IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received November 7, 2018, accepted November 13, 2018, date of publication November 19, 2018,
date of current version December 18, 2018.

Digital Object Identifier 10.1109/ACCESS.2018.2882210

An Energy Efficient and Reliable In-Network Data Aggregation Scheme for WSN

JINHUAN ZHANG[1], PENG HU[2], FANG XIE[1], JUN LONG[1], AND AN HE[1]
[1]School of Information Science and Engineering, Central South University, Changsha 410083, China
[2]Seneca College of Applied Arts and Technology, Toronto, ON M3J 3M6, Canada

Corresponding author: Jun Long (jlong@mail.csu.edu.cn)

This work was supported in part by the Natural Science Foundation of Hunan Province under Grant 2017JJ3417, in part by the Science and Technology Plan of Hunan under Grant 2016TP1003, and in part by the National Natural Science Foundation of China under Grant 61472450.

ABSTRACT Data aggregation can reduce the data transmission between the nodes, and thus save the energy and extend the life of the network. Many related researches on in-network data aggregation take the generalized maximum functions. For the cases that the original packets of N nodes aggregated into M $(1 < M < N)$ packets, it is a challenge to improve the energy efficiency and reduce the transmission delay under the transmission reliability guarantee. In this paper, a novel ring-based in-network data aggregation scheme is proposed to this problem. The network is partitioned into rings and the data aggregation is executed ring by ring from outside to inside. To ensure transmission reliability, the source or intermediate aggregating node unicasts multiple aggregated packet copies to its next hop node in the inner ring with the maximum residual energy. The reliability is higher with the more unicasting packet copies. However, more sending packets copies will lead to more additional energy cost. Besides, nodes close to the sink tend to relay more size of data packets and the energy is depleted more quickly than nodes far to the sink. Meanwhile, the nodes close to the sink need to relay the aggregated packets, which contain more information. If the number of packet copies is too small, the packets loss will greatly worse the transmission reliability. Based on this, the number of unicasting packet copies is adaptively adjusted through fuzzy logic. The proposed scheme adaptively unicasts variable number of aggregated packets copies continuously in a window according to the request transmission reliability and the imbalance of nodes energy cost. Our analysis and simulation results show the effectiveness of the proposed scheme.

INDEX TERMS In-network data aggregation, fuzzy logic, ring, reliability, lifetime.

I. INTRODUCTION

Artificial Intelligence, Big Data, Cloud Computing, Edge Computing and 5G communication networks are promoting the rapid development and wide application of the Internet of Things (IoT) [1]–[6]. IoT is attracting more and more investment. And it is developing more and more rapidly because of its wide applications, industrial prospects, and potentially huge commercial value. The rapid development of IoT also faces serious challenges [7]–[10]. With the rapid growth of interconnected IoT devices and the emergence of many IoT systems (such as intelligent transportation, smart grid, smart home, etc.), the amount of data that the network needs to collect, store, transmit and process increases exponentially. Thus, there is a contradiction among the rapid growth of data and the limited spectrum resources, transmission capacity of the network. The congestion caused by the huge amount of data becomes more serious, which leads to the data transmission delay sharply increased. Then, the quality of service (QoS) is declined and the quality of experience (QoE) is deteriorated. Therefore, with the constant growth of user numbers, the amount of data produced by IoT is explosively growing, which may cause the corresponding problems of greatly increased transmission delay and reduced data quality. It is urgent to study and solve these problems for the development of human-machine-thing interconnection.

Wireless sensor network (WSN) is an integral part of IoT [11]. To cope with the decline of network quality of service caused by the rapid growth of IoT data, it is an ideal and effective way to reduce the amount of data transmitted by WSN without increasing the investment of existing hardware infrastructure. Data aggregation can minimize redundant data transmission to reduce the amount of data transmitted in the network. Thus, appropriately reducing and minimizing the amount of data by data aggregation is an effective way to reduce the amount of network data transmissions [12], [13]. However, many of the existing data aggregation mechanisms assume that the original packets of N nodes can aggregate into one packet at the aggregation node. For example, in many existing studies on in-network data aggregation, they adopt the generalized maximum (GM) functions, such as in [14]. The final aggregated result corresponds to an element of the sensed data. That model is suitable to the Min, Max or Average cases. For the cases that the original packets of N nodes aggregated into M $(1 < M < N)$ packets, the aggregated packets have more size than that of the original packet, they have less packets to relay for the nodes far to the sink. Therefore, they have more remaining energy. Whereas, the nodes close to the sink tend to relay more size of data packets and the energy is depleted quickly.

In addition, it is particularly important to ensure reliable transmission of aggregated packets because the aggregated packets contain more information. To ensure the transmission reliability, re-transmission schemes (such as Automatic Repeat-reQuest, ARQ) are widely explored for packet loss recovery [15]. However, these schemes need to wait for the acknowledgement (ACK) reply. Because of the loss of ACK, unnecessary packets will be retransmitted, resulting in additional energy consumption and transmission delay. Broadcast is another way to reduce packet loss and guarantee the reliable transmission. Fig.1 shows an illustration of data aggregation with broadcast over lossy links [14]. In the figure, the solid line with arrow denotes the successful packet transmission. The dotted line with a cross represents the failed transmission. As shown in Fig.1, one source node S broadcasts its own generated packet to the inner ring (represented by red line with arrow). Although failed transmission happens, the packet is successfully transmitted because one intermediate node in the inner ring receives the packet. The receiving intermediate node aggregates and further broadcasts to next-hop nodes. The process is repeatedly until the packet is transmitted to the Sink. It is obviously that broadcasting will result in many redundant packets in the entire network transmission as shown in Fig. 1. Consequently, it consumes more additional energy, especially when the data aggregation model does not adopt the GM functions. This is inconsistent with reducing the amount of network transmission data and saving network bandwidth resources.

To address these issues, in this paper we propose a novel ring and fuzzy logic based in-network data aggregation scheme in considering the energy efficiency and network transmission reliability. The core idea of the proposed scheme is to adaptively unicast variable number of aggregated packet copies continuously in a window according to the request transmission reliability and the imbalance of nodes energy cost. The main contributions of the paper are:

(1) A ring based in-network data aggregation scheme is proposed to satisfy the network transmission reliability by unicasting multiple packets copies. The energy efficient and

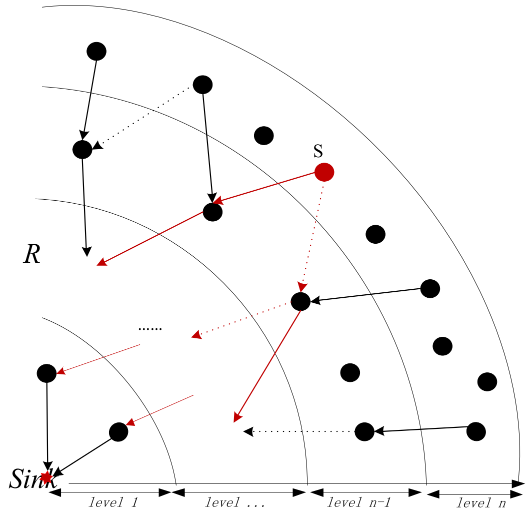


FIGURE 1. Data aggregation with broadcast over lossy wireless links.

reliable in-network data aggregation with unicast scheme can improve the energy efficiency with the satisfaction of network transmission reliability. In the scheme, multiple copies of data packets in the sending window can be transmitted continuously for data aggregation by unicast. For sending nodes, there is no need to wait for the confirming ACKs of the data packets. The receiving node does data aggregation adopting data aggregation model sequentially. The aggregated packets may have more size than that of the original packet, which conforms to information aggregation characteristics. Furthermore, it greatly reduces the transmission data compared with the broadcast data aggregation. Through without waiting for the ACK reply, it could reduce the end to end (E2E) delay of all information aggregated to the sink.

