I. INTRODUCTION

With the rapid development of microprocessor technology, there are more and more types of sensing devices [1]–[3], and the types of information and data that can be perceived are diversified. This has greatly developed the application field and scope of sensor nodes [4]–[6], thus greatly promoting the development of the network [7], [8]. According to current statistics, more than 5 billion terminal devices have been connected to the network, thus Internet of Things (IoTs) has been greatly developed [8], [9]. The network is undergoing a process from centralization to edge network development [6], [10], [11], all of which have benefited from sensing devices based on the perception and acquisition of data in the objective world [12]–[16]. Wireless sensor networks is

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

one of the most important networks, and thus has become the focus of research.

In wireless sensor networks (WSNs), a large number of sensor nodes are deployed in areas that need to be monitored. Once the interested event or physical phenomenon is detected, the perceived data will be sent to sink through multi-hop routing. Then sink analyzes the data and makes decisions on the event [17]–[20]. Data collection is the most basic and important function of WSNs [21], [21]–[25]. However, high lifetime [19], [23], [24], [26]–[29] and low latency [18], [19], [26], [30] data transmission are the two key issues of the most worthwhile research in data collection [31], [32]. Data collection is also the most energy-consuming operation in WSNs [1]–[4], [7], [13], [14], [23], [31]. Therefore, the designed data collection strategy governs the performance of WSNs [23], [31], [33]. Lifetime generally refers to the time when the first node in the network dies [19], [23], [24], [29]. Since sensor nodes are generally small in size and powered by batteries, their energy is extremely limited, so how to conserve energy to improve lifetime becomes the first important challenge in data collection [19], [23], [24], [29]. Another challenge to design an efficient data collection strategy is how to collect all data packets in the network to sink quickly [18], [19], [26], [30]. Because decision makers need to make decisions on the monitored objects based on the collected data, WSNs are often applied to such as industrial monitoring and event detection which need to make quick response. Therefore, rapid data collection is of great significance for these applications, and delayed data collection can cause significant losses.

Researchers have proposed numerous data collection strategies. According to whether the time division is strictly divided in the data collection, it can be divided into Time Division Multiple Access (TDMA) MAC protocol [7], [32] and contention-based MAC protocol [18], [19], [21], [22], [24]. The main characteristics of the data collection strategy based on TDMA are that the time of nodes in the network is divided into slots one by one, and the node can complete a transmission in one time slot [7], [32]. Each node is active only in the slot where the data operation is performed and sleeping in the slot where there is no data operation. Moreover, the slot in which the node performs data operations is carefully scheduled, so that there is no interference between the nodes performing data operations at the same time. It can clearly be seen that TDMA-based MAC protocol is very effective in energy consumption, and thus has been widely studied [7], [32].

There are already some data collection strategies based on TDMA. This type of TDMA-based data collection is mostly applicable to networks where nodes generate data that is fixed. It is generally assumed that each node will generate one packet in a data collection cycle, and these packets need to be sent to the sink. However, Zhao and Tang [7] believe that this static network traffic pattern that generates one packet for each cycle has some applications, but in more applications, the data generation of the network is often dynamic. Some nodes generate data in a cycle, while some nodes do not need to generate data because there is no event. Therefore, the number of packets that need to be collected in each cycle of data collection is not equal. The general TDMA scheduling algorithm is to arrange the data operation slot for each node according to the situation that each node generates data in each cycle of data collection. Although this data collection strategy can also be applied to networks with dynamic data generation, where the node's operation is empty as long as there is no data in the slot assigned to the node. However, this approach has the following disadvantages: First, the data collection delay of the network is large, as the time required for data collection has a great relationship with the number of collected data packets. When there are many nodes in the network and the amount of data is large, the time required for data collection is long. However, if the number of packets actually generated in the network is not large, data collection in a fixed manner in which each node generates data will cause a great delay. Second, the energy consumption of nodes increases, when there is no data to be sent, it is better to turn off the node and put it in sleep state to save energy. Obviously [18], [22], when a node has no data transmission, an empty operation will consume the node's energy.

For many applications, not every node generates one data packet in a cycle of data collection, Zhao et al. [7] proposed a data collection strategy for dynamic traffic patterns. The difficulty of the data collection strategy for WSNs with dynamic traffic patterns lies in that it is impossible to determine the number of packets to be transmitted by the nodes, so it is a great challenge to reasonably and economically arrange the active slots of the nodes. The main idea of scheduling strategy proposed by Zhao and Tang [7] is that each node sends continuous packets at its activation time slots, and if the receiver receives a packet from the sender, it will receive the packet sent by the sender continuously. After the sender sends all its own packets, it stops sending. While the receiver can confirm that all packets of the sender have been sent after experiencing an active slot with no data transmission, thus ending the data receiving of this node in advance and turning to other nodes for data operations. The essence of this approach is to arrange the data packets of the sender node in continuous active time slots, and add a slot of empty operation to indicate that the sender has finished sending the entire data packets, thereby realizing data collection of the dynamic data generation networks.

Although some scheduling strategies for data collection have been proposed, these strategies still have much room for improvement in reducing data collection delay, improving network lifetime, and balancing energy consumption. In this paper, a Reducing Delay and Maximizing Lifetime (RDML) data collection scheme is proposed for WSNs with dynamic traffic patterns. The main innovation points of our work are as follows:

(1) A data collection strategy with mixed node sending radius is proposed in this paper. In our proposed scheme, the data transmission radius is no longer unified to r, and

some nodes with energy remaining have a data transmission radius of $mr(m > 1)$. By adopting a data collection strategy with multiple data transmission radii, the following network performance can be achieved better than the previous strategy. (a) Owing to the data transmission radius of some nodes with residual energy is mr, the nodes in the near sink mr radius range can send data directly to sink. It can improve the network lifetime by avoiding the deficiency of the original strategy that only sink r range nodes send data to sink, which makes these nodes bear a huge amount of data. (b) Because some nodes adopt mr as their data transmission radius, they can advance mr distance to sink in one data transmission, thus speeding up data routing, which can reduce the time required for data collection. (c) Due to the remaining energy of the node is fully utilized to increase the radius of the node data transmission, the energy utilization rate can be effectively improved.

(2) For WSNs with dynamic traffic patterns, this paper completely proposes a TDMA data collection algorithm suitable for node sending radius r and mr mixing. In the previous data collection strategies, the data transmission radius of all nodes in the network is r, and when nodes send data, they only interfere with the data transmission of adjacent two-layer nodes, so the design of scheduling algorithm for data collection is relatively easy. When the sending radius of the node in the network is mr, the data transmission of the node will interfere with the transmission of upper and lower m-layer data, so its interference range is large. At the same time, different sending radii lead to great challenges in designing the data collection strategy. This paper is the first time to design a scheduling algorithm of TDMA with multiple transmission radii, which is innovative.

(3) Through theoretical analysis and experimental results, the RDML data collection strategy proposed in this paper can effectively improve network performance compared with previous strategies. The RDML scheme compared with previous strategies can reduce the maximum delay of data collection by 22.32% and improve lifetime as well as energy utilization by 0.86% - 84.40%, 23.33% - 88.00% respectively.

The rest of this paper is organized as follows: The related work are introduced in Section II. In Section III, the network model and problem statement are presented. Then, the design of RDML scheme is introduced in Section IV. Performance analysis of RDML scheme is presented in Section V. The experimental results and analysis are given in Section VI. Finally, Section VII provides conclusions.

II. BACKGROUND AND RELATED WORK

With the development of big data networks, the importance of data collection is increasingly reflected. In addition, with the development of current technology, the means and methods of data collection have become more and more diversified [5], [6], [8], [13], [16]. On the basis of a large amount of data collection, coupled with the development of artificial intelligence technology [34], [35], the hidden information in the data [36] can be fully mined, so that work can be carried out more effectively. It has been fully applied in social networks [9], [17], [36], product recommendation [36], location service [37], privacy and security protection [38]–[40]. In particular, with the development of sensing technology, various data can be perceived, transmitted, analyzed and processed anytime, anywhere. The emergence of Edge Network [41], [42], Fog Computing [43], [44], Transparent Computing [45], Cloud Computing [10], [17], [33], [43] and other network models and computing methods, as well as the rise of 5G network [46], [47] has pushed the current network development to a new height, and data collection plays an important role in it.

A. RESEARCH ON DATA COLLECTION STRATEGY

Data collection in WSNs has always been one of its most important functions. There have been many researchers who have proposed quite a number of data collection strategies. These data collection strategies can be divided into different categories based on different classification criteria. Firstly, according to whether competitive communication mode is adopted for data collection, it is divided into TDMA-based data collection schemes [7], [32] and contention-based data collection strategies [18], [19], [21], [24].

(1) Data collection scheme based on TDMA. In TDMA-based data collection, the time in the network is divided into slots with equal unit time intervals, and slot is the basic time unit for data operations [7], [32]. In general, a node can establish a connection and send or receive a data packet in one slot. And in this way, the clocks of the nodes in the network are synchronized. At present, there are two kinds of slot allocation strategies: fixed allocation and on-demand allocation. The fixed allocation strategy is to arrange the data operation slots for each node according to the scheduling algorithm before data collection. In the data collection, each node can complete the data collection according to the pre-arranged slots. This kind of data collection scheme can optimize scheduling in advance for the network generated by fixed data, so the scheduling efficiency is better, but it has poor adaptability to changes in the number of nodes. On-demand allocation is a dynamic time slot allocation technique, but not suitable for network topology updates. Obviously, the advantage of a TDMA-based data collection strategy is that the node does not need to be in the active state for a long time, but only when data operation is needed, it will be converted to the active state, thus saving energy. In addition, with a network based on TDMA-based data collection strategy, the additional communication required for data transmission between nodes is also relatively small. Because each node can perform the corresponding operation in the planned slot without prior communication. However, data collection networks based on TDMA require time synchronization between nodes, so the maintenance of additional time synchronization protocols also consumes energy. TDMA-based data collection strategies can also be divided into two types, one is data aggregation based on data collection strategy, the other is data collection strategy without data aggregation.

Since there is a correlation between the data collected by the sensor network nodes, when multiple data packets meet, it is able to aggregate smaller data. Therefore, the amount of data that the node needs to transmit can be reduced and the lifetime of the network can be increased, and thus the data collection approach based on data aggregation is widely used in WSNs. One of the special data aggregation TDMA data collection schemes has been the most studied [32]. This data aggregation means that when n packets meet, they can be merged into one packet. For example, in a wireless sensor network that monitors crops, the monitored data is the highest, lowest, and average temperature (or humidity) of the farm. Therefore, when multiple data packets meet, they can be merged into one data packet. In order to save energy consumption, the general data collection approach of this network is to convert the network into a tree. The data operation of each node complies with the following two-stage data collection rules, namely the data collection stage and the data transmission stage. The node only collects data in the first stage, that is, it only receives the data sent by the child nodes. When the first stage ends, the node fuses all the data packets it has collected into one data packet and sends it to its parent node. That is, only one data packet is sent during the data transmission stage, and the data packet is no longer received once the node performs data transmission. In this data collection strategy, the operation of the data begins with the leaf nodes of the tree, that is, the lowest layer of the network. When the parent node receives the data packets of all the son nodes, it sends data to the parent node of the upper layer. After data collection is performed layer by layer, sink receives the data of the entire network as the root node of the tree. Therefore, in this data collection strategy, on the basis of obeying the two stage data collection rules, the data operation can be performed concurrently without interference, so that the delay of data collection is minimized. Because in this kind of data collection strategy, the energy consumption between nodes is basically balanced. The advantage of this kind data collection approach, which abstracts the network into a tree structure, is that it can be applied to both networks with data aggregation and networks without data aggregation.

In the data aggregation network, the shortage of the tree structure network is that the data can only be collected layer by layer from the leaf layer, so the delay of data collection is still relatively large. Subsequent research has been improved. Li *et al.* [32] proposed a parallel data collection strategy. This strategy also has two stages: the first stage is the data collection and fusion within the cluster. Firstly, the network is divided into clusters, and each cluster member node sends data to the cluster head, which aggregates the data into a single packet. The second stage is the data collection between clusters. When all the cluster head data are merged into one data packet, all the cluster heads of the network will form a tree, and then the tree-based data collection strategy mentioned above can be used for data collection. However, the research of Li *et al.* [32] is different from the previous clustering data fusion strategy: In the previous clustering network, the size of each cluster is the same. But Li *et al.* [32] considered that in the clustered network with equalization, when the data collection in the cluster ends, the data collection of the cluster head node still needs to be collected layer by layer from the leaf nodes of the tree structure. At this time, the cluster head node in the near sink area needs to wait, which will cause a large delay. Therefore, Li *et al.* [32] proposed a data collection scheme with unequal clustering structure in which the near-sink area adopts a larger cluster and the far sink area uses a smaller cluster. In such an unequal clustering structure, since the clustering away from the sink area is small, the data collection in the cluster is terminated early and the data collection between the clusters is initiated. The clustering near the sink area is large, and the time for collecting data in the cluster is long. Therefore, when the data of the cluster head node far from the sink arrives, the data collection between the clusters exactly happens. In this way, data collection within the cluster and between clusters can be performed simultaneously in parallel, thereby reducing the time required for data collection. The difficulty of this scheme is that it is necessary to accurately control the data arrival time of the cluster head node away from the sink.

(2) Data collection scheme based on competition mechanism. In a competition-based data collection strategy, time does not need to be divided into slots and synchronization is not required. When a node needs to send data, it listens and preempts the channel first. Once the channel is contended, the node begins to transmit data [18], [19], [21], [22]. Obviously, the advantage of this strategy is that the time between nodes in the network does not need to be synchronized. When the node has data to send, it contends for the channel to send its own data. But in this way, the energy consumption of the node is larger than that based on the TDMA scheme. Because in the TDMA scheme, when the node has no data operation, it can get into sleep to save energy. However, in the strategy of competition mechanism, if the node has no data to send, it may be necessary to forward the data of other nodes as a relay node, so it is better to keep active. It may not be required to serve as a relay node in most of the time, but since it is unknown when other nodes have data to forward, so the node either remains active all the time or adopts the duty cycle mode to save energy [18], [22], [26]. The Duty cycle mode is a way of working in which nodes periodically wake up and sleep [18], [22], [26]. This mode saves more energy than the way the node stays active all the time, but it introduces more delays in data collection. This is because in the duty cycle mechanism, when the sender has data to send, its forwarding node may be in a sleep state. Therefore, the sender needs to wait for its forwarding node to wake up before data can be routed, thus causing a large delay. Kim *et al.* [26] proposed an anycast forwarding and sleep wakeup scheduling strategy. The anycast strategy reduces the delay by each node opportunistically forwarding the packet to the first neighboring node that wakes up between multiple candidate nodes. In addition, the scheme optimizes network performance by jointly controlling the system parameters of

the sleep-wake scheduling protocol and the anycast packet forwarding protocol, such as the node wake-up rate, forwarding set, and priority. In competitive data collection, there are also communication interference and conflict when multiple nodes wake up at the same time.

In terms of data collection, in fact, most of the data collection strategies are based on competition mechanism. In many studies, sometimes not called data collection strategies, but called various data routing algorithms, such as cluster routing algorithm [32], shortest routing algorithm, its essence is to collect network data into the sink. Clustering network data collection can adopt either the TDMA mechanism mentioned above or the competition mechanism.

According to different data generation patterns, the data collection strategy can be divided into data collection under dynamic data patterns [7] and data collection under static data pattern [32]. In general, most of the data collection strategies based on competitive mechanism can be adapted to the network of dynamic data generation with little change. Because no matter how much data is generated by the data collection of the competition mechanism, its operation mode is the same, which is mainly competition. After successful competition, data will be sent. But the disadvantage of the competition mechanism is that when multiple nodes compete for one forwarding node, network congestion may be caused, thereby increasing energy consumption and delay, and thus a load balancing routing algorithm is required. However, the TDMA mechanism is completely different. In the static data generation network, because each node is determined before data collection, it is possible to arrange the state of each node in each slot in advance, so that it can be accurately scheduled. However, in the network with dynamic data generation, how many active slots each node needs to be arranged, and when to arrange the active slot of the node cannot be determined in advance. And since the number of packets is uncertain, the receiver can not even determine how many packets the sender needs to send and when the receiver has finished sending. Therefore, the dynamic data collection strategy of the TDMA mechanism faces a large challenge. Zhao and Tang [7] proposed a data collection strategy for dynamic traffic patterns. The main idea of their proposed strategy is to arrange the data transmission of a node in its continuous activation slot. The sender sleeps after sending its own data, and the last slot after the receiver receives the data continuously is empty, which means that the sender's data packets have been sent.

B. RESEARCH ON LIFETIME

The performance of different data collection strategies is different, among which the two most important ones are lifetime and delay. This section first discusses the situation of lifetime with different data collection strategies.