(2) The proposed fuzzy logic-based scheme is with adaptive variable unicasting packet copies γ. In the proposed scheme, one node sends multiple packet copies in a window continuously to its next hop node by unicast. The probability of data packets successfully transmitted from one node to its next hop node on the routing path is higher with the more sending packet copies γ. However, more sending packets copies will lead to more additional energy cost. Furthermore, for the nodes far to the sink, they have less packets to relay. Therefore, they have more remaining energy and could send more data copies to ensure the successful transmission. Whereas, the nodes close to the sink tend to relay more size of data packets and the energy is depleted quickly. Thus, the small number of sending packet copies is suitable to prolong the network lifetime. However, the nodes close to the sink need to relay the aggregated packets, which contain more information. If the number of packet copies is too small, the packets loss will greatly worse the transmission reliability. The core idea of the manuscript is to adaptively determine the sending number of packet copiesγ to prolong the network lifetime and ensure the transmission reliability.

(3) Comprehensive simulations are conducted to verify the effectiveness of the proposed scheme. Comprehensive

simulations are implemented to explain the availability of the proposed scheme. Simulation results show that the proposed scheme could improve the energy efficiency subject to the guarantee of transmission reliability. In our simulation, the lifetime could be improved by 17.14% compared with invariable sending packet copies aggregation approach. It presents the effectiveness of our strategy.

The rest of this paper is organized as follows: the related works are reviewed in Section 2; Section 3 describes the system model; Section 4 explains the novel proposed scheme including discussion and performance analysis; simulation performance evaluations are presented in Section 5; and the conclusion is made in Section 6.

II. BACKGROUND AND RELATED WORK

To support the big data streaming under the QoS requirements, [16] presented the design of a CoNtainer-based virtualized networked computing architecture for the real-time support of IoT applications. Data compression can minimize redundant data to reduce the amount of data transmitted in the network. Reference [17] presented an accelerated distributed solving algorithm to guarantee the error minimization of reconstructed data and the fast-optimal value converge. In addition to the researches on architecture and data compression, data aggregation is much studied.

Data aggregation could reduce the amount of data transmitted by network without increasing the investment of existing hardware infrastructure and it is an ideal and effective way to cope with the problems caused by the current rapid growth of IoT data. Important data aggregation mechanisms include tree-based, cluster-based and centralized WSN data aggregation mechanisms [18]–[22]. There are quite numerous researches on tree-based data aggregation strategies. In tree-based data aggregation, nodes are organized into a tree, and the data is aggregated in the intermediate nodes during the process of propagation to the root of the tree, and eventually transmitted to the root node. The key of this mechanism is to establish an effective data aggregation tree. Typical algorithms are Dynamic Data Aggregation Tree (DDAT) and Energy-aware Distributed Heuristic Spanning Aggregation Tree (EADAT) [23]. DDAT reduces data transmission and energy consumption by minimizing the distance traveled. The main idea of EADAT is to close broadcasting of all leaf nodes to reduce the amount of broadcast data. In cluster-based data aggregation mechanism, the whole network is partitioned into different clusters. Each cluster is composed of multiple sensing nodes. The Low Energy Adaptive Clustering Hierarchy (LEACH) algorithm is a typical cluster-based data aggregation protocol. Unlike the cluster-based data aggregation protocol, (Power Efficient Gathering Sensor Information Systems) PEGASIS proposes a chain-based data aggregation algorithm from the idea of minimizing energy consumption. For cluster-based data aggregation, a data aggregation optimization algorithm for single cluster head and multi-cluster head is proposed to prolong the network lifetime [24], [25]. A Prolong-SEP (P-SEP) is presented in [26] to find an energy-efficient policy to opt cluster heads in the WSNs. Reference [27] proposes hybrid multi-hop routing (HYMN) algorithm, which uses a mixture of planar multi-hop routing with high transmission distance and hierarchical multi-hop routing with data aggregation. In data aggregation-based data collection strategies, many of the existing data aggregation mechanisms are based on the assumption that the original packets of N nodes can aggregate into one packet at the aggregation node. For the cases that the original packets of N nodes aggregated into M packets, it is a challenge to improve the energy efficiency and reduce the transmission delay under the transmission reliability guarantee. Reference [14] studied the implication of wireless broadcast for data aggregation in lossy wireless sensor networks. But in this broadcast form, the additional energy consumption is obvious, especially when the data aggregation model does not adopt the ideal assumption of aggregating into one packet, resulting in many redundant packets in the entire network transmission by broadcast.

Because the wireless data transmission link is unreliable, there are quite a lot of research to improve network reliability while ensuring high energy efficiency. The ARQ protocol is the representative protocol based on the failure retransmission mechanism. The other is proliferation routing (PR) based on the packet regeneration mechanism [28]. In the two representative mechanisms, the lost packets will be retransmitted or replicated after certain transmission route to ensure the reliable transmission to reach the destination. In PR strategy, in order to ensure the reliability every certain number of routing hop, the packets will replicate the same number of copies. It fails to combine the characteristics of uneven energy consumption and data aggregation in sensor networks, which cause high energy consumption in hotspots.

In this paper, we propose a novel ring and fuzzy logic based in-network data aggregation scheme considering the energy efficiency and network transmission reliability. The transmission reliability is ensured by sending several data copies with unicast. It could adaptively send different number of packet copies continuously in a window according to the request transmission reliability and the imbalance of nodes energy cost.

III. THE SYSTEM MODEL AND PROBLEM STATEMENT

A. THE SYSTEM MODEL

A typical and the most applied periodic data gathering wireless sensor network is considered in the paper. Assume there are N sensor nodes randomly deployed in a planar circular area. The radius of the area is R meters. Once the nodes are deployed, the location of the nodes are fixed. Assume the transmission range is r meters and it is identical for each node. Each sensor node periodically generates a packet and forwards the packet to the next hop intermediate node. The intermediate node does data aggregation and the aggregated data is transmitted to one node (called the sink) located at the center of the network through multi-hop relays. We assume the packet initiated by each node is with the same size. The final aggregated or the initiated packets finish their

TABLE 1. Parameters and values.

Parameter	Value
threshold (d_0) (m)	87
E_{elec} (nJ/bit)	50
ε_{fs}(pJ/bit/m²)	10
ε_{amp} (pJ/bit/m⁴)	0.0013

transmission to the next hop in one time slot. Each sensor node is equipped with GPS and knows about the distance to the sink and its one hop neighbors. The wireless link is lossy and we assume the links are spatial dependent and time independent [29]. For the link between node v_k^i in the ring k and one node v_{k-1}^j in the ring $k-1$, the successful transmission reliability is represented by p_{ij}.

In this paper, the energy consumption model in [30] is adopted. Following the model, the energy consumed for sending one-bit data is calculated as follows.

$$\begin{cases} E_t = E_{elec} + \varepsilon_{fs} d^2 (d < d_0) \\ E_t = E_{elec} + \varepsilon_{amp} d^4 (d \geq d_0) \end{cases} \tag{1}$$

And the energy consumed for receiving one-bit data is computed by

$$E_r = E_{elec} \tag{2}$$

The meaning of parameters in the formula can refer to [30]. Tab.1 shows the parameters and the corresponding values gennerally adopted in energy consumption model. Following the model, it is easy to calculate the energy consumption for sending or receiving data packets.

The data aggregation model follows the model in Ref. Reference [31] and [32] nodes execute data aggregation when receiving data from the other node. Incoming data is aggregated with existing data in order of arrival. For data aggregation, we use v_k^i to represent one node in ring k and ϑ_i to denote the original packet size of node v_k^i, $\varphi(i,j)$ to indicate the intermediate aggregation result of node $v_{k-1}^j$with incoming data from node v_k^i. It can be abbreviated as φ_j to denote the current intermediate aggregation result of node v_{k-1}^j. The final aggregation result of node v_k^i is represented by μ_i. For the node without receiving any incoming data, $\mu_i = \vartheta_i$. When node v_{k-1}^j receives data μ_i from node v_k^i, node v_{k-1}^j aggregates μ_i with its current data φ_j. $\varphi_j = \vartheta_j$ when the node v_{k-1}^j first receives the incoming data.

$$\varphi(i,j) = \beta(\mu_i + \varphi_j) \tag{3}$$

In (3), β is the compression ratio and it is between 0 and 1.

B. THE PROBLEM STATEMENT

In the paper, in one round of data aggregation, each node generates, aggregates and forwards sensing packets to the sink. We use the running rounds until the first node death to approximately represent **network lifetime**. Thus, the network lifetime is mainly determined by the node with the largest energy consumption. **Network transmission reliability** is ued to represent the statistical success rate of packets delivered from all nodes to the sink in QoS level. It is denoted by ϵ. In order to improve the transmission reliability, several packet copies are transmitted continuously. If there are γ packet copies transmitted for one source or intermediate node v_k^i, the one hop transmission reliability is $\varepsilon_i = 1-(1-p_{ij})^\gamma$. Therefore, the network transmission reliability of node v_k^i to the sink is $\epsilon_i = \prod_k \left(1 - \left(1-p_{ij}\right)^\gamma\right)$. It is obviously that the higher γ brings the higher transmission reliability. However, the higher γ causes additional energy cost and reduces the network lifetime. Furthermore, the data collection is executed ring by ring from outside to inside. Although the data aggregation is employed in the data collection, the node close to the sink is usually with higher data load, because it needs to relay the aggregated packet and the packet size is bigger with the more packet aggregation. There exists energy consumption imbalance. For the nodes far to the sink, they have less packets to relay. Therefore, they have more remaining energy and could send more packet copies to ensure the successful transmission. Whereas, the nodes close to the sink tend to relay more size of data packets and the energy is depleted quickly. Thus, the small number of sending packet copies is suitable to prolong the network lifetime. However, the nodes close to the sink need to relay the aggregated packets, which contain more information. If the number of packet copies is too small, the packets loss will greatly worse the transmission reliability.

Thus, our paper aims to prolong the network lifetime under the satisfaction of the network transmission reliability requirement ϵ. As mentioned above that the network lifetime mainly determined by the node with the largest energy consumption of the network, reducing the energy cost of the node with the largest energy consumption is the key. Therefore, the objective is to meet the following requirements for all the nodes in the network.

$$\begin{cases} \mathrm{Max}(T) = \mathrm{Min}[\max_{0<i<n} (Cos)_i] \\ Cos_i = N_i^t E_t + N_i^r E_r \\ \dfrac{\Sigma_k \Sigma_i \sigma(v_k^i)}{N} \geq \epsilon \end{cases} \tag{4}$$

where, N_i^t and N_i^r represent the packet size of sending and receiving data of node v_k^i, respectively. E_t and E_r respectively denote the energy consumption for transmitting and receiving one-bit data. And, Cos_i denotes the total energy consumption of node v_k^i. $\sigma\left(v_k^i\right)$ represents whether the information generated by node v_k^i successfully transmitted to the sink. If it is successfully transmitted, $\sigma\left(v_k^i\right) = 1$. Otherwise, $\sigma\left(v_k^i\right) = 0$. Hence, our goal is to minimize the largest node energy consumption under the guarantee of the network reliability ϵ.

IV. THE DESIGN OF THE PROPOSED SCHEME

A. OVERVIEW OF THE PROPOSED SCHEME

In the scheme, the data collection is executed ring by ring from outside to inside. The source or intermediate node sends multiple packet copies continuously by the window W with

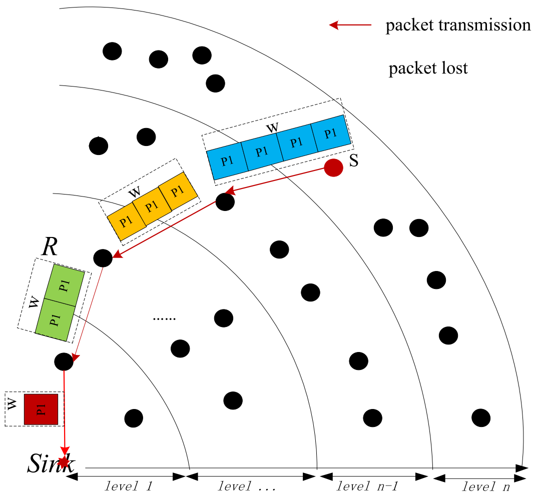


FIGURE 2. The sketch of data transmission of the node *S*.

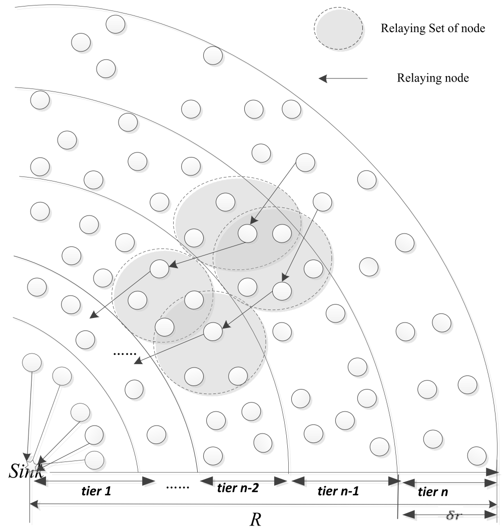


FIGURE 3. Ring partition and relaying node finding.

the size of γ. There is no need for sending nodes to wait for the ACK confirming of the sending packets. When the receiver receives multiple copies of the same data packet, only one copy is retained and the rest is deleted. Then, it will finish data aggregation following the data aggregation model as mentioned in the section 3.1. With the data transmitted and aggregated to the nodes close to the sink, the packets contain more information and the data loads of the nodes in the area close to the sink is becoming heavy. In order to prolong the network lifetime and ensure the transmission reliability, nodes in the area far to the sink send several more packet copies with unicast, and nodes close to the sink send fewer packet copies to reduce their energy cost. Taking into account that packets received by nodes close to the sink contain more information and the packets loss deteriorates the transmission reliability greatly, nodes close to the sink need to send a little more packet copies to ensure transmission reliability.

In the paper, the operation of the proposed scheme is illustrated in Fig.2. Data transmission starts from the outermost ring, and the source node in the outermost ring sends several packet copies to the next hop node after finding its next hop in the inner ring. As shown in Fig.2, the source node continuously transmits multiple copies of data packet P1 in the sending window with $\gamma = 4$. In the packet transmission procedure, some packets are lost due to the unreliable links, such as the 2^{nd} and 3^{rd} packets are lost in Fig.2. Although two packets are lost in the transmission, the transmission is successful if one packet arrives the receiver. There is no need for the receiving node to reply ACK. Then the intermediate receiving node does data aggregation. The receiving nodes in the inner ring do data aggregation sequentially. When all the nodes in the outer ring report their sensing data to the corresponding nodes in the inner ring, they begin their data transmission to the next hop nodes in their inner ring similarly. The scheme ensures the transmission reliability while reducing the transmission delay.

B. THE IMPPLEMENTATION OF THE PROPOSED SCHEME

In the section, the implementation of the proposed scheme is explained, including ring construction, relaying node selection and the number of packet copies determination.

(1) Ring partition and relaying node selection

Fig. 3 illustrates the partitioned ring and relaying node selection. As shown in Fig. 3, assume the width of the ring or tier is δr, the network is portioned into rings $\{T_k\}$, where $0 < \delta < 1$ to ensure each node in the outer ring finds its relaying node in the inner ring. For the node v_k^i in the ring T_k, the set of relaying nodes (donated RS) is composed of the nodes located within its sensing range and in the inner tier or ring T_{k-1}. That is $RS = \{v_{k-1}^j || v_{k-1}^j - v_k^i| \leq r\}$. The node v_k^i broadcasts a request message for the remaining energy of the relaying nodes in RS. Each node in the RS returns a packet containing its own remaining energy. After receiving the returned packets, the node v_k^i selects one node with the maximum residual energy as its next hop node in the transmission to the sink.