Because of the limited energy of sensor nodes, how to save energy and improve lifetime is the most critical factor in data collection strategies. Yetgin *et al.* [28] provided a broad overview of the definition of network lifecycle based on the application areas and design constraints of WSNs. After discussing the related network lifetime maximization techniques, Yetgin *et al.* [28] also provided some general design examples for maximizing the network lifetime of WSN to show potential improvements in different design standards. Therefore, from the point of view of the node, the energy consumption of the node should be minimized to extend its lifetime. It should be pointed out here that the lifetime of the network is not the same concept as the lifetime of the node [27]–[29]. This is because the lifetime of the network is defined as the time at which the first node in the network dies [19], [23], [24]. Therefore, to extend the lifetime of the network, it is necessary to maximize the lifetime of the node with the largest energy consumption in the network, so as to effectively improve the network lifetime. However, there is a unique phenomenon called "energy hole" in WSNs, which makes it particularly challenging to improve the network lifetime [24], [29]. Energy hole is such a phenomenon: Because in WSNs, the data of all nodes have to be routed to sink through multiple hops, which means that nodes within the r area of sink assume the relay of data of all nodes in the network, so their energy consumption is much higher than that of nodes in other areas, which is called hotspots [24]. As the energy consumption of hotspots is much higher than that of nodes in other areas, these nodes die prematurely, which affects the lifetime of the network [24]. Therefore, researchers realized that avoiding energy holes is the most important way to improve network lifetime. Thus, many strategies to avoid energy holes and improve network lifetime are proposed. The main ideas of these strategies are as follows: (a) The overall amount of data that needs to be transferred to the sink is reduced, thereby reducing the energy consumption of the hotspots nodes and improving the network lifetime. There are many such strategies, for example, data aggregation strategy can reduce the transmission of redundant data by utilizing the correlation between data [32], thus improving the network lifetime. There are also the recently developed compression-sensing strategies [22], as well as methods that use matrix completion techniques to reduce the amount of data and improve network lifetime [16], [23], [31]. In such a method, since there is a certain correlation between data, when the network lacks some data, it can be derived from the data of the adjacent location, thereby effectively reducing the collection of data to improve the network lifetime [16], [23], [31]. (b) A method of using different transmission radii. Different transmission radii can make the hotspots area of the network not in the same position, which can improve the network lifetime. Ju *et al.* [21] proposed the collection strategy of using a large sending radius when the energy absorbed by the external environment is sufficient, and using a small sending radius when the energy absorbed by the sensor node from the external environment is not sufficient, which can effectively reduce the energy consumption of hotspots area and improve the service life of the network. In addition, there are some joint design optimization schemes. Madan *et al.* [27] provided an iterative algorithm that alternates between link scheduling and computation of transmission powers and rates to maximize the lifetime of energy-constrained WSNs. Yetgin *et al.* [29] focused on cross-layer optimization of power allocation, scheduling, and routing operations, and proposed the optimal but excessive-complexity exhaustive search algorithm (ESA) and a near-optimal single objective genetic algorithm (SOGA) to maximize the network lifetime.

However, to the best of our knowledge, the current proposed strategies are to target the data collection strategy under the competition mechanism, but not the data collection strategy under the TDMA MAC protocol. Obviously, nodes with different transmission radii in TDMA-based data collection strategy can effectively improve the network performance. However, TDMA-based data collection algorithms face greater difficulties because mutual interference between multiple transmit radii makes the design very difficult [48], [49]. Moreover, it is more difficult to apply this strategy to the network generated dynamically by data, so there is no TDMA-based scheduling strategy of mixed sending radius. This paper is the first data collection strategy to attempt to use dynamic data generation networks based on TDMA with different transmission radii.

C. RESEARCH ON DELAY OPTIMIZATION

Delay research is also one of the most important research contents in WSNs [18], [19]. There are also many studies on reducing delay in wireless sensor networks. The delay for reducing data collection can be performed from multiple layers, such as the MAC protocol layer, the network routing layer, the application layer, and so on. In fact, wireless communication is used between nodes in a wireless sensor network. Because of the reliability of wireless communication, data transmission may fail. In order to ensure the reliability of data transmission, it is necessary to adopt a certain reliability guarantee method. The most commonly used protocol is send and wait automatic repeat-request (SW-ARQ) protocol. In the SW-ARQ protocol, Sender waits for the receiver to return an ACK to acknowledge the receipt of the packet after each packet is sent, and then starts the transmission of the next packet. If the ACK is not received until the timeout, the packet needs to be resent. It can be seen that the delay in wireless communication is relatively large.

From the network layer, the delay of data collection using competitive MAC protocols will be relatively large. Because in the competing MAC protocols, there are a large number of active nodes in the network, which leads to more interference between nodes, thus increasing delay in data retransmission, etc. Relatively speaking, the TDMA strategy is mainly applied to the occasion where the delay requirement is relatively high, and the delay of data collection is relatively small, but the time scheduling of data collection is strict.

III. THE SYSTEM MODEL AND PROBLEM STATEMENT

A. THE NETWORK MODEL

Similar to many studies [7], [18], [50], WSNs are organized into a tree structure with base station as the root for data collection. In a continuous data collection scenario, the sensor node samples local data in each sampling interval and then transmits the data to the base station. In addition to sending its own generated data, each node in the tree is also responsible for forwarding its child's data to its parent. Due to energy saving issues and the nature of monitoring applications, sensor nodes may not send data to the base station in each sampling interval. Therefore, the traffic patterns of data collection in the network exhibit unpredictable dynamic variability.

We consider that n homogenous sensor nodes and one sink node deployed in a WSN. The network can be regarded as $\{\aleph = s, v_1, v_2, v_3, v_4 \ldots v_n\}$, where s means the base station and v_i represents the i-th node. Those sensor nodes are randomly and uniformly deployed in the target area such as environment monitor, industrial field, smart field and so on [22]–[24]. The sensor nodes are powered by battery whose energy are limited, and the energy of the sink is unlimited. The network radius is R, the transmission radius of a sensor node is r. Each node, including sink node, has ω transmission power levels, and the transmission radius r of different transmission power levels is also different. In our research approach, the wireless sensor network can be converted into a complete k-ary routing tree of d levels, which can be represented in the Figure 1. Sink is the root of the tree, called level 0, and the other nodes in the tree are called level d according to the minimum hop counts to the sink. As shown in Figure 1, the left side is labeled with the level of the sensor nodes, and the right side is the maximum number of nodes per layer.

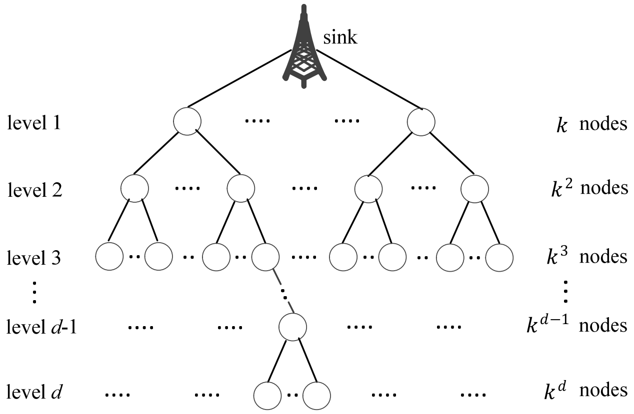


FIGURE 1. Network structure diagram.

The data sampling mode is: In each data sampling interval, each sensor node generates a data packet with a probability of p, and transmits its acquired data to the sink. For the sake of simplicity, it is assumed that the generation of data packets is independent on different sensor nodes and in different sampling intervals. The data collection approach adopts the Time Division Multiple Access (TDMA) mechanism, and the cycle T of each data collection is divided into $|T|$ time slots with the same length. Thus, one cycle can be expressed as $\{0, 1, 2, 3, \ldots, |T| - 1\}$ time slots. Each sensor node in the network has two states: activation and sleep. The TDMA mechanism assigns each wireless sensor network node independent time slots for data transmission. In other idle time slots, the node is in sleep state to save energy. Similar to other related studies [51]–[53], it is assumed that the clock is synchronized between sensor nodes. In the study of this paper, one slot is sufficient to complete the data acquisition and transmission of one hop irrespective of the transmission radius, and the data acquired in each sampling interval is suitable for the size of one data packet. Data packets generated by different nodes will not be aggregated on the path transmitted to the sink (the aggregation of the packets can be achieved by the extended algorithm, but it is limited by the size of some packets).

B. THE ENERGY CONSUMPTION MODEL

In this paper, a typical energy consumption model [14] is used. The energy consumption of the transmitted data is shown in equation (1), and the energy consumption of the received data is shown in equation (2).

$$\begin{cases} E_{send} = lE_{elec} + l\varepsilon_{fs}r^2 & if \;\; r < r_0 \\ E_{send} = lE_{elec} + l\varepsilon_{amp}r^4 & if \;\; r > r_0 \end{cases} \tag{1}$$

$$E_{receive} = lE_{elec} \tag{2}$$

In the equation, E_{elec} represents the energy dissipated by the transmitting circuit. If the transmission distance is less than the threshold value r_0, the power amplification loss adopts a free space model. If the transmission distance is greater than or equal to the threshold value r_0, the multipath attenuation model is adopted. ε_{fs}, ε_{amp} are the energy required for power amplification in the two models, respectively, and l represents the number of bits of data. In this paper, the specific settings of the above parameters are taken from the literature [14], as shown in Table 1.

TABLE 1. Network parameters.

Parameter	Value
Threshold distance (r_0) (m)	87
Sensing range (r) (m)	≤80
Eelec (nJ/bit)	50
ε_{fs}(pJ/bit/m^2)	10
ε_{amp}(pJ/bit/m^4)	0.0013
Initial energy(E_{init})(J)	0.5
l (bit)	16

C. THE CONFLICT HANDING MODEL

In order to avoid interference in the data collection process, the data reception and transmission of the node cannot be performed in the same time slot, and only non-conflicting nodes are allowed to perform data transmission in the same time slot. For any two nodes v_i and v_j in the network, when the formula (3) is satisfied, it can be ensured that the respective transmissions of v_i and v_j do not conflict.

$$\begin{cases} |v_i - v_j| > r_i + r_j & if \;\; l_{v_i} \neq l_{v_j} \\ v_i \text{ and } v_j \text{ are not brothers} & if \;\; l_{v_i} = l_{v_j} \end{cases} \tag{3}$$

where $|v_i - v_j|$ represents the shortest distance between nodes v_i and v_j, r_i and r_j represent the transmission radius of node i and node j, respectively, l_{v_i} and l_{v_j} represent the level where v_i and v_j are located, respectively. For wireless sensor networks where the transmission distance of each node is r, the transmission of node i and its sibling nodes, parent node and grandfather node are in conflict with each other. That is, nodes that are separated by two levels can transmit data in the same time slot. We assume that before each cycle of scheduling, each node knows about other nodes that are in conflict with it. This information can be obtained by using practical methods such as RID [50].

D. PROBLEM STATEMENT

As mentioned earlier, since TDMA-based strategies are more energy efficient than competition-based strategies, and the TDMA mechanism can avoid conflicts that are simultaneously awakened by multiple nodes in the contention mechanism by allocating time slots. This paper attempts to design an effective time-slot scheduling strategy based on TDMA protocol with different transmission radii, which is suitable for dynamic data generation networks. The ultimate goal is to reduce delays in data collection and improve the life cycle of the network. The specific discussion is as follows:

Definition 1: Minimization of network data collection delay (denoted as D). The network data collection delay D refers to the time when data packets (if any) of all nodes in the network are transmitted to the sink. The aim of this paper is to reduce the delay of network data collection. The time required for the data generated by node v_i to arrive at sink is denoted by d_i, it can be calculated as follows.

$$\min(D) = \min(\max_{1 \leq i \leq n} (d_i)) \tag{4}$$

Definition 2: Maximization of network lifetime (denoted as ℓ). The life cycle of a network can be defined as the time from the start of data collection to the death of the first node. Affected by the death node, the data of some nodes could not be transmitted to the sink, which damaged the function of the whole network and prevented it from working normally. In this paper, the energy consumption e^i of the node v_i is mainly composed of three parts: the energy consumption e_r^i of the received data packet, the energy consumption e_s^i of the transmitted data packet, and the energy consumption e_l^i of the idle monitoring data. The initial energy of node v_i is denoted by E_{init}. Thus, the lifetime of node v_i in the network is $E_{init} \big/ (e_r^i + e_s^i + e_l^i)$. While maximizing the life of the network is to maximize the node with the least lifetime in the network, and it can be expressed as follows.

$$\max(\ell) = \max\left(\min_{1\le i\le n}\left(E_{init} \Big/ (e_r^i + e_s^i + e_l^i)\right)\right) \tag{5}$$

Definition 3: Maximization of effective energy utilization rate (denoted as φ). φ refers to the ratio of the energy consumed by the sensor network at the end of its life cycle to its initial energy. The aim in this paper is to make full use of energy left in nodes far from the sink, then maximize energy utilization of the whole network. Thus, the effective energy utilization rate is as follows.

$$\max(\varphi) = \max\left(\left(\sum\nolimits_{1\le i\le n} e^i\right) / \left(\sum\nolimits_{1\le i\le n} E_{init}\right)\right) \tag{6}$$

Obviously, the goal of data collection scheduling strategy is to minimize delay, D, maximize the network life, ℓ, and effective energy utilization, φ, which can be summarized as follows.

$$\begin{cases} \min(D) = \min\left(\max_{1\le i\le n}(d_i)\right) \\ \max(\ell) = \max\left(\min_{1\le i\le n}\left(E_{init} / \left(e_r^i + e_s^i + e_l^i\right)\right)\right) \\ \max(\varphi) = \max\left(\left(\sum_{1\le i\le n} e^i\right) / \left(\sum_{1\le i\le n} E_{init}\right)\right) \end{cases} \tag{7}$$

For the convenience of readers, notations used in the paper are listed in Table 2, where the columns "Notation" contains the names of variables, and their definitions are presented in the columns "Definition".

TABLE 2. Notations used in the paper.

Notation	Definition
k	The degree of the tree network.
d	The level of the tree network.
p	The probability that a node will generate one packet.
r	The transmission radius of a node.
E_{init}	Initial energy of the node.
E_s	The energy consumption of one packet sent by a node with transmission distance of mr.
E_r	The energy consumption of one packet received by a node with transmission distance of mr.
E_l	The energy consumption for idle listening in one time slot by a node with transmission distance of mr.
e_s	The energy consumption of one packet sent by a node with transmission distance of r.
e_r	The energy consumption of one packet received by a node with transmission distance of r.
e_l	The energy consumption for idle listening in one time slot by a node with transmission distance of r.
e_{re}^i	Residual energy of each node at level i.
Δ_i	Subtree size of mr transmission distance rooted at node of level i.
δ^i	The energy consumption of each node in the i-th layer in a life cycle.
ℓ	Life cycle of the network.
φ	Energy efficiency rate of the network.

IV. THE DESIGN OF RDML SCHEME

A. RESEARCH MOTIVATION

In the previous data collection scheduling, the sending radius adopted by each node is r. Therefore, data transmission only occurs between two adjacent levels of the network, that is to say, nodes at level d will only send data packets to nodes at level $d-1$. When the node has data to send, the data transmission of the node is performed as much as possible under the premise that the data transmission does not cause conflict, so that the delay required for data collection is minimized. However, we have observed that if some nodes in the network have a transmission radius of m times r, that is, when the transmission radius is mr, the data of the level d nodes can be transmitted to the nodes of level $d-m$ in one slot only by one hop. Therefore, the distance of one-hop data transmission is increased, the time required for data collection of the entire network is reduced, and the lifetime and energy utilization rate is improved.

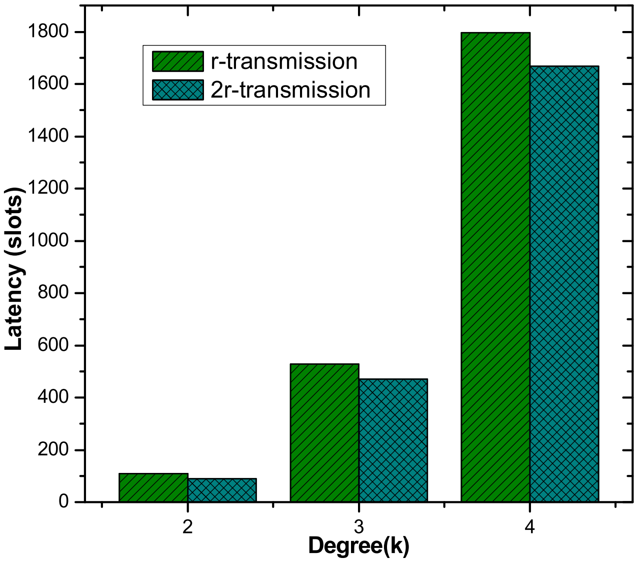


FIGURE 2. r-transmission and $2r$-transmission network delay under $d = 5$, $r = 10$.

We first use a specific experiment to illustrate that when the sending radius of some nodes in the network is mr, the time required for data collection can be effectively reduced. In this paper, we use the network with the sending radius r as the basic network. When the maximum sending radius of the nodes in the network is mr, we call this mr-transmission network. In this way, when $m = 1$, it is the network used in the previous scheduling strategy, which is called r-transmission network. When $m = 2$, it is called $2r$-transmission network. Figure 2 is a comparison of the time required for a cycle of data collection after the transmission radius of the non-hotspots nodes are increased to $2r$, and the time required for the data collection by all the nodes in the previous strategy adopting r. As can be seen from the experimental results, when the transmission radius of some nodes in the network increases to $2r$, the time required for the network to complete a cycle of data collection is significantly reduced.