(2) The value of data copies γ determination

In the scheme, there is no need for each node to wait for the returned ACK information for the sending packets to reduce the transmission delay. As the above mentioned, one node unicast multiple packet copies to its next hop node in the proposed scheme to ensure transmission reliability. The probability of packets successfully transmitted from one node to its next hop node on the routing path is $\epsilon_i = 1-(1-p_{ij})^\gamma$. From this we can see that the more sending copies of one packet, the transmission reliability is higher, which will lead to more additional energy cost. For the nodes far to the sink, they have less packets to relay. Therefore, they have more remaining energy and could send more packets copies to ensure the successful transmission. Whereas, the nodes close to the sink tend to relay more size of data packets and the energy is depleted quickly. Thus, the number of sending packet copies should be small to prolong the network lifetime. However, the nodes close to the sink need to relay the aggregated packets, which contain more information. If the number of packet copies is too small, the packets loss will greatly worsen the transmission reliability. In order to prolong the network lifetime under the assurance of transmission reliability, it needs to adaptively determine an optimized γ.

In the work, the number of sending packet copies depends on two factors including distance to the sink (Dis for simplification) and transmission reliability request (Re). A typical fuzzy logic system is created for making decision to compute γ. It firstly maps the crisp inputs into linguistic variables by membership functions. Then, it uses a series of IF-THEN rules to inference the inputs to the outputs represented by linguistic variables. A crisp output value from the fuzzy space is obtained by the defuzzifier.

1) DISTANCE TO THE SINK

For the nodes far to the sink, they have less packets to relay. Therefore, they have more remaining energy and could send more data copies to ensure the successful transmission. Whereas, the nodes close to the sink tend to relay more size of data packets and the energy is depleted quickly. Thus, the number of sending packet copies is small to prolong the network lifetime.

The distance to the sink is defined as:

$$Dis(i) = \frac{d_{i,sink}}{R} \tag{5}$$

where, d is the Euclidean distance from the node v_k^i to the sink, R is the radius of the circular network. The value range of $Dis\,(i)$ is (0, 1]. Nodes farthest to the Sink will have the maximum number of sending packet copies. As illustrated in Fig. 2, the source node S in the outmost ring have the maximum γ.

2) TRANSMISSION RELIABILITY LEVEL

The transmission reliability level is defined as the requested transmission reliability. That is:

$$Re = \epsilon \tag{6}$$

where, ϵ represents the request of QoS level transmission reliability. The value range is (0,1). If the request transmission reliability denoted by Re is higher, the γ is greater.

Although in the data aggregation strategy in this paper, multiple packets from the same source node arrive at the receiving node and the receiving node saves one packet and deletes the rest, it aggregates the receiving one packet with its original (or intermediate aggregation result) into one packet with more size and contains more information. Therefore, a sensor with bigger receiving packets should send more greater γ copies of aggregated packets to avoid the more information loss and ensure the transmission reliability. In the process of data collection, this will cause the nodes close to the sink tend to have a large amount of data load, correspondingly, the energy consumption is depleted quickly.

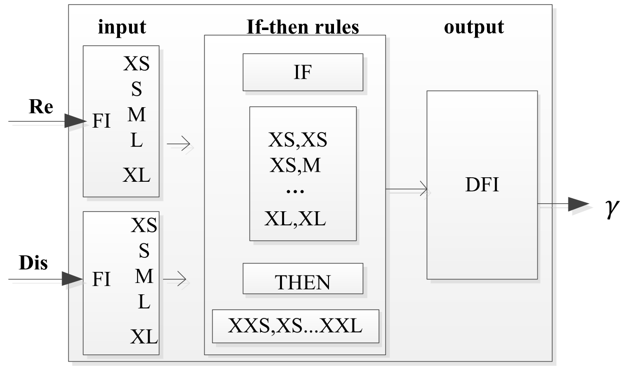


FIGURE 4. Two inputs and one output fuzzy logic system.

And, nodes far away from Sink have less data load and less energy consumption.

The γ value of the sending packet copies of each node is inferenced applying a fuzzy logic system as shown in Fig. 4. The fuzzy-logic based γ determination consists of two inputs and one output. The inputs include the request transmission reliability *Re* and nodes distance to the sink *Dis*. The output of the number of the sending packet copies γ is inferenced according to the inputs. If the source or intermediate nodes have longer distance *Dis* to the sink and the high *Re*, they would have more sending data copies. Otherwise, there are less data copies to be transmitted. More detailed fuzzy inference could be seen in the simulation section.

Therefore, under the condition of guaranteeing transmission reliability, reducing node energy consumption is the core problem to be solved in this paper. In the proposed scheme, the number of the unicasting packet copies of each node is decided according to the node distance to the sink and the request transmission reliability. The fuzzy logic system is explored to inference the result. The detailed implementation of the proposed scheme is explained in algorithm 1.

C. PERFORMANCE ANALYSIS OF THE PROPOSED SCHEME

The energy efficient and reliable in-network data aggregation with unicast scheme ensures the energy efficiency under the guarantee of network transmission reliability. Packet copies in the configurable sending window are unicasted to the next hop node and there is no need to wait for the ACK confirmation for sending nodes. The key idea of the manuscript is to adaptively determine the unicasting number of packet copies γ to improve the network lifetime and satisfy the transmission reliability. In this section, the performance is discussed including conditions under which the proposed scheme can be effective.

1) DATA LOADS AND ENERGY COST

Theorem 1: Assume the total ring of the network is $\aleph$ and the initiated packet of each node is with the same size of τ (bits). After one round of data aggregation with unicast to the sink,

Algorithm 1 Procedure of the Proposed Scheme

Begin
//ϵ:the network reliability
//p: the one hop transmission reliability
// $\aleph$: the total number of rings
For $k = \aleph$:-1:1
  for each node v_k^i in T_k
    sets up RS in inner T_{k-1}
     selects one node in RS with the maximum remaining energy as its next hop node
     fuzzy inferences the sending packets copies γ according to Dis and Re in the fuzzy space
      unicasting the γ copies to the next hop node continuously in the window
       the next hop node stores only one packet from the same node and does aggregation sequentially
   end for
End For
End

the data loads of node v_{k-1}^j is calculated by:

$$D_{k-1,r}^j = \gamma_k(1-(1-p)^{\gamma_k})\sum_{\sigma=1}^{\sigma=\alpha}\beta^{\alpha+1-\sigma}\mu_k^\sigma \quad (7)$$

$$D_{k-1,s}^j = \gamma_{k-1}\mu_{k-1}^j \quad (8)$$

where, γ_k represents the packet copies for nodes in the ring k need to send and a denotes the number of receiving packets from different nodes in sequence, and $\mu_{k-1}^i = \beta^\alpha\tau + \gamma_k(1-(1-p)^{\gamma_k})\sum_{\sigma=1}^{\sigma=\alpha'}\beta^{\alpha+1-\sigma}\mu_k^\sigma$.