However, the energy consumption of the node is related to the transmission radius of the node in the power of 2 or even 4th. Since the energy of the sensor nodes is very limited, increasing the transmission radius of the node to mr will increase the energy consumption, which may cause the network lifetime to decrease. However, after the following analysis, we can completely increase the transmission radius of some nodes in the network to mr without affecting the network lifetime.

In WSNs, since sink is the data collection center of the entire network, when the node adopts sending radius r, all data packets will be forwarded to sink through the node within the area of sink r. Thus, the amount of data carried by nodes in the area of sink r is much larger than that of nodes in other areas. These nodes are also called hotspots nodes. As a result, the hotspots nodes often die first, resulting in the end of the life cycle of the entire network. While non-hotspots nodes have a large amount of energy surplus, according to [12], due to the influence of hotspots, when the network dies, the remaining energy in the network is as high as 80%. In theory, if we only increase the transmit power of the non-hotspots nodes for mr transmission, we can share the amount of data that the hotspots nodes bear. Instead of reducing lifetime, this will improve the lifetime of the entire network. At the same time, the mr transmission strategy also reduces the number of hops that packets are sent to the sink, thereby reducing the delay in data collection.

In order to verify the effectiveness of the strategy in this paper, we use the network as shown in Figure 1, through a specific example to analyze and illustrate that the wireless sensor network has a large amount of residual energy, which can extend the life cycle of the network on the premise of increasing the transmission radius of some nodes to mr. Considering a worst case scenario, each node in the tree produces one packet (ie, $p = 1$) during a cycle of data collection. Then, the node sends packets to sink continuously in its transmission slot, so there is no energy consumption for idle listening. For a r-transmission network, each node of level d has only the energy consumption of sending packets. Each node from level $d-1$ to level 1 has both the energy consumption of sending and receiving data packets. In a data collection cycle, the number of packets received by each node at level $d-1$ is k, at level $d-2$ is $k+k^2$, and so on, and finally at level 1 is $k+k^2+\ldots+k^{d-1}$, ie $\sum_{n=1}^{d-1} k^n$. Since the node itself also generates one data packet, the number of packets to be sent by each node is one more than the number of packets to be received. Thus, the amount of data packets carried by each node at level $i(1 \leq i \leq d)$ can be expressed as $\sum_{n=0}^{d-i} k^n$.

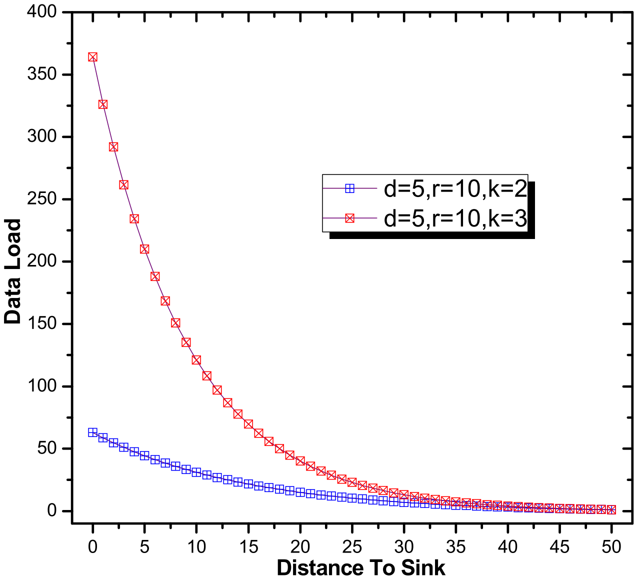


FIGURE 3. The amount of data transmitted by the node increases as the distance from the sink node decreases.

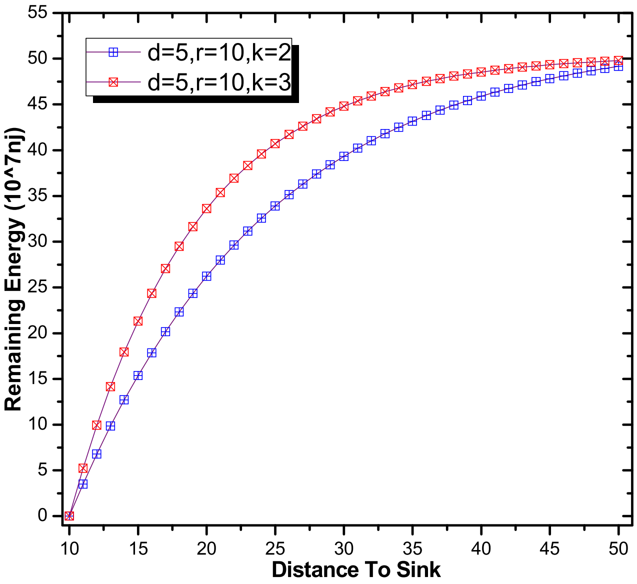


FIGURE 4. The remaining energy of the nodes at different distances from the sink in the network.

Then combined with equation (1) and (2), the energy consumption of each node at level 1 during a data collection cycle is $e_r \cdot \sum_{n=1}^{d-1} k^n + e_s \cdot \sum_{n=0}^{d-1} k^n$. Obviously, each node of level 1 belongs to hotspots nodes, so when the node of level 1 dies, the life cycle of the entire network is over. Thus, the life cycle of the network can be expressed as

$$\ell = E_{init}/(e_r \cdot \sum_{n=1}^{d-1} k^n + e_s \cdot \sum_{n=0}^{d-1} k^n).$$

Thus, the remaining energy e_{re}^i of each node at level i in the network can be expressed as

$$e_{re}^i = E_{init} - ((e_r + e_s) \cdot \sum_{n=0}^{d-i} k^n - e_r) \cdot E_{init}/(e_r \cdot \sum_{n=1}^{d-1} k^n + e_s \cdot \sum_{n=0}^{d-1} k^n).$$

We take the tree's level d to be 5, the transmission distance r to be 10, and k to be 2 and 3, respectively. The values of the other parameters are taken from Table 1, and the following Figure 3 and Figure 4 are obtained. Figure 3 shows the amount of data that each node in the near-sink area and the far-sink area bears. It can be seen that the near-sink area node bears much more data than the far sink area. Figure 4 shows the remaining energy in different areas of the network. It can be seen that there is a lot of energy remaining in the far sink area. Therefore, we can use this part of the remaining energy to increase the transmission radius of some nodes to mr.

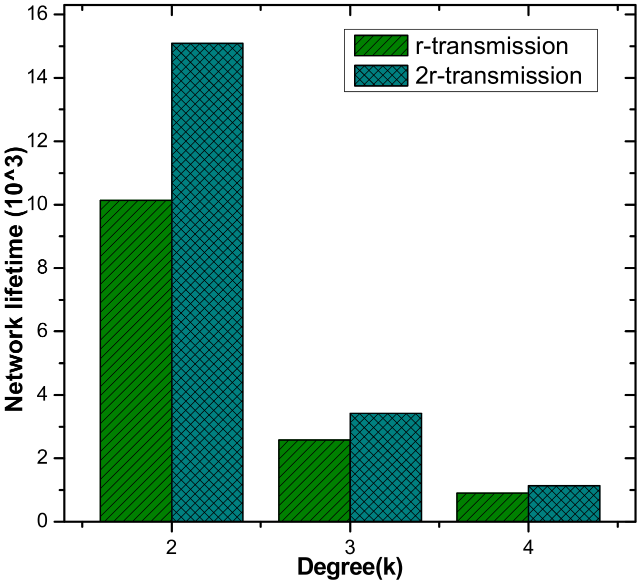


FIGURE 5. r-transmission and $2r$-transmission network lifetime under $d = 5, r = 10$.

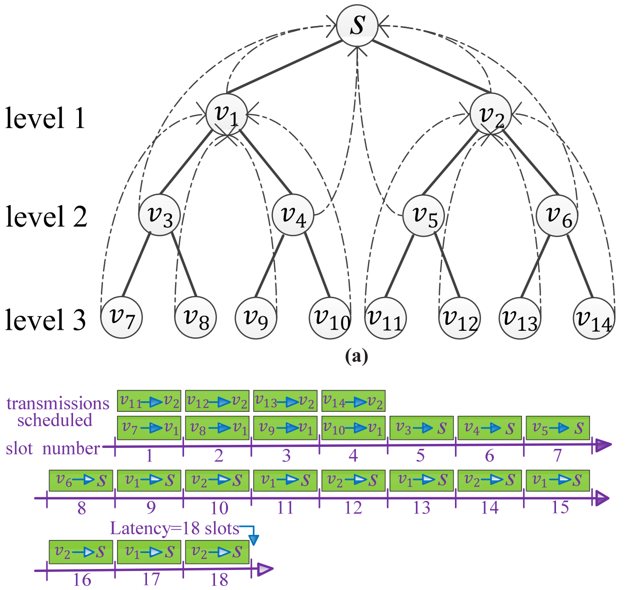
Not only that, the new scheduling strategy can significantly improve the network lifetime. In the original scheduling strategy, the transmission radius of all nodes in the network is r. Therefore, all data in the network must be sent to the sink through the nodes in the area of sink r. That is to say, the nodes in the network near the sink r area bear the data forwarding of all nodes in the network, so the energy consumption is very heavy, and its lifetime determines the life of the entire network. When the transmission radius of the non-hotspots node in the network increases to $2r$, the data of the nodes in the area of sink $2r$ can be directly sent to the sink, so that the amount of data undertaken by the hotspots nodes is greatly reduced. This can effectively improve the lifetime of the entire network. Figure 5 is a comparison of the network life cycle that increases the transmission distance of non-hotspots nodes to $2r$ and all nodes are transmitted at a distance of r. It can be seen that when k is taken as 2, 3 and 4, the life cycle of $2r$-transmission network is longer than that of r-transmission network. Among them, when k is taken as 2, the life cycle is increased by 48.77%. This shows that the data collection strategy of making the transmission radius of some nodes as mr proposed in this paper can effectively improve the network performance.

Although the previous examples show that the proposed strategy can effectively improve network performance, the effective data collection strategy in the sensor network of dynamic data generation still faces some significant challenges. The major challenge is how to optimize the active slot in each node's data collection process, so that the number of active slots required by the node is as small as possible to save energy. At the same time, the data transmission slot of the node is optimally arranged so that as many nodes as possible can transmit data in parallel without interference, so as to minimize the delay of data collection. In the strategy of this paper, the data transmission radius of some nodes is mr, when the node sends data, its interference range is m times larger than the original strategy, so its scheduling is more complex and more challenging. Furthermore, this paper is aimed at a network of dynamic data generation, that is, whether a node generates data in a data collection cycle is uncertain. This is very different from previous studies, in which each node produces a fixed data packet in a cycle of data collection. Therefore, slot allocation is more complex. Finally, according to the energy consumption of the nodes in the network, it is also a certain challenge to select which nodes with energy remaining to adopt mr distance for data transmission.

B. AN EXAMPLE OF THE PROPOSED SCHEME

From the previous discussion, the new data collection strategy proposed in this paper is very complicated and challenging. In this section, we first illustrate our data collection strategy through a specific time slot scheduling example. Since the data generation of nodes is uncertain, the number of packets that the same node needs to receive and send will be greatly different in different data collection cycles, which means that the number of slots that nodes need to send and receive is also greatly different. Therefore, the problem of scheduling strategy is how to effectively reduce the unnecessary idle listening in data collection under dynamic traffic patterns. It is better to make each node's data packets can be transmitted continuously, so as to effectively reduce the time and energy consumption required for data collection.

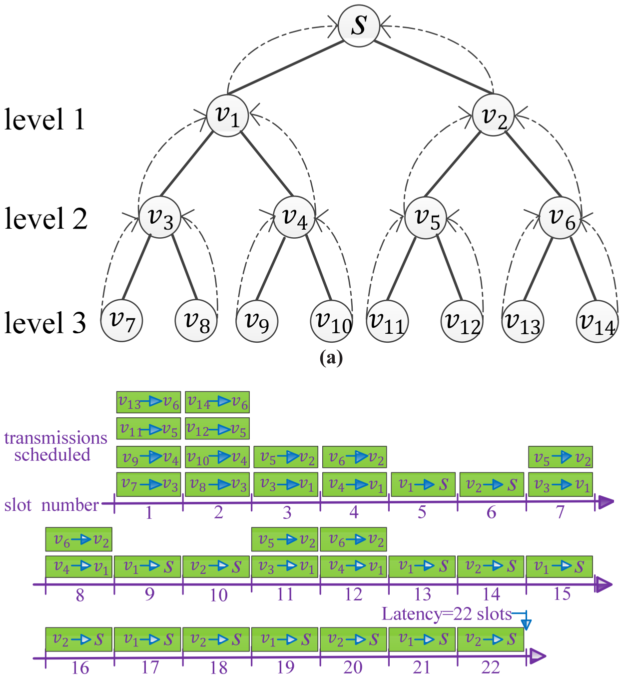
The general idea of our proposed strategy is to allow each node to transmit all packets (if any) in its consecutive transmission slots starting from its first transmission slot in a data collection cycle. If the node does not send a data packet in the scheduled transmission time slot, indicating that the node has no data packets to be transmitted or all the data packets have been transmitted, then the data collection of the node is completed. The principle of scheduling is that the nodes below level m (including level m) can transmit mr distance in one hop, and the nodes above level m can send data packets to sink by one hop, and multiple rounds of data collection are required for a cycle. A round of data collection means that from the bottom of the network, each node makes one data transmission to its next hop node in one time slot (nodes that do not conflict can be transmitted in the same time slot). Then, the data transmission of the upper level node is performed until the node that is r distance away from the sink performs one data transmission to the sink. Thus, each node at level $i (1 \le i \le d)$ collects at most the subtree size rounds of the mr transmission distance rooted at itself. If the node does not transmit one data packet in its activation slot or reaches its maximum data collection rounds, the node's data collection is completed. And during the data collection, nodes of non-conflicting levels can be transmitted concurrently. In this way, multiple rounds of collection are performed. Finally, if the node at level 1 does not receive the data packet in its activation time slot or reaches its maximum collection rounds, the data collection of the network in this cycle is completed.



(b)

FIGURE 6. **(a) $2r$-transmission network nodes transmission topology. (b) The entire $2r$-transmission network schedule when each node generates one data packet.**

We start with an example to illustrate the scheduling process of our proposed strategy. Figure 6(a) shows the topology diagram that each node can transmit data packets to its next hop node (indicated by the dotted line with arrows in the figure) when the binary tree of level 3 adopts $2r$ transmission distance. When each node generates one data packet (ie, $p = 1$), Figure 6(b) shows the specific process of time slot scheduling in the $2r$-transmission network. According to our scheduling principle, the maximum number of data collection rounds for each node at level 3, 2, 1 is 1, 1, 5, respectively. Packets of nodes v_7, v_8, v_9 and v_{10} can be directly transmitted to node v_1 by one hop. Similarly, packets of nodes v_{11}, v_{12}, v_{13} and v_{14} are directly transmitted to node v_2. Since the transmissions of the sibling nodes collide with each other, in the first round of transmission, nodes v_7 and v_{11} end the data collection after performing one data transmission in the slot 1. Similarly, nodes v_8 and v_{12}, v_9 and v_{13}, v_{10} and v_{14} end data collection after one data transmission in time slots 2, 3, 4, respectively. Nodes v_3, v_4, v_5 and v_6 end the data collection after one data transmission in slot 5, 6, 7 and 8. Then according to this schedule, v_1 and v_2 are assigned to slots 9 and 10. Thus, after the first round of data collection, the bottom level of the network becomes 1. In the second round of transmission, slot 11 is assigned to node v_1 and slot 12 is assigned to node v_2. Similarly, in the third round of transmission, nodes v_1 and v_2 are assigned to slots 13 and 14, respectively. Thus, by the fifth round of transmission, nodes v_1 and v_2 are assigned to time slots 17 and 18, respectively, reaching the maximum number of data collection rounds for this level nodes, and then the node goes to sleep and ends the data collection. At this time, the data collection of the entire network is completed, which takes 18 time slots and has no idle listening.



(b)

FIGURE 7. **(a) r-transmission network nodes transmission topology. (b) The entire r-transmission network schedule when each node generates one data packet.**

Figure 7(a) is the topology diagram that each node in the r-transmission network can transmit data packets to its next hop node. It can be seen that the next hop node of each node is the parent node of its binary tree. When $p = 1$, Figure 7(b) shows the specific scheduling process of data transmission. The maximum number of data collection rounds for each level 3, 2, 1 node is 1, 3, 7, respectively. The first round of scheduling begins with level 3, and nodes v_7, v_9, v_{11} and v_{13} end data collection after simultaneous data transmission in time slot 1, and nodes v_8, v_{10}, v_{12} and v_{14} simultaneously terminate data collection after transmitting one data packet in time slot 2. Thus, the nodes v_3 and v_5, v_4 and v_6 of level 2 perform one data transmission in slots 3 and 4, respectively. Then, to level 1, v_1 and v_2 perform one data transmission in slots 5 and 6, respectively. The second round starts from level 2, v_3 and v_5, v_4 and v_6 perform one data transmission in slots 7 and 8, respectively, and v_1 and v_2 of level 1 perform one data transmission in slots 9 and 10, respectively. By analogy, finally, the data collection of the network is completed at the end of time slot 22, and there is no idle listening.

TABLE 3. Number of rounds for *r*-transmission and 2*r*-transmission network schemes when *p* = 1.

Round	End slot (r-transmission network)	End slot ($2r$-transmission network)
1	6	10
2	10	12
3	14	14
4	16	16
5	18	18
6	20	
7	22	

Table 3 lists the number of rounds required for the two schemes to complete data collection in one cycle and the time slots at the end of each round. By comparison, it can be found that when each node in the three-layer binary tree network generates one data packet, both schemes have no idle listening. However, the data collection of $2r$-transmission network scheme is 2 less rounds and 4 less time slots than that of r-transmission network scheme.

Since this paper studies the wireless sensor network generated by dynamic data. During a data collection cycle, the network traffic pattern changes abruptly, and each node in the network generates a packet to be sent with a probability of 0.5. The set of nodes that actually generate the data packet is $\{v_1, v_2, v_3, v_5, v_6, v_9, v_{12}\}$. The specific processes of the $2r$-transmission network and r-transmission network scheduling are given in Figure 8(a) and Figure 8(b), respectively. In the figure, the actual transmission only occurs in the green rectangle, and the yellow rectangle indicates that the receiving node has listening, but actually the sending node has no data transmission. The gray rectangle indicates that the receiving node is not listening, and the sending node has no data transmission, that is, the node has finished data collection and is in a sleep state.

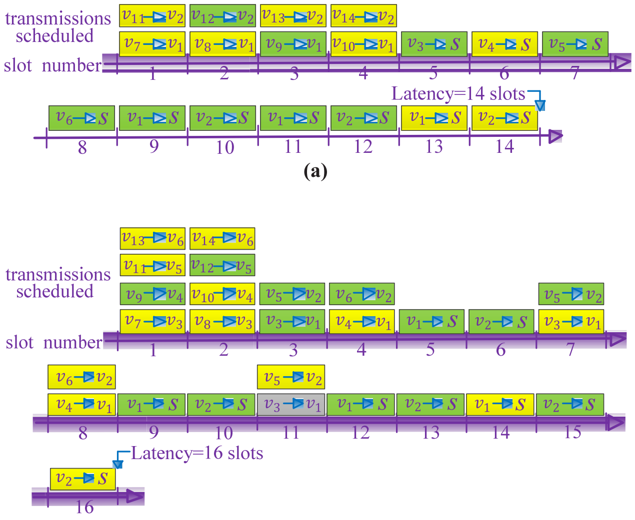


FIGURE 8. (a) The entire 2*r*-transmission network schedule with dynamic traffic patterns. (b) The entire *r*-transmission network schedule with dynamic traffic patterns.

We illustrate the $2r$-transmission network scheduling process for dynamic data generation in conjunction with Figure 8(a). Similar to Figure 6(b), in the first round of scheduling, time slot 1 is assigned to nodes v_7 and v_{11}, and time slots 2, 3, 4 are assigned to v_8 and v_{12}, v_9 and v_{13}, v_{10} and v_{14}, respectively. Then nodes v_3, v_4, v_5 and v_6 are assigned to time slots 5, 6, 7 and 8, and v_1 and v_2 to time slots 9 and 10, respectively. In the second round of scheduling, the time slot 11 is assigned to the node v_1, and the transmission of the node v_2 is performed in the time slot 12. In the third round of scheduling, nodes v_1 and v_2 are assigned to slots 13 and 14, respectively. Since there are no data packet transmissions in slots 13 and 14, the nodes v_1 and v_2 end their data collection, and the network collection of data in this cycle is completed. Although sink has completed the collection of all packets at the end of the second round (end of time slot 12), it has to continue to listen for possible transmissions in slots 13 and 14. The reason is that sink can not know the situation of packets generated by each node in the network in advance. However, in time slots 11 and 12, node v_1 and v_2 still send data to sink, and the maximum number of collection rounds is not reached. In order to prevent any possible loss of data packets, sink has to continue listening to the next transmission slots 13 and 14 of nodes v_1 and v_2 to sink. In time slots 13 and 14, sink does not receive the packets sent by v_1 and v_2, at which time the data collection of the network is completed. Therefore, the latency of the entire network data collection is 14 time slots.

TABLE 4. Number of rounds for *r*-transmission and 2*r*-transmission network schemes when *p* = 0.5.

Round	End slot (r-transmission network)	End slot ($2r$-transmission network)
1	6	10
2	10	12
3	13	14
4	15	
5	16	

Table 4 shows the number of rounds required for two schedules to collect data in one cycle and the time slots at the end of each round. The comparison shows that when $p = 0.5$, adopting the $2r$-transmission network scheduling to complete data collection requires 14 time slots, 9 idle listening and 3 rounds. Compared with r-transmission network, the total data collection time slot is 2 less, the number of idle listening is 4 less, and the number of required rounds is 2 less.

C. OUR PROPOSED SCHEDULING ALGORITHM RDML

In view of this, we propose the RDML scheduling algorithm (Reducing Delay and Maximizing Lifetime) that increases the transmission radius of some nodes to mr. We record the next hop node of node v transmission data as P_v, and assume that each sensor node v knows its next hop P_v. In general, such information can be easily obtained at each node after the routing tree is constructed. In addition, each node v also knows the set $I(v)$ of other nodes that conflict with it, and level $i(1 \leq i \leq d)$ also knows the set $I(i)$ of levels in conflict with it. This can be obtained by having the sensor node broadcast [54] to its neighbors or using a method such as RID [50].

Algorithm 1 RDML Scheduling Algorithm

```
1: let $N$ be the list of all nodes in a post-order traversal;
2: For each node $v$ in $N$ Do
3:   $T(v) = \emptyset$;
4:   $Tag(v) = 1$;
5:   $R_v = 0$;
6: End for
7: $depth = d$;
8: While $d > 0$ Do
9: For each node $v$at the level $d$ Do
10:   If $Tag(v) = 1, \notin v_i I(v_j)$ and$\notin v_j I(v_i)$ then
11:     simultaneous transmission of $v_i$ and $v_j$;
12:     $t++$;
13:     $T(v_i) = T(v_i) \cup \{t\}, T(v_j) = T(v_j) \cup \{t\}$;
14:     $T_{v_i}++; T_{v_j}++$;
15:   End if
16:   If $P_v$ receives no data from $v$or$R_v = \Delta_v$ then
17:     $Tag(v) = 0$;
18:   End if
19:  End for
20:   Ifeach node $v$ at level $d$ has $Tag(v) = 0$ then
21:     $depth--$;
22:   End if
23:  $d--$;
24: End while
25: If $depth = 0$ and each node $v$ in $N$ has $Tag(v) = 0$ then
26:   transmission finished;
27: End if
28: Else
29:  $d = depth$;
30: If $\notin dI(i)$ then
31:  simultaneous transmission of level $d$ and $i$;
32: End if
33: Goto line 8;
```

Algorithm 1 gives the pseudocode of the RDML scheduling algorithm. The algorithm uses post-order traversal, and each round of scheduling allocates time slots from the bottom level to the top level. The meaning of each parameter is as follows: $T(v)$ records the time slot allocated to node v. $Tag(v)$ is the state value of node v. Taking 1 means that the data collection of the node has not been completed, and taking 0 means that the node has finished data collection. d represents the level of the tree, and R_v represents the actual number of data collection rounds of the node v.Δ_v represents the subtree size of the mr transmission distance with the node v as the root, that is, the maximum number of transmission rounds of the node v. v_i and v_j are any two nodes of level d. t is the transmission time slot, and the initial value is 0. $depth$ is the maximum level of nodes in the network that have not completed data collection.

In the RDML scheduling process, $T(v)$ is used to record the transmission time slot of the node, initially $\emptyset$. All nodes do not start data collection at the initial stage, and the actual data collection round R_v is 0 (steps 2-6). Each round of data collection starts with the node at the lowest level level d. If one node is not in the conflict set of another node, they can transmit data in the same time slot (steps 9-15). If the node does not transmit a packet to its next hop node in the transmission slot, or has reached its maximum number of collection rounds, the node's data collection is complete (steps 16-18). If all nodes of level d have completed data collection, the maximum level of the nodes that have not completed data collection is updated to $depth - 1$ (steps 20-22). Then continue to collect the nodes of level $d-1$, $d-2$, …, 1 level by level. When d is 0, the current round of data collection is completed (steps 8-24). Then, the next round of data collection begins with the maximum level $depth$ of the actual unfinished data collection nodes (steps 29). If d and i are levels that do not conflict with each other, they can be transmitted in parallel (steps 30-32). This is performed in multiple rounds until the $depth$ is 0 and all nodes in N have finished the data collection, the network data for this cycle are all collected (steps 25-27).

V. PERFORMANCE ANALYSIS

We analyze the performance of the RDML algorithm by increasing the transmission radius of some nodes to $2r$. In fact, the analysis of increasing the transmission radius of some nodes to mr is similar. We analyze and compare the theoretical performance of the RDML algorithm with the advanced TPO algorithm. In the TPO algorithm, all nodes adopt the same transmission radius r for data collection, which is an efficient data collection scheme for dynamic data generation networks.

A. ANALYSIS OF ENERGY CONSUMPTION

In our analysis, we focus on the energy consumed by sensor nodes in transmitting data, receiving data, and idle listening. To simplify presentation, we shall denote the size of the $2r$ transmission distance subtree rooted at a node located at level i of the k-ary tree by Δ_i. E_s, E_r and E_l represent the energy consumption of sending one data packet, receiving one data packet and one idle listening at a node with a hop transmission distance of $2r$, respectively. d represents the level where the leaf nodes are located, for each $1 \leq i \leq d$, there is

$$\Delta_i = \sum\nolimits_{j=0}^{\left|\frac{d-i}{2}\right|} k^{2\cdot j} \tag{8}$$

Theorem 1: In a data collection cycle, the expected energy consumption of each node at the i-th($1 \leq i \leq d$) level of the RDML algorithm is

$$\delta_{RDML}^{i} = \begin{cases} p \cdot E_s,\ \ if\ i = d\ or\ i = d-1 \\ p \cdot ((\Delta_i - 1) \cdot E_r + \Delta_i \cdot E_s) + E_l \\ \qquad \cdot k^2 \cdot \left(1 - p^{\Delta_{i+2}}\right),\ \ if\ 2 \leq i \leq d-2 \\ p \cdot ((\Delta_1 - 1) \cdot E_r + \Delta_1 \cdot e_s) + E_l \\ \qquad \cdot k^2 \cdot \left(1 - p^{\Delta_3}\right),\ \ if\ i = 1 \end{cases} \tag{9}$$

Proof: For the RDML algorithm, during a data collection cycle, the nodes of level d and level $d-1$ in the k-ary routing tree simply send the generated packets (if any) to their next hop parents without listening to any other nodes. Therefore, the expected energy consumption of each node at level d and level $d-1$ is $p \cdot E_s$.

Considering each internal node v of level $i(2 \leq i \leq d-2)$, v receives and sends data packets by $2r$ distance, where the number of data packets to be transmitted is $p \cdot \Delta_i$. Therefore, for node v, the energy consumption of transmitting and receiving data is $p \cdot ((\Delta_i - 1) \cdot E_r + \Delta_i \cdot E_s)$. In addition, node v also listens to one idle time slot of each child node unless the child node sends data packets in all of its transmission time slots, which occurs when each node in the $2r$ transmission distance subtree rooted at the child node generates one packet. Therefore, the probability that node v performs idle listening for one slot of a child is $1 - p^{\Delta_{i+2}}$. Since the number of last hop nodes of node v has k^2, the expected energy consumption of v for idle listening is $E_l \cdot k^2 \cdot (1 - p^{\Delta_{i+2}})$. Thus, the expected total energy consumption of node v is $p \cdot ((\Delta_i - 1) \cdot E_r + \Delta_i \cdot E_s) + E_l \cdot k^2 \cdot (1 - p^{\Delta_{i+2}})$.

Since each node of level 1 can send data packets to sink through r distance, the energy consumption of sending data packets is $p \cdot \Delta_1 \cdot e_s$. Therefore, the expected energy consumption of each level 1 node is $p \cdot ((\Delta_1 - 1) \cdot E_r + \Delta_1 \cdot e_s) + E_l \cdot k^2 \cdot (1 - p^{\Delta_3})$. ■

Theorem 2: In a data collection cycle, the ratio ψ_{en}^{i} of the proposed algorithm to the expected energy consumption of each node at the i-th ($1 \leq i \leq d$) level of the previous TPO algorithm is $\psi_{\text{en}}^{i} = \delta_{RDML}^{i} / \delta_{TPO}^{i}$, where

$$\delta_{TPO}^{i} = \begin{cases} p \cdot e_s, \quad if\ i = d \\ p \cdot \left(\sum_{j=1}^{d-i} k^j \cdot e_r + \sum_{j=0}^{d-i} k^j \cdot e_s\right) + \mathrm{e}_l \cdot k \\ \cdot \left(1 - p^{\sum_{k=j}^{d-i-1} k^j}\right),\ \ if\ 1 \leq i \leq d-1 \end{cases} \tag{10}$$

Proof: For TPO scheduling, the leaf nodes of level d in one cycle only transmit data without receiving and listening, so the energy consumption of the level d node is $p \cdot e_s$.

The energy consumption of each internal node v of other level $i(1 \leq i \leq d-1)$ is composed of three parts: transmission, reception and listening. The size of the r transmission distance subtree rooted at v is $p \cdot \sum_{j=0}^{d-i} k^j$, so the energy consumed by receiving and sending data packets is $p \cdot (\sum_{j=1}^{d-i} k^j \cdot e_r + \sum_{j=0}^{d-i} k^j \cdot e_s)$. Similarly, node v also needs to listen to one idle time slot of each child to ensure the integrity of data transmission. The probability of this occurrence is $1 - p^{\sum_{j=0}^{d-i-1} k^j}$. And each node v has k child nodes, so the expected energy consumption of v for idle listening is $\mathrm{e}_l \cdot k \cdot (1 - p^{\sum_{j=0}^{d-i-1} k^j})$. Therefore, the expected total energy consumption of node v is $p \cdot \left(\sum_{j=1}^{d-i} k^j \cdot e_r + \sum_{j=0}^{d-i} k^j \cdot e_s\right) + \mathrm{e}_l \cdot k \cdot (1 - p^{\sum_{j=0}^{d-i-1} k^j})$. Finally, the ratio ψ_{en}^{i} of the expected energy consumption of each node at level $i(1 \leq i \leq d)$ of the RDML and the TPO is given by formula (10). ■

B. ANALYSIS OF NETWORK LIFETIME

As mentioned earlier, the life cycle of a sensor network is mainly determined by the node in the network that consumes the most energy. In the study of this paper, each node has the same initial energy, so the node with the most energy consumption is the first to die. Then the data transmission of other nodes will be interrupted by the influence of dead nodes, which leads to sink unable to complete the data collection of the entire network. Therefore, when calculating the life cycle of the network, it is to calculate the lifetime of the node with the highest energy consumption in the network.

Theorem 3: A sensor network that adopts RDML scheme for data collection, whose life cycle can be expressed as

$$\begin{aligned} \ell_{RDML} = E_{init} / \max\{& p \cdot ((\Delta_2 - 1) \cdot E_r + \Delta_2 \cdot E_s) \\ & + E_l \cdot k^2 \cdot (1 - p^{\Delta_4}), p \cdot ((\Delta_1 - 1) \cdot E_r + \Delta_1 \cdot e_s) \\ & + E_l \cdot k^2 \cdot (1 - p^{\Delta_3})\} \end{aligned} \tag{11}$$

Proof: Since RDML scheduling algorithm transmits data packets from bottom to top by $2r$ distance. Therefore, all data packets in the network are either transmitted to sink by $2r$ distance of level 2 nodes or to sink by r distance of level 1 nodes. Thus, the nodes that consume the most energy in the network are located at level 1 or level 2. According to formula (9), the expected energy consumption of each level 2 and level 1 node in one cycle is$p \cdot ((\Delta_2 - 1) \cdot E_r + \Delta_2 \cdot E_s) + E_l \cdot k^2 \cdot (1 - p^{\Delta_4})$ $p \cdot ((\Delta_1 - 1) \cdot E_r + \Delta_1 \cdot e_s) + E_l \cdot k^2 \cdot (1 - p^{\Delta_3})$,so the life cycle of the network can be collectively given by formula (11).