Proof: Under the assumption that the total ring of the network is $\aleph$ and the initiated packet size of each node is τ bits. For the node in ring $\aleph$,

$$\mu_\aleph^i = \tau \quad (9)$$

For the node $v_{\aleph-1}^j$ in ring $\aleph-1$, if it sequentially receives a_1 packets from ring $\aleph$, the receiving data loads is computed by:

$$D_{\aleph-1,r}^j = \sum_{i=1}^{a_1}\gamma_\aleph(1-(1-p)^{\gamma_\aleph})\mu_\aleph^i \quad (10)$$

Similarly, when node $v_{\aleph-2}^j$ in ring $\aleph-2$ has a_2 different packets received, the data loads is obtained by:

$$D_{\aleph-2,r}^j = \sum_{i=1}^{a_2}\gamma_{\aleph-1}(1-(1-p)^{\gamma_{\aleph-1}})\mu_{\aleph-1}^i \quad (11)$$

$\mu_{\aleph-1}^i$ is calculated by analogy. Following the data aggregation model, if it has one packet to receive from the outer ring for node $v_{\aleph-1}^i$, the following is obtained:

$$\varphi_{\aleph-1}^{i,1} = \beta\,[\varphi_{\aleph-1}^{i,0} + \gamma_\aleph\left(1-(1-p)^{\gamma_\aleph}\right)\mu_\aleph^1] \quad (12)$$

If it has two packets to sequentially receive, then

$$\begin{aligned}\varphi_{\aleph-1}^{i,2} &= \beta\left(\varphi_{\aleph-1}^{i,1} + \gamma_\aleph\left(1-(1-p)^{\gamma_\aleph}\right)\mu_\aleph^2\right)\\ &= \beta 6(\beta(\varphi_{\aleph-1}^{I,0} + \gamma_\aleph\left(1-(1-p)^{\gamma_\aleph}\right)\mu_\aleph^1)\\ &\quad + p\gamma_\aleph\left(1-(1-p)^{\gamma_\aleph}\right)\mu_\aleph^2)\\ &= \beta^2(\varphi_{\aleph-1}^{i,0} + \gamma_\aleph\left(1-(1-p)^{\gamma_\aleph}\right)\mu_\aleph^1)\\ &\quad + \beta\gamma_\aleph\left(1-(1-p)^{\gamma_\aleph}\right)\mu_\aleph^2\\ &= \beta^2\tau + \gamma_\aleph\left(1-(1-p)^{\gamma_\aleph}\right)[\beta^2\mu_\aleph^1 + \beta\mu_\aleph^2] \quad (13)\end{aligned}$$

Similarly, the final aggregation packet size for node v_{k-1}^i is:

$$\begin{aligned}\mu_{k-1}^i &= \beta^\alpha\tau + \gamma_k\left(1-(1-p)^{\gamma_k}\right)\\ &\quad\times\left(\beta^a\mu_k^1 + \beta^{\alpha-1}\mu_k^2 + \cdots + \beta\mu_k^j\right)\\ &= \beta^\alpha\tau + \gamma_k(1-(1-p)^{\gamma_k})\sum_{\sigma=1}^{\sigma=a}\beta^{\alpha+1-\sigma}\mu_k^\sigma \quad (14)\end{aligned}$$

where, a represents the number of node v_{k-1}^i different receiving packets. Thus, the data load of each node is calculated as the following:

$$D_{k-1,r}^j = \gamma_k(1-(1-p)^{\gamma_k})\sum_{\sigma=1}^{\sigma=\alpha}\beta^{\alpha+1-\sigma}\mu_k^\sigma \quad (15)$$

$$D_{k-1,s}^j = \gamma_{k-1}\mu_{k-1}^j \quad (16)$$

■

On the basis of the results obtained above and the typical energy consumption model described in equation (1), the energy cost on sending and receiving the data packet for v_{k-1}^j can be calculated respectively by the following equations.

$$E_{k-1,r}^j = D_{k-1,r}^j E_{elec} \quad (17)$$

$$E_{k-1,s}^j = D_{k-1,s}^j(E_{elec} + \varepsilon_\theta d^\beta) \quad (18)$$

Therefore, the energy cost for node v_{k-1}^j is:

$$E_{k-1}^j = E_{k-1,r}^j + E_{k-1,t}^j \quad (19)$$

After the specific parameters are replaced and further simplified, the following result is obtained.

$$\begin{aligned}E_{k-1}^j &= \gamma_k\left(1-(1-p)^{\gamma_k}\right)E_{elec}\sum_{\sigma=1}^{\sigma=\alpha}\beta^{\alpha+1-\sigma}\mu_k^\sigma\\ &\quad + \gamma_{k-1}\mu_{k-1}^j(E_{elec} + \varepsilon_\theta d^\beta) \quad (20)\end{aligned}$$

Although the final mathematical expressions are complicated, it is clear to see that the energy cost is greater with the bigger value of γ_k under other given the parameters. For the packet transmission with broadcast, if the one hop number is greater than γ_k, it consumes more energy.

2) DELAY

Packet transmission with broadcast is more energy cost. Although stop-and-wait ARQ scheme is highly energy efficient, it needs more time for waiting the ACK and even retransmitting the lost packet. Assume one hop transmission time is h, for the SW the statistical one hop time needed for successful transmitting one packet is $t_{sw} \geq ph + (1-p)(2h+h) = 3h(1-p)$. However, for the proposed scheme, the transmission time for one hop is $t = h$. Thus, we get the following:

$$t_{sw} - t \geq 2h - 2ph = 2h\,(1-p) > 0 \quad (21)$$

For the lossy wireless sensor networks, $p < 1$, so $t_{sw} - t \geq 2\tau\,(1-p) > 0$, that is $t_{sw} > t$.

Algorithm 2 Finding the Optimized γ_u

1: **Begin**
2: //ϵ: the request network transmission reliability
3: // $\aleph$: the total number of rings
4: // E_{max}: the maximum energy cost of one node
5: $E_{max} = \text{inf}$, $E_k^{max} = 0$
6: $\gamma_u = \gamma_{\max}$;
7: **For** $k = \aleph{:}-1{:}1$
8: for each node v_k^j in T_k
9: calculate the energy cost E_k^j
10: if $E_k^j > E_k^{max}$
11: $E_k^{max} = E_k^j$
12: **end**
13: end for
14: **End For**
15: **If** $E_k^{max} < E_{max}$ && satisfying ϵ
16: $E_{max} = E_k^{max}$
17: $\gamma_{\max} = \gamma_{\max} - 1$
18: Go to step 6;
19: **Else**
20: **return** $\gamma_u = \gamma_u + 1$
21: **End**
22: **End**

3) FUZZY OUTPUT SPACE

In the proposed scheme, the fuzzy output is γ and we know that the minimum value of γ is 1. Therefore, the lower bound of the fuzzy output space is 1. If we use γ_{max} to denote the maximum value of γ and γ_u to represent the upper bound of the fuzzy output space which satisfies the request transmission reliability ϵ and has the lowest energy cost simultaneously. According to $1-(1-p)^{\gamma_{\max}} \leq \epsilon^{1/\aleph}$, the following can be obtained:

$$\gamma_{\max} \leq \frac{\log(1-\epsilon^{1/\aleph})}{\log(1-p)} \tag{22}$$

The enumeration method is applied to find the optimal γ_u to ensure the optimization goal. The detailed implementation is shown in Algorithm 2.

V. SIMULATION

In this section, simulations are explored to evaluate and prove the proposed scheme in the performance of extending the network lifetime under the constraint of the network reliability.

A. PARAMETER SETTINGS

The simulation is implemented on a two-dimensional circular wireless sensor network as described in Section 3. It consists of 200 sensor nodes. The radius is $R = 100$(m). In addition, the transmission range is assumed $r = 30$(m). The initiated packet size of each node in the network is fixed 200 bits. And we assume the one hop transmission reliability is $p = 0.7$ and the request network transmission reliability $\epsilon = 0.8$. The compression ratio $\beta = 0.8$ when receiving node implements

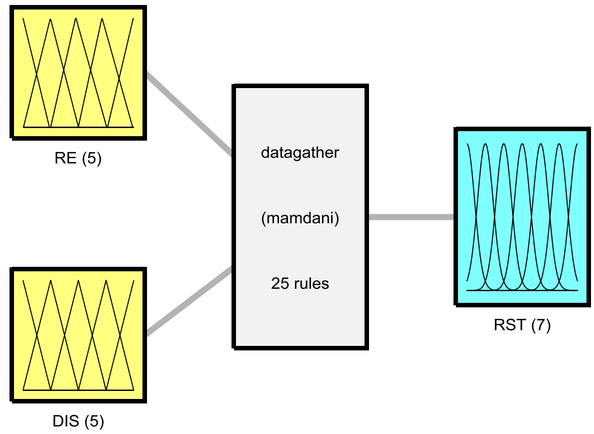


FIGURE 5. The fuzzy logic system to inference γ.