More specifically, according to formula (8), we can get

$$\begin{cases} \Delta_1 = k^0 + k^2 + k^4 + \ldots + k^{d-2} \\ \Delta_2 = k^0 + k^2 + k^4 + \ldots + k^{d-2}\ \ if\ d\ is\ even \end{cases} \tag{12}$$

$$\begin{cases} \Delta_1 = k^0 + k^2 + k^4 + \ldots + k^{d-1} \\ \Delta_2 = k^0 + k^2 + k^4 + \ldots + k^{d-3}\ \ if\ d\ is\ odd \end{cases} \tag{13}$$

When d is even, according to formula (12), $\Delta_1 = \Delta_2$, so $p \cdot \Delta_1 = p \cdot \Delta_2$, which means that in a data collection cycle, each node at level 1 and level 2 has the same number of data packets to send. However, the node of level 1 consumes less energy to transmit data to sink than that of level 2, so the node at level 2 first dies. Therefore, for even-level tree networks, the network lifecycle of RDML algorithm can be directly expressed as

$$\ell_{RDML} = E_{init}/(p \cdot ((\Delta_2 - 1) \cdot E_r + \Delta_2 \cdot E_s) + E_l \cdot k^2 \cdot (1 - p^{\Delta_4})).$$

When d is odd, according to formula (13), it is obvious that the expected number of data packets to be sent in one cycle of each level 2 node is $p \cdot \Delta_2$, less than the number of each level 1 node $p \cdot \Delta_1$. However, the node of level 1 sends one data packet with less energy consumption than that of level 2. Therefore, the maximum energy consumption of the node can be expressed as $\max\{p \cdot ((\Delta_2 - 1) \cdot E_r + \Delta_2 \cdot E_s) + E_l \cdot k^2 \cdot (1 - p^{\Delta_4}), p \cdot ((\Delta_1 - 1) \cdot E_r + \Delta_1 \cdot e_s) + E_l \cdot k^2 \cdot (1 - p^{\Delta_3})\}$, so the network life cycle of RDML algorithm is as formula (11). ■

Theorem 4: The network expected lifetime ratio ψ_{lt} of the RDML algorithm proposed in this paper relative to the previous TPO algorithm is expressed as

$$\begin{aligned}\psi_{lt} = (p \cdot & \left(\sum\nolimits_{j=1}^{d-1} k^j \cdot e_r + \sum\nolimits_{j=0}^{d-1} k^j \cdot e_s\right) \\ & + \mathrm{e}_l \cdot k \cdot \left(1 - p^{\sum_{j=0}^{d-2} k^j}\right))/\max\{p \cdot ((\Delta_2 - 1) \cdot E_r \\ & + \Delta_2 \cdot E_s) + E_l \cdot k^2 \cdot (1 - p^{\Delta_4}), p \cdot ((\Delta_1 - 1) \cdot E_r \\ & + \Delta_1 \cdot e_s) + E_l \cdot k^2 \cdot (1 - p^{\Delta_3})\}\end{aligned} \tag{14}$$

Proof: Since all nodes use r distance to transmit data when adopting TPO algorithm, the data packets generated by all nodes in the network are sent to sink through level 1 nodes. Finally, the node with the highest expected energy consumption is the node at level 1 of the k-ary tree. Because each level 1 node has the expected energy consumption of $p \cdot \left(\sum_{j=1}^{d-1} k^j \cdot e_r + \sum_{j=0}^{d-1} k^j \cdot e_s\right) + \mathrm{e}_l \cdot k \cdot (1 - p^{\sum_{j=0}^{d-2} k^j})$ in a data collection cycle. Thus, the network life cycle of the TPO algorithm can be expressed as $\ell_{TPO} = E_{init}/(p \cdot \left(\sum_{j=1}^{d-1} k^j \cdot e_r + \sum_{j=0}^{d-1} k^j \cdot e_s\right) + \mathrm{e}_l \cdot k \cdot (1 - p^{\sum_{j=0}^{d-2} k^j}))$, according to Theorem 3, the network life cycle ratio ψ_{lt} of the RDML and TPO algorithm can be expressed as formula (14). ■

C. ANALYSIS OF ENERGY UTILIZATION RATIO

According to formula (6), the energy utilization of WSN is the ratio of the energy consumed by all nodes in the network at the end of its life cycle to the initial energy of the entire network. Initially, each node in the network has the energy of E_{init}. Since the k-ary tree has d levels and there are k^i nodes at each level i, the total initial energy of the network is $\sum_{i=1}^{d} k^i \cdot E_{init}$.

Theorem 5: When the sensor network adopts the RDML scheme for data collection, its network energy utilization can be expressed as

$$\begin{aligned}\varphi_{RDML} = (p \cdot E_s \cdot (k^d + k^{d-1}) & + \sum\nolimits_{i=2}^{d-2} k^i (p \cdot ((\Delta_i - 1) \cdot E_r \\ & + \Delta_i \cdot E_s) + E_l \cdot k^2 \cdot (1 - p^{\Delta_{i+2}})) \\ & + k \cdot (p \cdot ((\Delta_1 - 1) \cdot E_r + \Delta_1 \cdot e_s) \\ & + E_l \cdot k^2 \cdot (1 - p^{\Delta_3}))) \cdot \ell_{RDML}/(\sum\nolimits_{i=1}^{d} k^i \cdot E_{init})\end{aligned} \tag{15}$$

Proof: The total network energy consumption is the sum of the energy consumption of each node. For the RDML algorithm, the total energy consumption of the network in a data collection cycle consists of the following three parts: (1) The energy consumption of all nodes of level d and level $d - 1$ without any receiving and listening is $p \cdot E_s \cdot (k^d + k^{d-1})$. (2) The energy consumption of all nodes at the level $i (2 \leq i \leq d - 2)$ of receiving, idle listening and sending data packets through $2r$ distance is $\sum_{i=2}^{d-2} k^i (p \cdot ((\Delta_i - 1) \cdot E_r + \Delta_i \cdot E_s) + E_l \cdot k^2 \cdot (1 - p^{\Delta_{i+2}}))$. (3) The energy consumption of all level 1 nodes of the $2r$ distance receiving, idle listening and r distance sending packets is $k \cdot (p \cdot ((\Delta_1 - 1) \cdot E_r + \Delta_1 \cdot e_s) + E_l \cdot k^2 \cdot (1 - p^{\Delta_3}))$. Therefore, the expected total energy consumption of all nodes in the network during a data collection cycle is given by $p \cdot E_s \cdot (k^d + k^{d-1}) + \sum_{i=2}^{d-2} k^i (p \cdot ((\Delta_i - 1) \cdot E_r + \Delta_i \cdot E_s) + E_l \cdot k^2 \cdot (1 - p^{\Delta_{i+2}})) + k \cdot (p \cdot ((\Delta_1 - 1) \cdot E_r + \Delta_1 \cdot e_s) + E_l \cdot k^2 \cdot (1 - p^{\Delta_3}))$. According to Theorem 3, the network life cycle of the RDML algorithm is ℓ_{RDML}. Therefore, the network energy utilization of the RDML algorithm can be expressed as formula (15). ■

Theorem 6: The ratio of the RDML algorithm in this paper to the network energy utilization ratio of the previous TPO algorithm can be expressed as follows

$$\begin{aligned}\psi_{eur} = (p \cdot E_s \cdot (k^d + k^{d-1}) & + \sum\nolimits_{i=2}^{d-2} k^i (p \cdot ((\Delta_i - 1) \cdot E_r \\ & + \Delta_i \cdot E_s) + E_l \cdot k^2 \cdot (1 - p^{\Delta_{i+2}})) + k \cdot (p \cdot ((\Delta_1 - 1) \\ & \cdot E_r + \Delta_1 \cdot e_s) + E_l \cdot k^2 \cdot (1 - p^{\Delta_3}))) \cdot \psi_{lt}/(p \cdot e_s \cdot k^d \\ & + \sum\nolimits_{i=1}^{d-1} k^i (p \cdot \left(\sum\nolimits_{j=1}^{d-i} k^j \cdot e_r \right. \\ & \left. + \sum\nolimits_{j=0}^{d-i} k^j \cdot e_s\right) + \mathrm{e}_l \cdot k \cdot (1 - p^{\sum_{j=0}^{d-i-1} k^j})))\end{aligned} \tag{16}$$

Proof: The expected total energy consumption of TPO algorithm in a data collection cycle consists of two parts: (1) The energy consumption of all level d leaf nodes that only send data packets is $p \cdot e_s \cdot k^d$. (2) There are both receiving data packets and idle listening, as well as the energy consumption of all level $i (1 \leq i \leq d - 1)$ internal nodes that send data packets, which is $\sum_{i=1}^{d-1} k^i (p \cdot (\sum_{j=1}^{d-i} k^j \cdot e_r + \sum_{j=0}^{d-i} k^j \cdot e_s) + \mathrm{e}_l \cdot k \cdot (1 - p^{\sum_{j=0}^{d-i-1} k^j}))$. Thus, the total energy consumption of the network in one cycle is $p \cdot e_s \cdot k^d + \sum_{i=1}^{d-1} k^i (p \cdot (\sum_{j=1}^{d-i} k^j \cdot e_r + \sum_{j=0}^{d-i} k^j \cdot e_s) + \mathrm{e}_l \cdot k \cdot (1 - p^{\sum_{j=0}^{d-i-1} k^j}))$. According to Theorem 4, the network life cycle of the TPO algorithm is ℓ_{TPO}. Therefore, the energy utilization φ_{TPO} of the network can be given by $(p \cdot e_s \cdot k^d + \sum_{i=1}^{d-1} k^i (p \cdot (\sum_{j=1}^{d-i} k^j \cdot e_r + \sum_{j=0}^{d-i} k^j \cdot e_s) + \mathrm{e}_l \cdot k \cdot (1 - p^{\sum_{j=0}^{d-i-1} k^j}))) \cdot \ell_{TPO}/(\sum_{i=1}^{d} k^i \cdot E_{init})$. Since $\psi_{eur} = \varphi_{RDML}/\varphi_{TPO}$, combined with Theorem 5, the formula (16) can be obtained by simplifying. ■

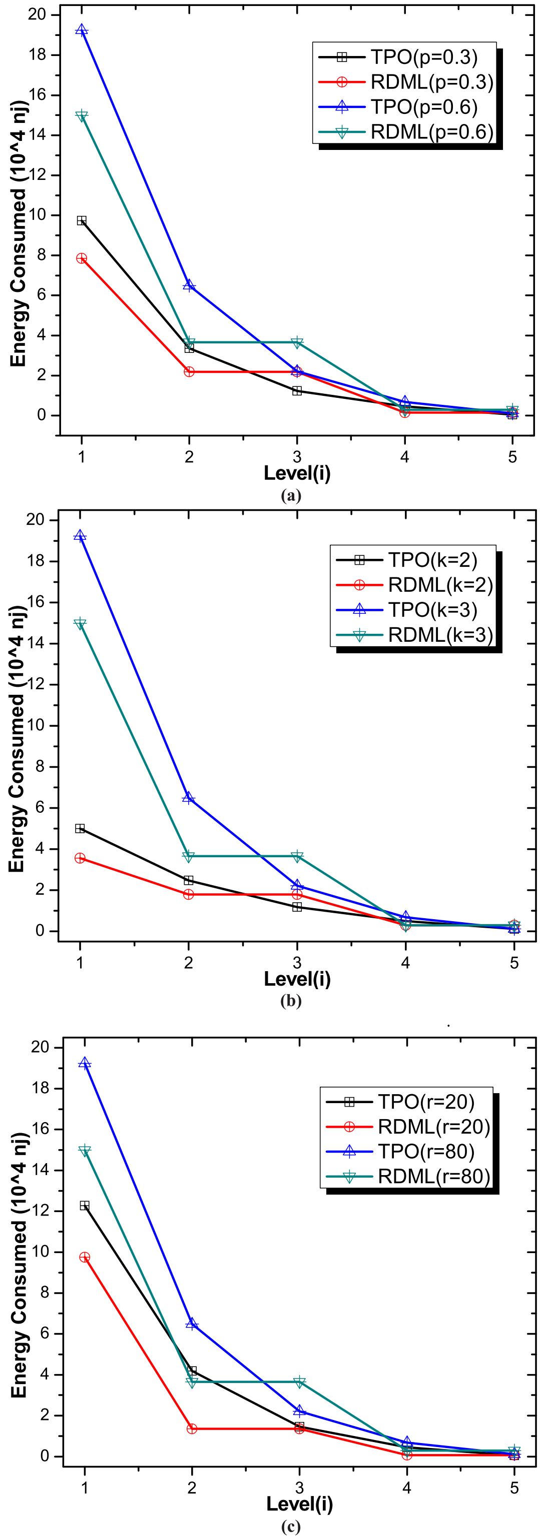
D. NUMERICAL RESULTS AND COMPARISON

Now, we show the numerical results obtained from the above sections analysis. The energy consumption value of sending and receiving data is calculated by equation (1) (2), while the energy consumption of idle listening is related to the length of slot, which is taken as 800 nJ. We take the transmission radius r as 80, the tree's level d is 5, k is 3, and the probability p of node generating a data packet as 0.6. Firstly, the energy consumption of nodes at different levels in one cycle can be obtained by changing r, k and p respectively, as shown in Figure 9(a), 9(b) and 9(c).

As can be seen from Figure 9(a), the energy consumption of RDML scheme increases intermittently with the decrease of levels, while that of TPO scheme increases continuously. When the network generates data packets with the same probability, the energy consumption of hotspots nodes of RDML scheme (ie, level 1 and level 2 nodes) is less than that of TPO. And when p increases from 0.3 to 0.6, the energy consumption of both schemes is increased. This is because in the RDML scheme for $2r$ distance transmission, the nodes of each level only need to bear the data packets of the downstream interval level nodes, while in the TPO scheme, the nodes of each level have to bear the data packets of all the downstream level nodes. The increase in p causes an increase in the amount of data packets generated by nodes in the network, so that the energy consumption of the two schemes also increases. In addition, the energy consumption of nodes in the third and fifth layers of RDML scheme is relatively higher, because the nodes at the fifth layer only send data, and it takes more energy to send data by $2r$ distance. Although the number of packets received by the level 3 node is less than that of the TPO, it is not enough to offset the higher energy consumption of sending and receiving packets with $2r$ distance. As the layer is closer to the sink, the advantage of the nodes at the same layer adopting the RDML scheme to receive less packets than the TPO is more obvious, so the energy consumption is less than that of the TPO. Compared with TPO scheme, the expected energy consumption of RDML hotspots nodes is reduced by 19.32% - 43.64%.

Obviously, analogous changes are also observed in Figure 9(b) and 9(c). Similar to the analysis in Figure 9(a), when k changes from 2 to 3, it increases the number of nodes in the network. Similarly, when r increases from 20 to 80, it also increases the transmission energy consumption of a single packet. Compared with the TPO scheme, the expected energy consumption of the RDML hotspots nodes in Figure 9(b) and Figure 9(c) is reduced by 22.07% - 43.64% and 20.53% - 67.66%, respectively.

Next, we change the parameters separately and get the network expected life cycle comparison charts of the two schemes. Figure 10(a) shows the expected life cycle of the network under RDML and TPO schemes when r is 80, d is 5, k is 3, and p is 0.3, 0.6, and 0.9, respectively. Figure 10(b), Figure 10(c) and Figure 10(d) respectively show the network expected life cycle of the two schemes when kr and d change.



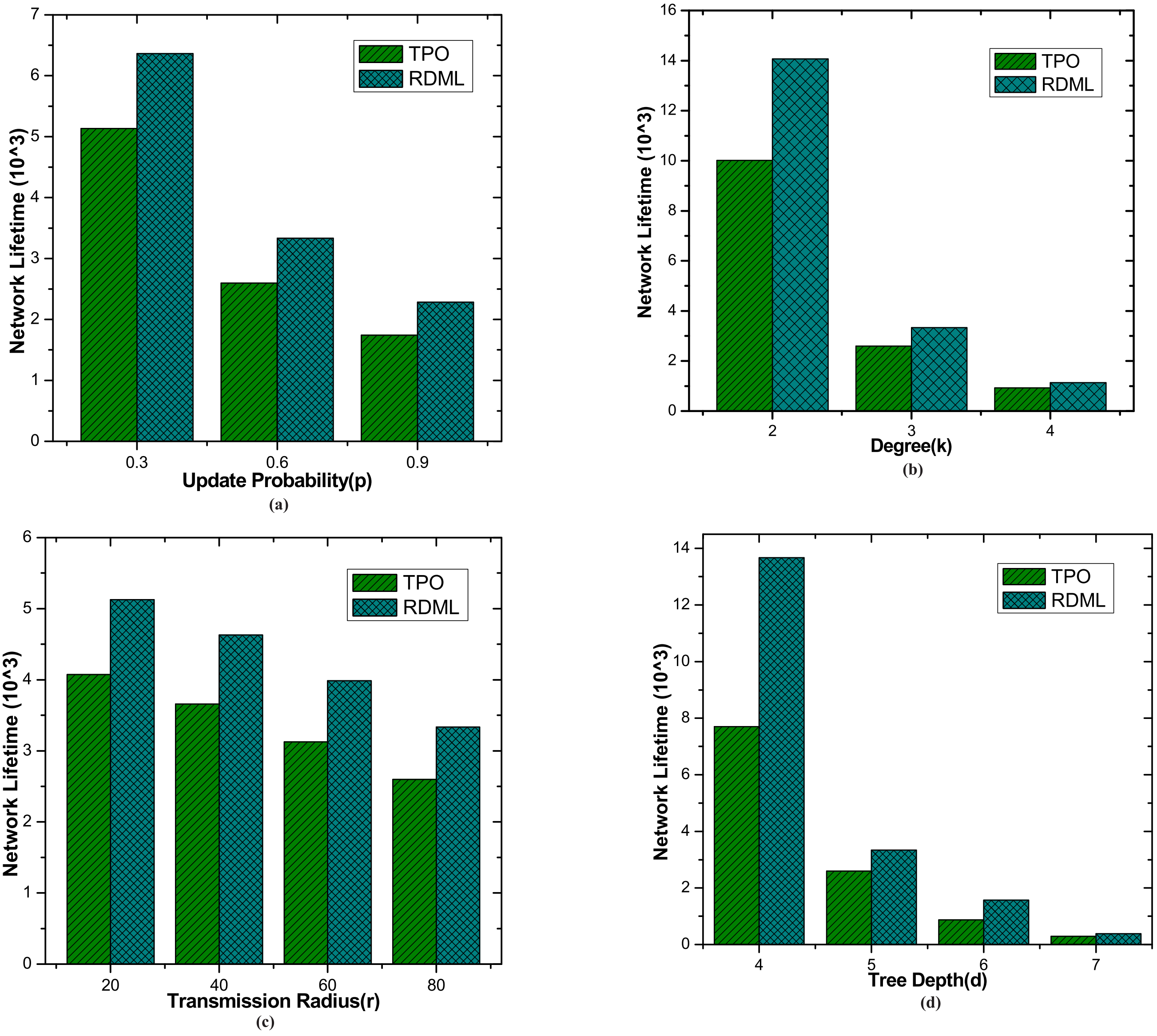
(a)