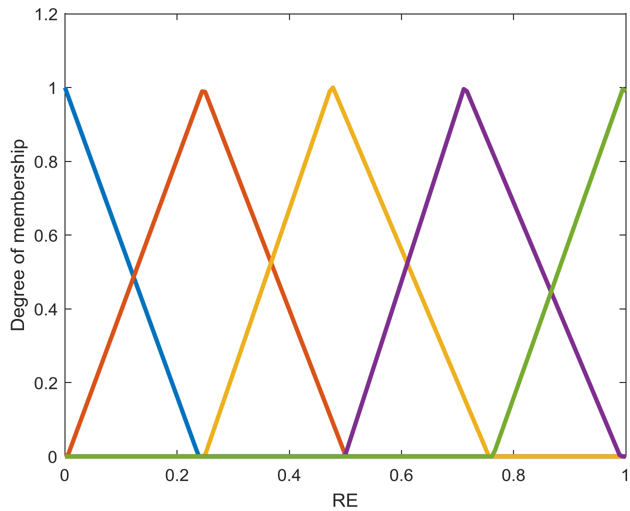


FIGURE 6. Membership function of input 'RE'.

data aggregation in sequential. All the nodes (except the sink) initialize the identical energy of 1 (J). Other parameters are shown in Tab. 1. Following (1) and (2), the energy cost of each node is calculated in the simulation based on the sending and receiving data loads. MATLAB is applied to run simulation 1000 times and the average results are recorded.

As shown in Fig.5, the configurable γ is determined by the fuzzy logic system including two input factors of request transmission reliability and distance to the sink. In the fuzzification, the linguistic terms are used to represent the fuzzy sets. The linguistic terms include extreme, extreme small (XXS), extreme small (XS), small (S), medium (M), large (L), extreme large (XL) and extreme, extreme large (XXL). For the input variables, the fuzzy sets are represented $Re = \{XS, S, M, L, XL\}$ and $Dis = \{XS, S, M, L, XL\}$ respectively. The inferenced output value of γ is represented as RST $= \{XXS, XS, S, M, L, XL, XXL\}$. The simple triangular membership function shown in Fig. 6 and Fig. 7 are applied to convert the crisp input values of Re and Dis to corresponding fuzzy sets. From the figures, the domain of all inputs

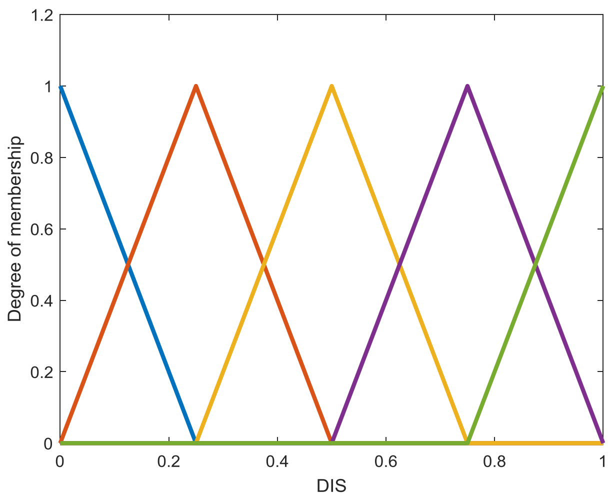


FIGURE 7. Membership function of input 'DIS'.

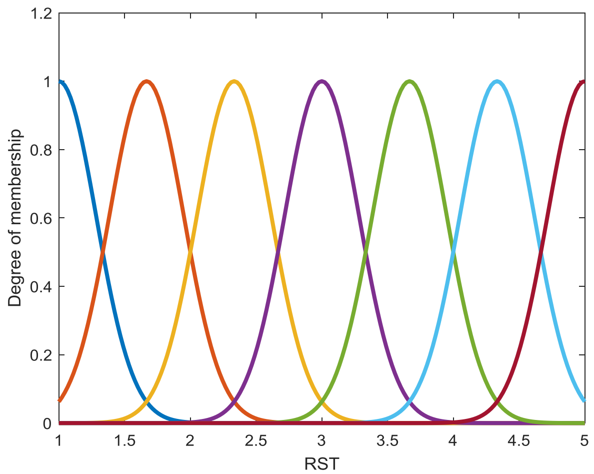


FIGURE 8. Membership function of output 'RST'.

TABLE 2. Fuzzy rules.

Re/Dis	XS	S	M	L	XL
XS	XXS	XXS	XS	XS	XS
S	XS	S	S	S	S
M	XS	S	M	M	M
L	XS	M	L	L	XL
XL	S	M	L	XL	XXL

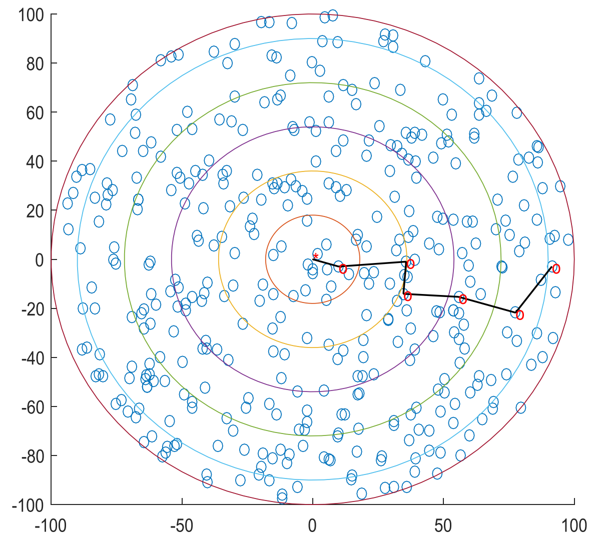


FIGURE 9. An illustration of one transmission path from one node in the outmost ring to the sink.

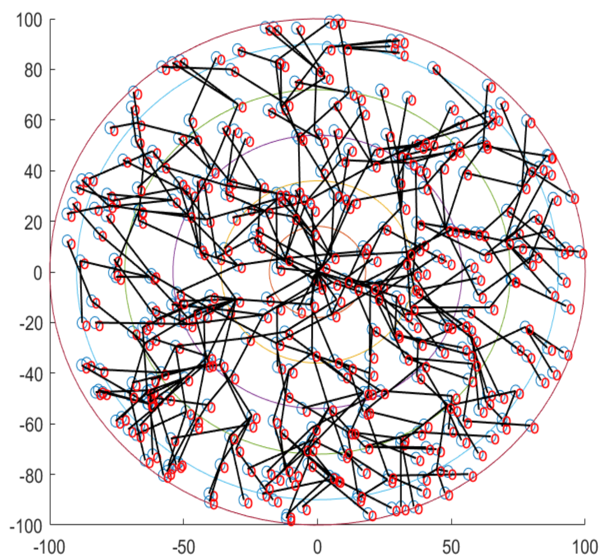


FIGURE 10. An illustration of all routing paths for all nodes.

is [0, 1]. The gauss membership functions shown in Fig. 8 are used for converting the output variable RST. Algorithm 2 is applied to find the upper bound value of the fuzzy output space which satisfies the request transmission reliability ϵ and has the lowest energy cost simultaneously. The fuzzy rules shown in Tab. 2 are applied to process the fuzzy values in the fuzzy inference phase. Besides, And method(min), Or method (max), Implication (min), Aggregation (probor) are applied in the fuzzy inference. The crisp output value γ is obtained in the defuzzification using the centroid method.

B. PERFORMANCE EVALUATION

1) DIFFERENT SENDING COPIES γ

In this section, the influence of the sending copies γ on the performance of the proposed scheme is evaluated. When the nodes with different distance to the sink sending the identical number of packets copies γ, the simulations on performance metrics of statistical reliability and energy cost are presented.

As we know that node v_k^i in ring k selects one node in ring $k-1$ with the maximum residual energy as its next hop node, the packet routing path may be different in different data gathering rounds. In the simulation, an illustration of packets

TABLE 3. Performance under different γ.