(b)

(c)

FIGURE 9. **(a) Expected energy consumption of nodes at different levels in one cycle under $r = 80$, $d = 5$ and $k = 3$. (b) Expected energy consumption of nodes at different levels in one cycle under $r = 80$, $d = 5$ and $p = 0.6$. (c) Expected energy consumption of nodes at different levels in one cycle under $d = 5$, $k = 3$, and $p = 0.6$.**

It can be seen from Figure 10(a) - Figure 10(d) that, with the increase of p, k, r and d, the expected network life cycle of both schemes starts to decrease, but the expected



(a)

(b)

(c)

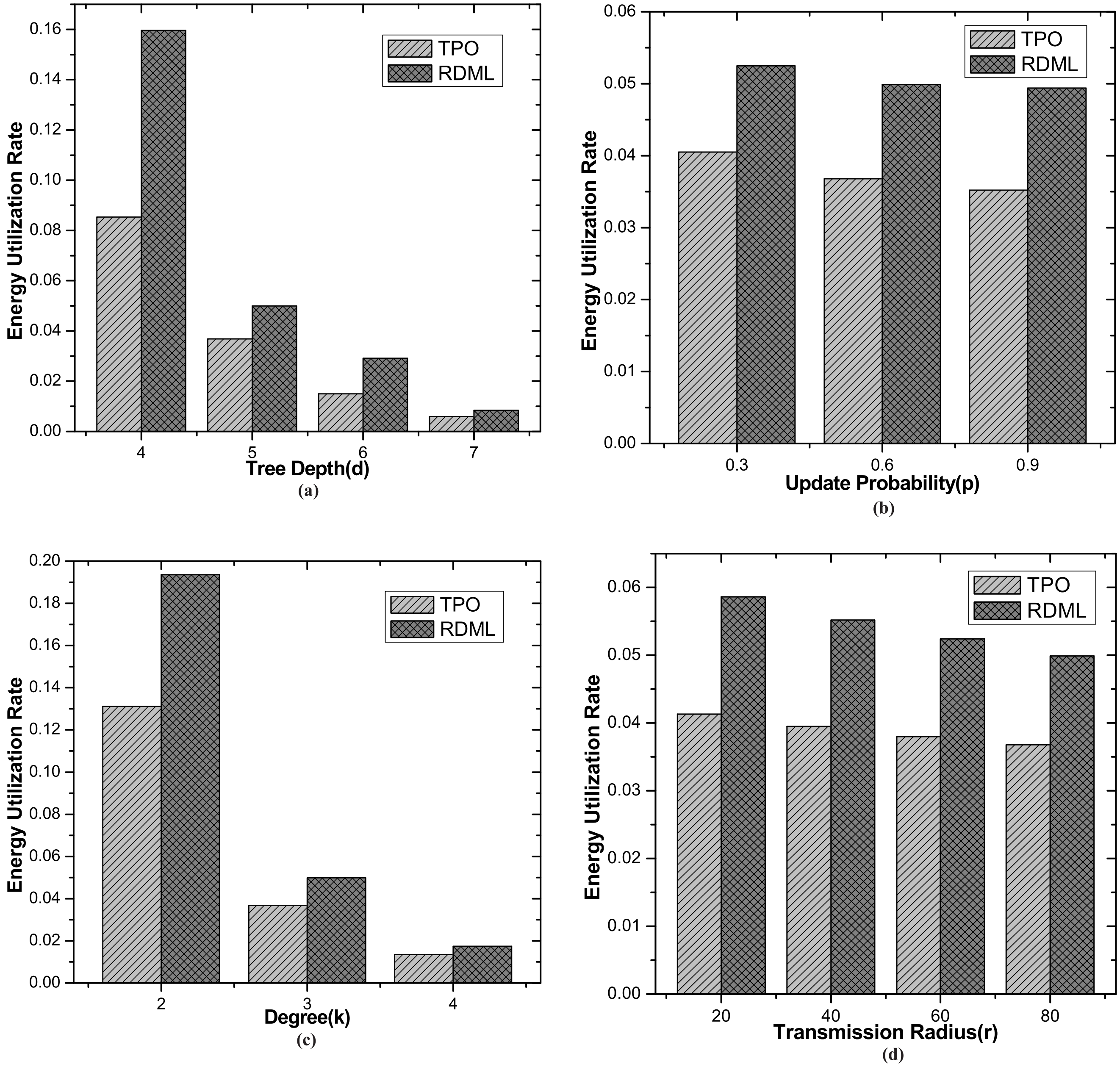
(d)

FIGURE 10. (a) RDML and TPO network expected lifetime under $r = 80$, $d = 5$ and $k = 3$. (b) RDML and TPO network expected lifetime under $r = 80$, $d = 5$ and $p = 0.6$. (c) RDML and TPO network expected lifetime under $d = 5$, $k = 3$ and $p = 0.6$. (d) RDML and TPO network expected lifetime under $r = 80$, $k = 3$ and $p = 0.6$.

network life cycle of RDML scheme is always better than that of TPO. This is because the life cycle of the network is determined by the lifetime of the hotspots nodes, while the increase of each parameter will lead to an increase in the energy consumption of the network in a cycle. As can be seen from Figure 9(a) - 9(c) above, the energy consumption of hotspots nodes using RDML scheme is always less than that of TPO in a data collection cycle. In addition, the increase of k and d significantly increases the number of network nodes, so that the expected network lifetime of the two schemes is reduced rapidly.

In Figure 10(d), it is easy to observe that compared with TPO scheme, the expected lifetime of RDML scheme increases more significantly when d is even. As proved in Theorem 3, when d is even, level 1 and level 2 nodes theoretically bear the same amount of data, which is more balanced than when d is odd. Therefore, RDML scheme improves the network life cycle more obviously when the level of the tree is even. In Figure 10(a) - Figure 10(d), the expected network life cycle of the RDML scheme increased by 23.94% - 31.31%, 22.04% - 40.50%, 25.83% - 28.33% and 28.33% - 81.01%, respectively, relative to TPO.

In addition, we also obtain the network expected energy utilization rate of the two schemes by changing the value of each parameter. First, we change the number of layers of the tree and get Figure 11(a) below. The network expected energy utilization of both schemes decreases as the number of tree layers increases, but the network expected energy



(a)

(b)

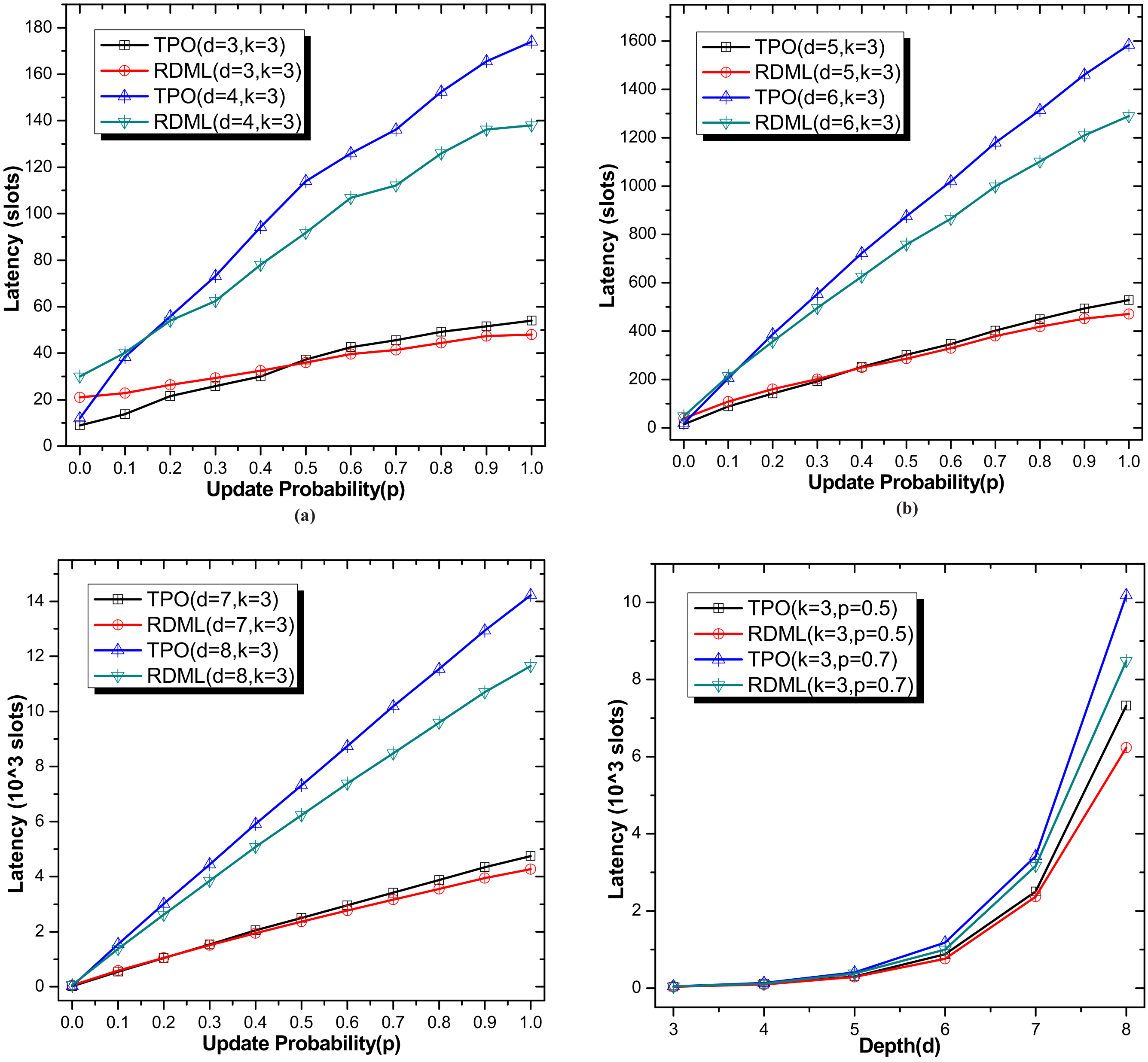
(c)

(d)

FIGURE 11. (a) RDML and TPO expected energy utilization rate under $r = 80$, $k = 3$ and $p = 0.6$. (b) RDML and TPO expected energy utilization rate under $d = 5$, $r = 80$ and $k = 3$. (c) RDML and TPO expected energy utilization rate under $d = 5$, $r = 80$ and $p = 0.6$. (d) RDML and TPO expected energy utilization rate under $d = 5$, $k = 3$ and $p = 0.6$.

utilization of RDML is always higher than that of TPO. This is because the network lifetime of the two schemes decreases sharply as the number of tree layers increases (as shown in Figure 10(d)), which leads to the decrease of network energy utilization of both schemes. However, the network energy consumption of RDML scheme is more balanced than that of TPO, so that the expected energy utilization rate of RDML is always higher than that of TPO. Besides, we observe a similar phenomenon as in Figure 10(d). When the level of the tree is even, the network expected energy utilization of the RDML scheme increases more obviously than when the level of the tree is odd. Specifically, in contrast with TPO scheme, when d is 4 and 6, RDML increases by 87.10% - 94.00%, while when d is 5 and 7, RDML increases by 35.60% - 42.37%. This is also the result of the energy consumption of even layer is more balanced than that of odd layer.

Then, we take the level d of the tree as 5, and change the probability of network generating data packet p, the degree of k and the transmission distance r to get the network expected energy utilization of the two schemes as shown in Figure 11(b), Figure 11(c) and Figure 11(d). Similarly, with the increase of the probability p of data packet generation,



(a)

(b)

(c)

(d)

FIGURE 12. (a) Average latency of data collection under $k = 3$, $d = 3$ and $d = 4$. (b) Average latency of data collection under $k = 3$, $d = 5$ and $d = 6$. (c) Average latency of data collection under $k = 3$, $d = 7$ and $d = 8$. (d) Average latency of data collection under $k = 3$, $p = 0.5$ and $p = 0.7$.

the degree of k and the transmission distance r, the expected energy utilization of both schemes decreases gradually, but the network energy utilization of RDML is always greater than that of TPO. Compared with TPO, the network expected energy utilization rate of RDML increased by 29.63% - 40.34% when p increased from 0.3 to 0.9, 28.89% - 47.56% when k increased from 2 to 4, and 35.60% - 41.89% when r increased from 20 to 80.

In summary, from the comparison of the numerical results of the expected energy consumption, life cycle and energy utilization of the above two schemes, RDML can use the remaining energy to make some nodes transmit data at $2r$ distance, which has better performance than TPO. Theoretical analysis shows that the expected energy consumption of RDML is reduced by 19.32% - 67.66% compared with the TPO hotspots nodes, the network life cycle is increased by 22.04% - 81.01%, and the energy utilization rate is increased by 28.89% - 94.00%.

VI. ANALYSIS OF EXPERIMENTAL RESULTS

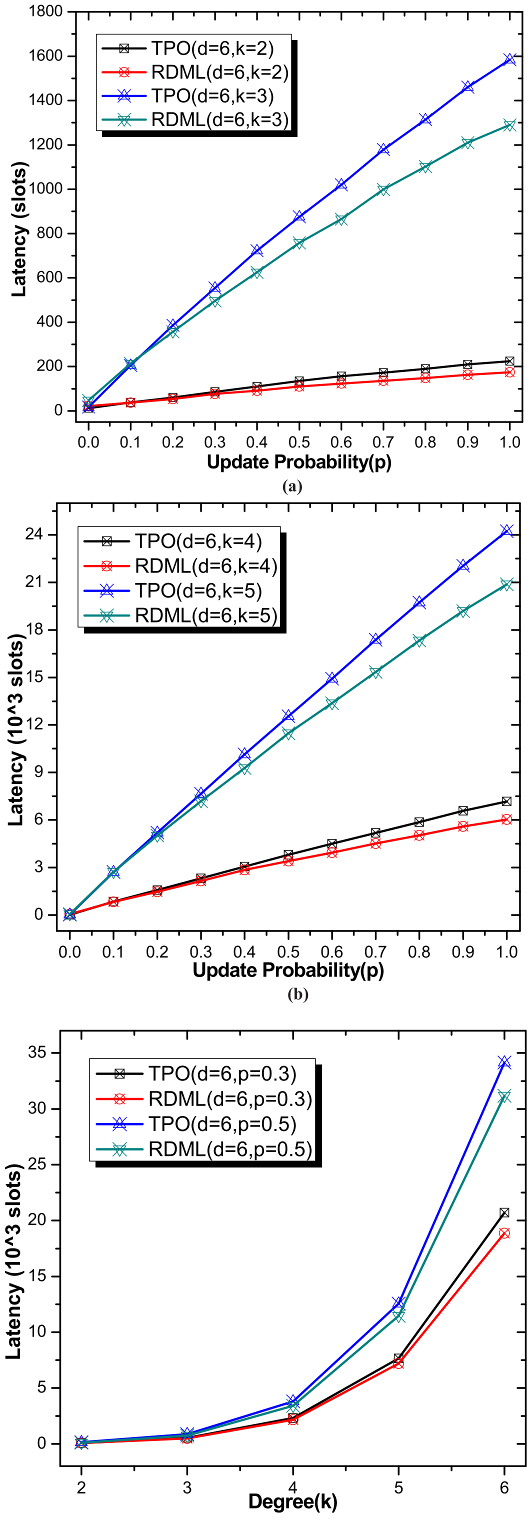
Our simulation experiments are carried out under the environment of Matlab. Matlab is a very excellent numerical calculation tool. It is a high-level language and interactive environment for algorithm development, data visualization and data analysis. The constants used in the experiment are consistent with Table 1: The energy E_{elec} of the transmitting circuit loss is 50 nJ/bit, and the energy ε_{fs} required for power amplification is 10 pJ/bit/m^2. Each node has the same initial energy E_{init} 0.5J, the number of bits per packet l is 16 bits.

A. DELAY ANALYSIS

First, we consider the data collection delay of the two schemes in one cycle when the level d of tree network and the probability p of data packets generated by nodes change. In order to reduce the error of the experiment, we conducted multiple experiments to take the average value. We fix the degree k of the tree as 3, and increase the level of the tree from 3 to 8, to obtain the average waiting time of data collection of RDML and TPO schemes when p changes from 0 to 1 for each node, as shown in Figure 12(a) - 12(c).

It can be seen that with the increase of p, the waiting time for data collection of both schemes in Figure 12(a) - 12(c) increases, but the data collection delay of RDML gradually starts to be less than that of TPO, and the larger p is, the more obvious the delay reduction is. The reason is that the increase of p increases the number of data packets that the network needs to collect, thus increasing the delay of data collection of the two schemes. At the same time, the increase of p also increases the number of rounds of data collection by the two scheduling algorithms in one cycle. In the RDML scheme for $2r$ distance transmission, the number of child nodes per node is at most k^2, while each TPO node is at most k. Therefore, in the first round of data collection, RDML scheduling takes more time slots than TPO. When the probability p is very low, the data collection can be completed in the early rounds, at this time the TPO delay is less. However, as p increases, the data needs multiple rounds to be collected. Since the maximum amount of packets transmitted to sink in each round of RDML is $k^2 + k$, while the TPO is at most k. So RDML can complete data collection with less rounds, thus reducing latency. As p increases, the number of rounds required to complete data collection also increases, so the advantage of RDML in reducing latency becomes more apparent.

In addition, with the increase of tree level, RDML scheduling has an advantage over TPO in the waiting time of data collection under lower probability of packet generation, and even tree level is more advantageous than odd tree level. Specifically, when d is 4, 6 and 8, the lowest probability that RDML scheduling has less latency than TPO scheduling data collection drops from 0.15 to 0.12 and approaches 0 gradually. When the tree level is 3, 5 and 7, the lowest probability decreases from 0.46 to 0.37 to 0.2. The reason is that the increase of tree level increases the number of rounds of data collection, so it can delay less than TPO scheduling at lower probability. When d is even, the transmission of even and odd layers in the tree has the same number of hops, but when d is odd, the transmission of odd layers in the tree is one more hop than that of even layers. Therefore, when the tree level is even, it can have a delay advantage over TPO scheduling when generating packets with less probability. When d increases from 3 to 8, the delay of RDML scheduling data collection is reduced by 0.95% - 20.69% compared with TPO.



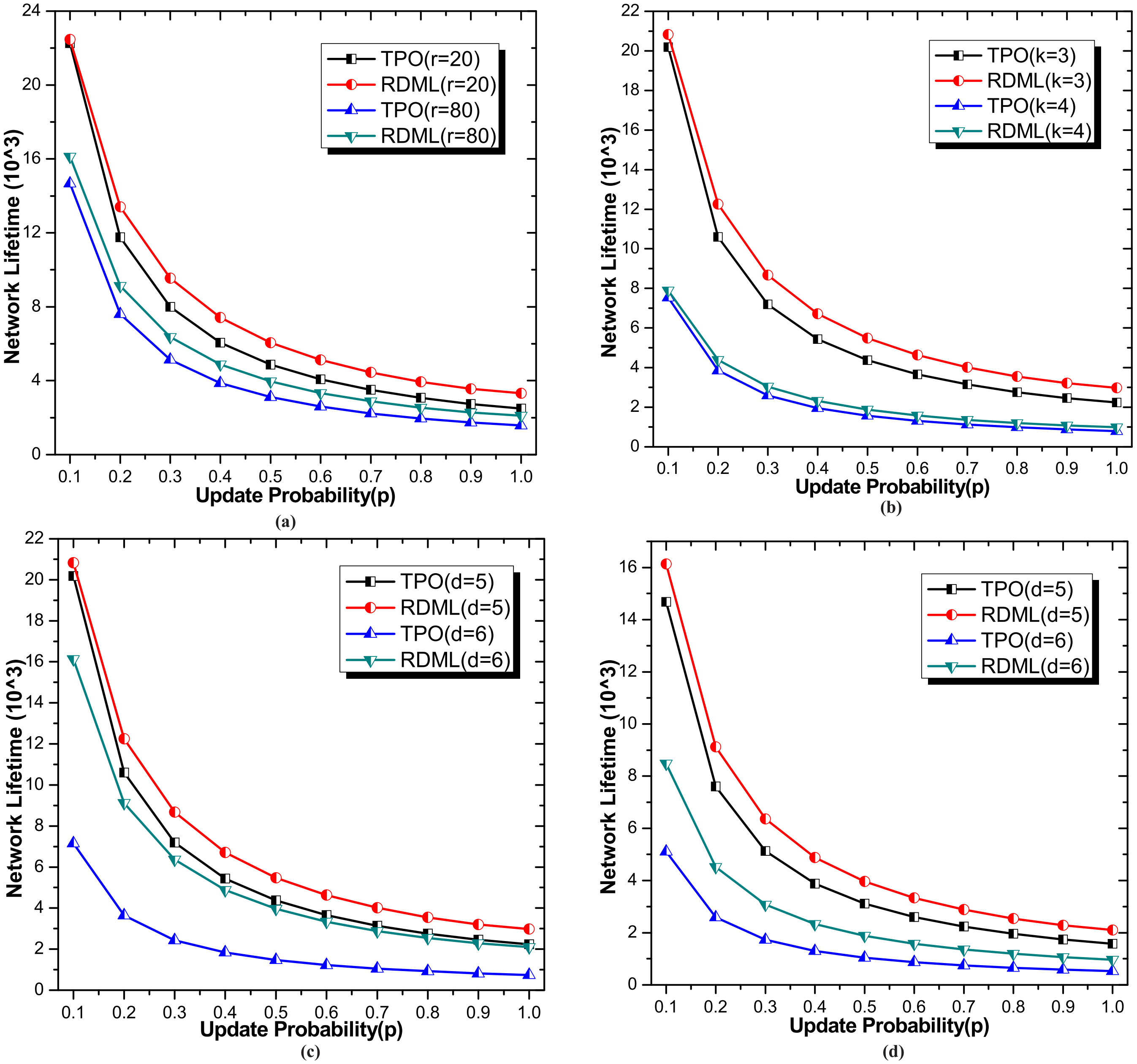
(a)

(b)

(c)

FIGURE 13. **(a) Average latency of data collection under $d = 6$, $k = 2$ and $k = 3$. (b) Average latency of data collection under $d = 6$, $k = 4$ and $k = 5$. (c) Average latency of data collection under $d = 6$, $p = 0.3$ and $p = 0.5$.**

Figure 12(d) shows the relationship between the data collection delay of the two schedules and the tree level when



(a)

(b)

(c)

(d)

FIGURE 14. **(a) RDML and TPO network lifetime under $d = 5$, $k = 3$, $r = 20$ and $r = 80$. (b) RDML and TPO network lifetime under $d = 5$, $r = 40$, $k = 3$ and $k = 4$. (c) RDML and TPO network lifetime under $k = 3$, $r = 40$, $d = 5$ and $d = 6$. (d) RDML and TPO network lifetime under $k = 3$, $r = 80$, $d = 5$ and $d = 6$.**

k is 3 and p is 0.5 and 0.7 respectively. It can be seen that the average waiting time of both scheduling data collection increases rapidly with the increase of the tree level, but the delay of RDML scheduling is always lower than TPO. Compared with TPO, the delay of RDML scheduling data collection is reduced by 3.22% - 19.47%.

Next, we consider the average data collection delay of the two scheduling schemes when the degree of k and the probability of packet generation p change. Figure 13 (a) and Figure 13 (b) show the average waiting time for data collection of the two schemes when d is 6 and k is increased from 2 to 5, respectively. It can be observed that as long as the probability of packet generation is greater than 0.2, our RDML scheduling has less latency than the TPO. But at the same time, with the increase of k, the probability of the lowest packet generation that RDML scheduling is less than the average data collection waiting time of TPO is also slightly increased. Specifically, when k is 2 and 5 respectively, the lowest p increases from 0.1 to 0.13. This is because the increase of k causes the number of packets of the child nodes to be received by each node to increase. Therefore, when the tree is "fat", it reduces the advantage of RDML scheduling delay. Figure 13 (c) shows the relationship between the data collection delay and k of the two schedules

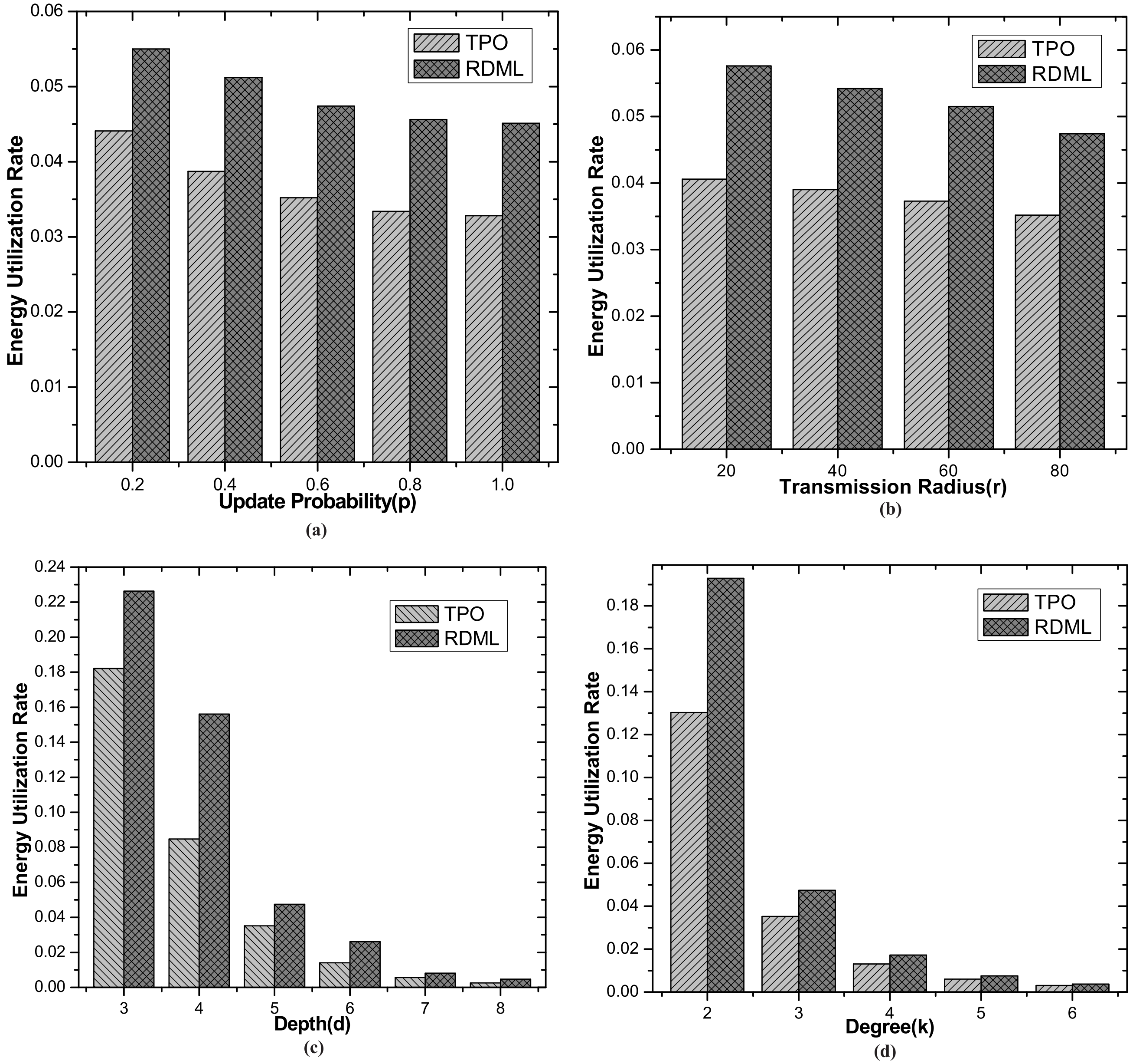


FIGURE 15. **(a) RDML and TPO energy utilization rate under $k = 3$, $d = 5$ and $r = 50$. (b) RDML and TPO energy utilization rate under $k = 3$, $d = 5$ and $p = 0.6$. (c) RDML and TPO energy utilization rate under $k = 3$, $r = 80$ and $p = 0.6$.(d) RDML and TPO energy utilization rate under $r = 80$, $d = 5$ and $p = 0.6$.**

when d is 6 and p is 0.3 and 0.5, respectively. Obviously, RDML scheduling is always less than the delay of TPO. Compared with TPO, the delay of RDML scheduling data collection is reduced by 1.68% - 22.32%.

The above experiments show that RDML scheduling has a great advantage over TPO scheduling in data collection delay, and the effect of even tree layers is better than that of odd tree layers. This advantage of RDML scheduling is more obvious when the tree network is thinner and the probability of data packet generation is higher. The experimental results show that RDML scheduling can reduce the maximum delay by 22.32% compared with TPO scheduling.

B. COMPREHENSIVE EFFECTS OF DIFFERENT PARAMETERS

We obtained the network life cycle and energy utilization ratio of the two schemes under different parameters through simulation experiments. Figure 14(a) shows how the network lifetime of the two schemes varies with probability p, when d is 5, k is 3 and transmission distance r is 20 and 80, respectively. Obviously, as the probability of packet generation increases, the network life cycle of both schemes decreases. However, regardless of whether r is 20 or 80, RDML scheduling is always longer than the network life cycle of TPO scheduling, which is increased by 0.86% - 33.14%. When the data packet is generated with a low probability ($p = 0.1$), and r is 20 and 80 respectively, the network life cycle of RDML is increased by 0.86% and 9.91%, respectively. This is because the energy consumption of the node sending the data packet is related to r^2. When the probability of the network generating the data packet is low, the energy consumption is greatly affected by the r change.

Figure 14(b) shows that when d is 5, r is 40 and k is taken as 3 and 4 respectively, the life cycle of the network under different probability of generating data packets. Figure 14(c) and Figure 14(d) show the network life cycle changes of two schemes where k is 3, d is 5 and 6, r is 40 and 80, respectively, and p is increased from 0.1 to 1. Obviously, the increase of each parameter leads to the decrease of the network life cycle of both, but our RDML scheduling network lifetime is always better than TPO, and the advantage of even tree level is obviously greater than odd tree level. This is basically consistent with our previous theoretical analysis. The experimental data show that the network life cycle of RDML is 3.20% - 84.40% higher than that of TPO.

When d is 5, r is 80 and k is 3, the change of network energy utilization rate with p is shown in Figure 15(a). When d is 5, k is 3 and p is 0.6, the variation of network energy utilization with r is shown in Figure 15(b). It can be seen that as the probability of packet generation and the transmission radius increase, the energy utilization rate of both schemes begins to decrease, but the network utilization of RDML scheduling is always greater than TPO. Compared with TPO, RDML's network energy utilization increased by 24.72% - 37.50% and 34.66% - 41.87% respectively.

Figure 15(c) and Figure 15(d) show the changes of network energy utilization rate with the depth of tree and the degree of k, respectively. Similarly, as the tree depth and the degree of k increase, the network energy utilization of both schedules begins to decline, but RDML is higher than TPO. In addition, when the RDML tree level is even, the effect is more obvious than the odd level, which is consistent with our previous analysis. Compared with TPO, the network energy utilization of RDML in the two graphs increased by 24.34% - 88.00% and 23.33% - 48.04% respectively.

From the experimental results, we can find that our RDML scheduling has better network life cycle and energy utilization under different parameters than TPO scheduling. This is basically consistent with our theoretical analysis in Chapter 5. The experimental results show that compared with TPO scheduling, RDML scheduling increases the network life cycle by 0.86% - 84.40%, and the energy utilization rate increases by 23.33% - 88.00%.

VII. CONCLUSION

In this paper, we have proposed a continuous data collection strategy RDML for dynamic traffic patterns. The RDML scheme uses the remaining energy in the network to increase the transmission radius of some nodes to mr. Therefore, it has better performance in terms of network life cycle, energy utilization rate and data collection delay than previous strategies. We theoretically analyzed the performance of the proposed scheme and compared it with the advanced TPO scheme. We also compared the performance under different parameters by simulation experiments. The results of analysis and simulation experiments show that: (1) The RDML scheme reduces the amount of data undertaken by hotspots nodes and balances the energy consumption of the network. Therefore, its network life cycle and energy utilization rate are better than TPO scheme. (2) RDML scheme uses continuous transmission strategy to reduce idle listening, and enables some nodes to transmit data through mr distance, which reduces the number of rounds of data collection in a cycle. Therefore, there is less latency than the TPO scheme. For future work, we plan to show the effectiveness of the proposed scheme on a real sensor network test platform.