γ	reliability	Maximum energy cost (10^{-3} J)
2	67.34%	0.3149
3	89.50%	0.5467
4	96.04%	0.7211
5	98.76%	0.8628

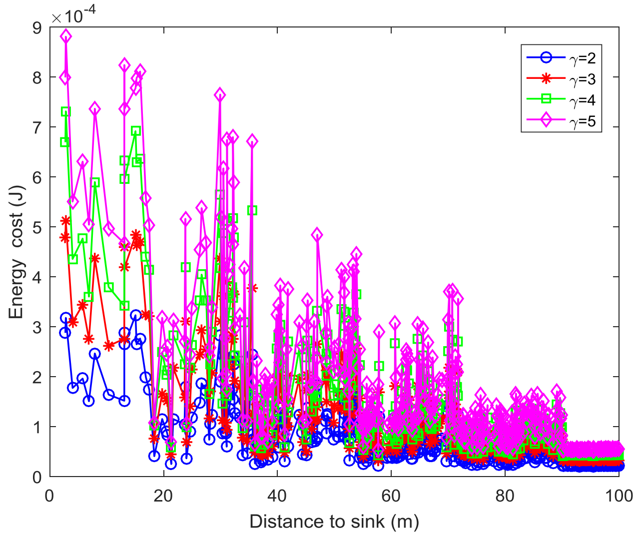


FIGURE 11. Comparison of energy consumption of each node with different sending packet copies γ in one round of data aggregation to the sink.

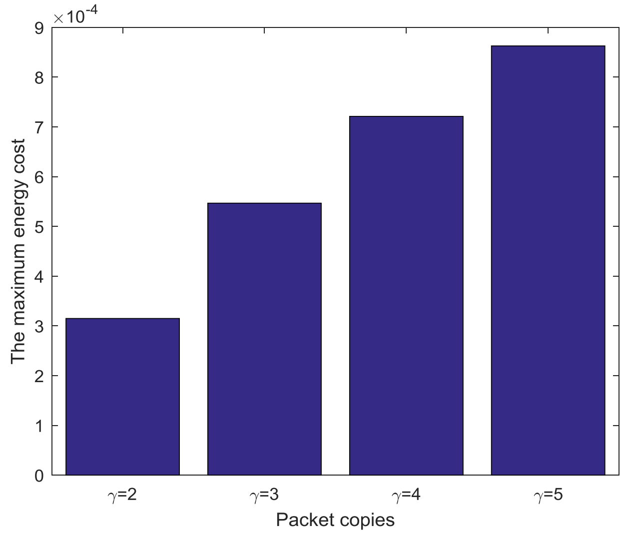


FIGURE 12. Comparison of the maximum energy cost with different sending packet copies γ after one round of data aggregation to the sink.

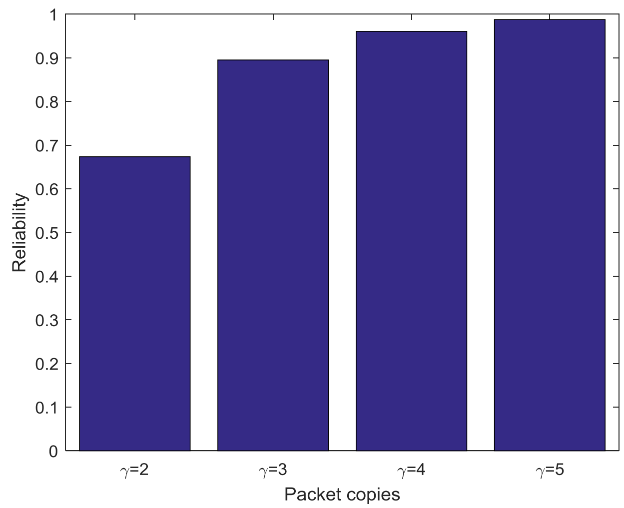


FIGURE 13. Comparison of transmission reliability with different sending packet copies γ after one round of data aggregation to the sink.

transmission path from one source node in the outmost ring to the sink following the proposed scheme is shown in Fig. 9. Fig.10 shows the packets routing paths of all nodes following the routing finding. When different γ packet copies are unicasted, the simulation results and comparison on statistical reliability and energy cost are demonstrated in Fig.11- Fig.13. To demonstrate the results more clearly, the detailed values are shown in Tab. 3. Fig.11 shows the comparison of energy cost of each node after one round of data aggregation to the sink under different γ values. From Fig.11, we can see that overall there is more energy cost for nodes close to the sink. Because the nodes need to receive and forward the bigger and bigger aggregated packets. Besides, the larger value of γ causes larger energy cost for source and intermediate nodes, because a larger γ indicates more sending data copies in the transmission. More concretely, the energy cost when $\gamma = 3$ is larger than that $\gamma = 2$ in general. When $\gamma = 5$, the node energy cost is overall the largest compared with $\gamma = 2$, $\gamma = 3$ and $\gamma = 4$. It is clearly illustrated by Fig.12. Fig.12 shows the maximum energy cost comparison with different unicasting packet copies γ after one round of data aggregated to the sink. From the figure, it can be seen that the bigger γ leads to larger the maximum node energy cost. Correspondingly, the network lifetime is shortened with bigger γ. However, the bigger γ could provide higher transmission reliability, which is demonstrated in Fig.13. As shown in the figure, the network reliability is 98.76%, 96.04%, 89.50% and 67.34% when $\gamma = 5, \gamma = 4, \gamma = 3$ and $\gamma = 2$. As mentioned in the parameter settings, the request transmission reliability is ϵ =80%. This means that $\gamma \geq 3$ could meet the request transmission reliability. However, the larger γ will result in excess energy consumption.

2) COMPARISON WITH BROADCAST

In the part, the proposed scheme is compared with the method of data aggregation with broadcast over lossy links in [14] under the case of the same network ring partition. Fig. 14 shows an illustration of packets broadcasting from one source node to the sink and Fig.15 illustrates the broadcasting

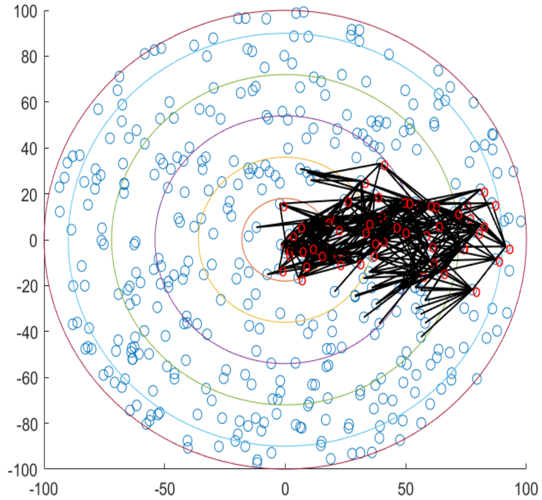


FIGURE 14. An illustration of broadcast from one node in the outmost ring to the sink.

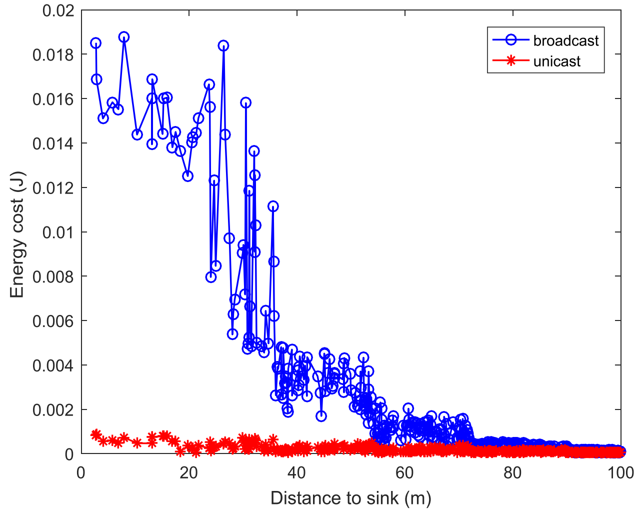


FIGURE 16. Energy cost comparison of each node after one round of data aggregation to the sink.