REFERENCES

[1] H. M. Abdulsalam, B. A. Ali, and E. AlRoumi, "Usage of mobile elements in Internet of Things environment for data aggregation in wireless sensor networks," *Comput. Elect. Eng.*, vol. 72, pp. 789–807, Nov. 2018.
[2] X. Xiang, W. Liu, A. Liu, N. Xiong, Z. Zeng, and Z. Cai, "Adaptive duty cycle control based opportunistic routing scheme to reduce delay in cyber physical systems," *Int. J. Distrib. Sensor Netw.*, vol. 15, no. 4, 2019. doi: 10.1177/1550147719841870.
[3] W. Osamy, A. M. Khedr, A. Aziz, and A. A. El-Sawy, "Cluster-tree routing based entropy scheme for data gathering in wireless sensor networks," *IEEE Access*, vol. 6, pp. 77372–77387, 2018.
[4] I. L. Vasconcelos, I. C. Martins, C. M. S. Figueiredo, and A. L. L. Aquino, "A data sample algorithm applied to wireless sensor network with disruptive connections," *Comput. Netw.*, vol. 146, pp. 1–11, Dec. 2018.
[5] H. Teng, Y. Liu, A. Liu, N. N. Xiong, Z. Cai, T. Wang, and X. Liu, "A novel code data dissemination scheme for Internet of Things through mobile vehicle of smart cities," *Future Gener. Comput. Syst.*, vol. 94, pp. 351–367, May 2019.
[6] T. Li, K. Ota, T. Wang, X. Li, Z. Cai, and A. Liu, "Optimizing the coverage via the UAVs with lower costs for information-centric Internet of Things," *IEEE Access*, vol. 7, pp. 15292–15309, 2019.
[7] W. Zhao and X. Tang, "Scheduling sensor data collection with dynamic traffic patterns," *IEEE Trans. Parallel Distrib. Syst.*, vol. 24, no. 4, pp. 789–802, Apr. 2013.
[8] H. Teng, W. Liu, T. Wang, A. Liu, X. Liu, and S. Zhang, "A cost-efficient greedy code dissemination scheme through vehicle to sensing devices (V2SD) communication in smart city," *IEEE Access*, vol. 7, pp. 16675–16694, 2019.
[9] Y. Liu, A. Liu, N. N. Xiong, T. Wang, and W. Gui, "Content propagation for content centric networking from location-based social networks," *IEEE Trans. Syst., Man, Cybern. Syst.*, to be published. doi: 10.1109/TSMC.2019.2898982.
[10] T. Wang, G. Zhang, A. Liu, M. Z. A. Bhuiyan, Q. Jin, "A secure IoT service architecture with an efficient balance dynamics based on cloud and edge computing," *IEEE Internet Things J.*, to be published. doi: 10.1109/JIOT.2018.2870288.
[11] Z. Zhou, D. Zhao, L. Liu, and P. Hung, "Energy-aware composition for wireless sensor networks as a service," *Future Gener. Comput. Syst.*, vol. 80, pp. 299–310, Mar. 2018.
[12] A. Liu, J. Min, K. Ota, and M. Zhao, "Reliable differentiated services optimization for network coding cooperative communication system," *Comput. Syst. Sci. Eng.*, vol. 33, no. 4, pp. 235–250, Jul. 2018.
[13] Y. Ren, W. Liu, T. Wang, X. Li, N. Xiong, and A. Liu, "A collaboration platform for effective task and data reporter selection in crowdsourcing network," *IEEE Access*, vol. 7, pp. 19238–19257, 2019.
[14] B. Huang, W. Liu, T. Wang, X. Li, H. Song, and A. Liu, "Deployment optimization for data centers in vehicular networks," *IEEE Access*, vol. 7, pp. 20644–20663, 2019.
[15] J. Chen, K. Hu, Q. Wang, Y. Sun, Z. Shi, and S. He, "Narrowband Internet of Things: Implementations and applications," *IEEE Internet Things J.*, vol. 4, no. 6, pp. 2309–2314, Dec. 2017.

[16] X. Liu, Y. Liu, N. Zhang, W. Wu, and A. Liu, ''Optimizing trajectory of unmanned aerial vehicles for efficient data acquisition: A matrix completion approach,'' *IEEE Internet Things J.*, vol. 6, no. 2, pp. 1829–1840, Apr. 2019.

[17] D. Zhang, R. Shen, J. Ren, and Y. Zhang, ''Delay-optimal proactive service framework for block-stream as a service,'' *IEEE Wireless Commun. Lett.*, vol. 7, no. 4, pp. 598–601, Aug. 2018.

[18] Y. Liu, A. Liu, N. Zhang, X. Liu, M. Ma, and Y. Hu, ''DDC: Dynamic duty cycle for improving delay and energy efficiency in wireless sensor networks,'' *J. Netw. Comput. Appl.*, vol. 131, pp. 16–27, Apr. 2019.

[19] W. Qi, W. Liu, X. Liu, A. Liu, T. Wang, N. N. Xiong, and Z. Cai, ''Minimizing delay and transmission times with long lifetime in code dissemination scheme for high loss ratio and low duty cycle WSNs,'' *Sensors*, vol. 18, no. 10, p. 3516, 2018. doi: 10.3390/s18103516.

[20] Z. Zhou, D. Zhao, G. Hancke, L. Shu, and Y. Sun, ''Cache-aware query optimization in multiapplication sharing wireless sensor networks,'' *IEEE Trans. Syst., Man, Cybern. Syst.*, vol. 48, no. 3, pp. 401–417, Mar. 2018.

[21] X. Ju, W. Liu, C. Zhang, A. Liu, T. Wang, N. Xiong, and Z. Cai, ''An energy conserving and transmission radius adaptive scheme to optimize performance of energy harvesting sensor networks,'' *Sensors*, vol. 18, no. 9, p. 2885, 2018. doi: 10.3390/s18092885.

[22] X. Xiang, W. Liu, N. Xiong, H. Song, A. Liu, and T. Wang, ''Duty cycle adaptive adjustment based device to device (D2D) communication scheme for WSNs,'' *IEEE Access*, vol. 6, pp. 76339–76373, 2018.

[23] J. Tan, W. Liu, M. Xie, H. Song, A. Liu, M. Zhao and G. Zhang, ''A low redundancy data collection scheme to maximize lifetime using matrix completion technique,'' *J. Wireless Commun. Netw.*, vol. 5, no. 1, p. 5, 2019. doi: 10.1186/s13638-018-1313-0.

[24] C. Xu, J. Ren, D. Zhang, Y. Zhang, Z. Qin, and K. Ren, ''GANobfuscator: Mitigating information leakage under GAN via differential privacy,'' *IEEE Trans. Inf. Forensics Security*, to be published. doi: 10.1109/TIFS.2019.2897874.

[25] T. Wang, J. Zhou, A. Liu, M. Z. A. Bhuiyan, G. Wang, and W. Jia, ''Fog-based computing and storage offloading for data synchronization in IoT,'' *IEEE Internet Things J.*, to be published. doi: 10.1109/JIOT.2018.2875915.

[26] J. Kim, X. Lin, N. B. Shroff, and P. Sinha, ''Minimizing delay and maximizing lifetime for wireless sensor networks with anycast,'' *IEEE/ACM Trans. Netw.*, vol. 18, no. 2, pp. 515–528, Apr. 2010.

[27] R. Madan, S. Cui, S. Lall, and A. Goldsmith, ''Cross-layer design for lifetime maximization in interference-limited wireless sensor networks,'' *IEEE Trans. Wireless Commun.*, vol. 5, no. 11, pp. 3142–3152, Nov. 2006.

[28] H. Yetgin, K. T. K. Cheung, M. El-Hajjar, and L. H. Hanzo, ''A survey of network lifetime maximization techniques in wireless sensor networks,'' *IEEE Commun. Surveys Tuts.*, vol. 19, no. 2, pp. 828–854, 2nd. Quart., 2017.

[29] H. Yetgin, K. T. K. Cheung, M. El-Hajjar, and L. Hanzo, ''Network-lifetime maximization of wireless sensor networks,'' *IEEE Access*, vol. 3, pp. 2191–2226, 2015.

[30] M. Huang, W. Liu, T. Wang, H. Song, X. Li, and A. Liu, ''A queuing delay utilization scheme for on-path service aggregation in services oriented computing networks,'' *IEEE Access*, vol. 7, pp. 23816–23833, 2019.

[31] J. Tan, W. Liu, T. Wang, N. N. Xiong, H. Song, A. Liu, and Z. Zeng, ''An adaptive collection scheme-based matrix completion for data gathering in energy-harvesting wireless sensor networks,'' *IEEE Access*, vol. 7, pp. 6703–6723, 2019.

[32] Z. Li, Y. Liu, A. Liu, S. Wang, and H. Liu, ''Minimizing convergecast time and energy consumption in green Internet of Things,'' *IEEE Trans. Emerg. Topics Comput.*, to be published. doi: 10.1109/TETC.2018.2844282.

[33] W. Zhang, W. Liu, T. Wang, A. Liu, Z. Zeng, H. Song, and S. Zhang, ''Adaption resizing communication buffer to maximize lifetime and reduce delay for WVSNs,'' *IEEE Access*, vol. 7, pp. 48266–48287, 2019.

[34] M. Luo, K. Wang, Z. Cai, A. Liu, Y. Li, and C. F. Cheang, ''Using imbalanced triangle synthetic data for machine learning anomaly detection,'' *Comput., Mater. Continua*, vol. 58, no. 1, pp. 15–26, 2019.

[35] W. Liu, P. Zhuang, H. Liang, J. Peng, and Z. Huang, ''Distributed economic dispatch in microgrids based on cooperative reinforcement learning,'' *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 29, no. 6, pp. 2192–2203, Jun. 2018.

[36] Y. Liu, A. Liu, X. Liu, and X. Huang, ''A statistical approach to participant selection in location-based social networks for offline event marketing,'' *Inf. Sci.*, vol. 480, pp. 90–108, Apr. 2019.

[37] X. Deng, L. He, X. Li, Q. Liu, L. Cai, and Z. Chen, ''A reliable QoS-aware routing scheme for neighbor area network in smart grid,'' *Peer-to-Peer Netw. Appl.*, vol. 9, no. 4, pp. 616–627, Jul. 2016.

[38] X. Peng, J. Ren, L. She, D. Zhang, J. Li, and Y. Zhang, ''Boat: A block-streaming App execution scheme for lightweight IoT devices,'' *IEEE Internet Things J.*, vol. 5, no. 3, pp. 1816–1829, Jun. 2018.

[39] L. Hu, A. Liu, M. Xie, and T. Wang, ''UAVs joint vehicles as data mules for fast codes dissemination for edge networking in smart city,'' *Peer-to-Peer Netw. Appl.*, to be published. doi: 10.1007/s12083-019-00752-0.

[40] C. Zhang, Y. Lin, L. Zhu, A. Liu, Z. Zhang, and F. Huang, ''CNN-VWII: An efficient approach for large-scale video retrieval by image queries,'' *Pattern Recognit. Lett.*, vol. 123, pp. 82–88, May 2019.

[41] X. Deng, T. He, L. He, J. Gui, and Q. Peng, ''Performance analysis for IEEE 802.11s wireless mesh network in smart grid,'' *Wireless Pers. Commun.*, vol. 96, no. 1, pp. 1537–1555, Sep. 2017.

[42] W. Yang, W. Liu, Z. Zeng, A. Liu, G. Huang, N. N. Xiong, and Z. Cai, ''Adding active slot joint larger broadcast radius for fast code dissemination in WSNs,'' *Sensors*, vol. 18, no. 11, p. 4055, 2018. doi: 10.3390/s18114055.

[43] T. Wang, G. Zhang, M. D. Z. A. Bhuiyan, A. Liu, W. Jia, and M. Xie, ''A novel trust mechanism based on fog computing in sensor—Cloud system,'' *Future Gener. Comput. Syst.*, to be published. doi: 10.1016/j.future.2018.05.049.

[44] X. Kui, Y. Sun, S. Zhang, and Y. Li, ''Characterizing the capability of vehicular fog computing in large-scale urban environment,'' *Mobile Netw. Appl.*, vol. 23, no. 4, pp. 1050–1067, Aug. 2018.

[45] J. Ren, H. Guo, C. Xu, and Y. Zhang, ''Serving at the edge: A scalable IoT architecture based on transparent computing,'' *IEEE Netw.*, vol. 31, no. 5, pp. 96–105, Aug. 2017.

[46] F. Wang, W. Liu, T. Wang, M. Zhao, M. Xie, H. Song, X. Li, and A. Liu, ''To reduce delay, energy consumption and collision through optimization duty-cycle and size of forwarding node set in WSNs,'' *IEEE Access*, vol. 7, pp. 55983–56015, 2019. doi: 10.1109/ACCESS.2019.2913885.

[47] J. Tan, W. Liu, T. Wang, S. Zhang, A. Liu, M. Xie, M. Ma, and M. Zhao, ''An efficient information maximization based adaptive congestion control scheme in wireless sensor network,'' *IEEE Access*, to be published. doi: 10.1109/ACCESS.2019.2915385.

[48] Q. Li, A. Liu, T. Wang, M. Xie, and N. Xiong, ''Pipeline slot based fast rerouting scheme for delay optimization in duty cycle based M2M communications,'' *Peer-to-Peer Netw. Appl.*, to be published. doi: 10.1007/s12083-019-00753-z.

[49] J. Ren, Y. Zhang, N. Zhang, D. Zhang, and X. Shen, ''Dynamic channel access to improve energy efficiency in cognitive radio sensor networks,'' *IEEE Trans. Wireless Commun.*, vol. 15, no. 5, pp. 3143–3156, May 2016.

[50] Y. Wu, K. S. Liu, J. A. Stankovic, T. He, and S. Lin, ''Efficient multichannel communications in wireless sensor networks,'' *ACM Trans. Sensor Netw.*, vol. 12, no. 1, pp. 1–24, 2016.

[51] A. Liu, Y. Hu, and Z. Chen, ''An energy-efficient mobile target detection scheme with adjustable duty cycles in wireless sensor networks,'' *Int. J. Ad Hoc Ubiquitous Comput.,*, vol. 22, no. 4, pp. 203–225, 2016.

[52] W. Zhao, X. Tang, and L. Xu, ''Constructing routing structures for sensor data collection with dynamic traffic patterns,'' *Comput. Netw.*, vol. 142, pp. 272–290, Sep. 2018.

[53] K. Zhang, Z. Feng, and X. Li, ''Weight-based link scheduling for convergecast in WirelessHART network,'' *Int. J. Distrib. Sensor Netw.*, vol. 12, no. 7, Jul. 2016, Art. no. 4594183.

[54] Q. Chen, H. Gao, Z. Cai, L. Cheng, and J. Li, ''Distributed low-latency data aggregation for duty-cycle wireless sensor networks,'' *IEEE/ACM Trans. Netw.*, vol. 26, no. 5, pp. 2347–2360, Oct. 2018.

QIANG WANG is currently pursuing the master's degree with the School of Computer Science and Engineering, Central South University, China. His research interests include data process, crowd sensing networks, and wireless sensor networks.

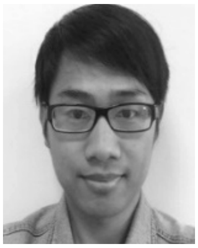
WEI LIU received the Ph.D. degree in computer application technology from Central South University, China, in 2014. He is currently an Associate Professor and a Senior Engineer with the School of Informatics, Hunan University of Chinese Medicine, China. His research interests include software engineering, data mining, and medical informatics. He has published over 20 papers in the related fields.

TIAN WANG received the B.Sc. and M.Sc. degrees in computer science from Central South University, in 2004 and 2007, respectively, and the Ph.D. degree from the City University of Hong Kong, in 2011. He is currently an Associate Professor with the National Huaqiao University of China. His research interests include wireless sensor networks, social networks, and mobile computing.

MING ZHAO received the M.Sc. and Ph.D. degrees in computer science from Central South University, China, in 2003 and 2007, respectively, where he is currently a Professor with the School of Software. He is also a member of the China Computer Federation (CCF). His major research interest includes wireless networks.

XIONG LI received the master's degree in mathematics and cryptography from Shanxi Normal University (SNNU), China, in 2009, and the Ph.D. degree in computer science and technology from the Beijing University of Posts and Telecommunications (BUPT), China, in 2012. He is currently an Associate Professor with the School of Computer Science and Engineering, Hunan University of Science and Technology (HNUST), China. He has published more than 70 referred journal papers in his research interests, which include cryptography, information security, cloud computing security, and so on. He has served as a TPC member of several international conferences on information security and reviewer for more than 30 ISI indexed journals. He was a recipient of the *Journal of Network and Computer Applications* 2015 Best Research Paper Award. He is currently an Editor of *Telecommunication Systems* and *KSII Transactions on Internet and Information Systems*.

MANDE XIE was born in 1977. He received the Ph.D. degree in circuit and system from Zhejiang University, in 2006. He is currently a Professor with Zhejiang Gongshang University. His research interests include wireless sensor networks (WSNs), social networks, and privacy preservation.



MING MA is currently pursuing the Ph.D. degree with the Department of Computer Science, Stony Brook University, USA. His research interests include geometric modeling, digital geometric processing, and wireless sensor networks.

GUOPING ZHANG was born in 1977. She received the master's degree in circuit and system from Hangzhou Dianzi University, in 2006. She is currently a Senior Experimenter with Zhejiang SCI-Tech University. Her research interests include wireless sensor networks (WSNs) and wireless rechargeable sensor networks (WCSN).

ANFENG LIU received the M.Sc. and Ph.D. degrees in computer science from Central South University, China, in 2002 and 2005, respectively, where he is currently a Professor with the School of Information Science and Engineering. His major research interest includes wireless sensor networks. He is a member (E200012141M) of the China Computer Federation (CCF).

• • •