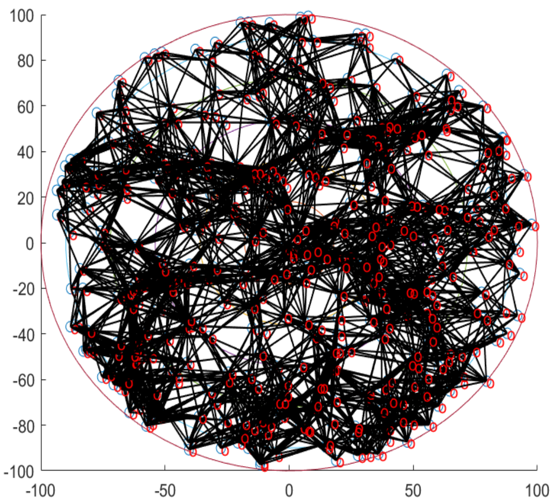


FIGURE 15. An illustration of broadcast of all nodes in the network.

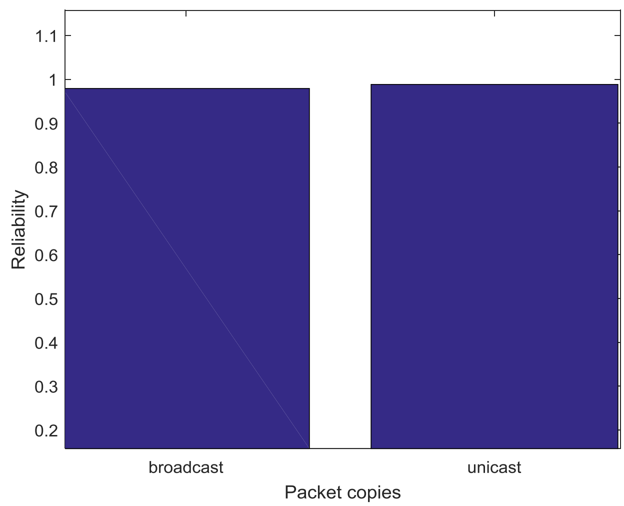


FIGURE 17. Transmission reliability comparison between broadcast and unicast ($\gamma = 5$).

of all nodes in the network. From the figure, we can see that packets broadcasting causes many nodes participating in the data transmission, which leads to additional energy cost. This also can be seen from Fig. 16. The energy consumption comparison of each node is illustrated in Fig.16. In the figure, we compare the packets broadcasting with the case in that each node unicasts $\gamma = 5$ packet copies to its next hop node. From Fig.16, it can be seen that the energy cost of packets broadcasting is far greater than packets unicasting with $\gamma = 5$ packet copies. However, there is no big difference in the transmission reliability, which can be seen from Fig.17.

C. PERFORMANCE EVALUATION ON γ INFERENCED BY FUZZY LOGIC

In this part, configurable γ inferenced by fuzzy logic is evaluated. In order to illustrate the effectiveness of the variable γ, comparisons on energy cost and transmission reliability with invariable γ are demonstrated and discussed. To further illustrate the effectiveness, we also compared the network lifetime in rounds.

Although packets that are continuously unicast with multiple copies have superiority in energy cost, it does not consider the imbalance of data loads and energy cost in the network. The advance of the proposed scheme is that it could automatically adjust the number of sending packet copies for each node according to nodes distance and the request transmission reliability. It could reduce the energy cost of nodes close to the sink as much as possible. According to algorithm 2, the optimized upper bound of fuzzy output is 5. Therefore, the output space ranges from 1 to 5.

The detailed comparison results are shown in Tab. 4. Fig.18 shows the comparison of each node energy cost after one round of data aggregation under different cases.

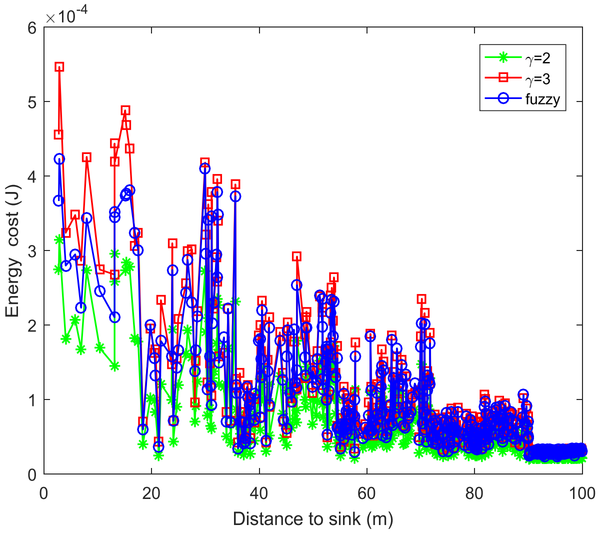


FIGURE 18. **Energy cost comparison of each node after one round of data aggregation to the sink.**

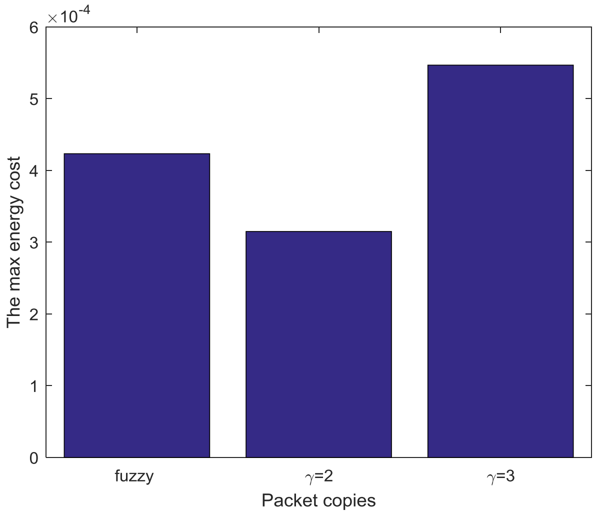


FIGURE 19. **The maximum energy cost comparison after one round of data aggregation to the sink.**

TABLE 4. **Comparisons result.**

parameter s	γ=2	fuzz y	γ=3
Maximum energy cost	$3.1491*10^{-4}$ (J)	$4.2332*10^{-4}$ (J)	$5.4674*10^{-4}$ (J)
Reliability	67.43%	82.58%	89.50%
Lifetime	47.08 (rounds)	41.21 (rounds)	35.11(rou nds)

To clearly state the comparison, Fig. 19 shows the maximum energy cost comparison. The network transmission reliability comparison is illustrated in Fig. 20. From the table and figures, it can be seen that it is failure to meet the request transmission reliability under the fixed $\gamma = 2$. And, there is

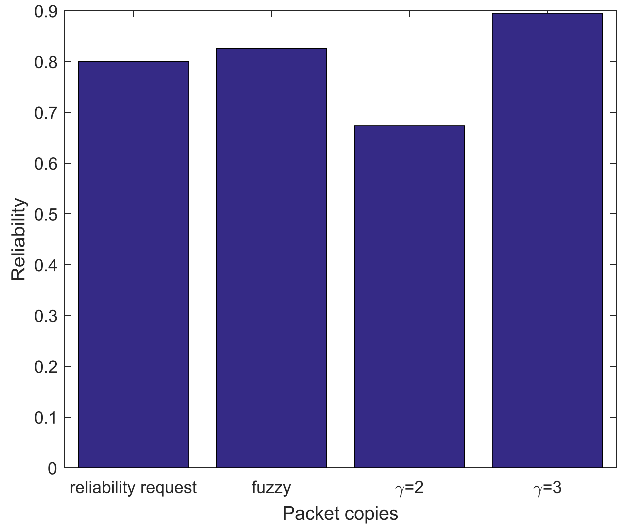


FIGURE 20. **Transmission reliability comparison.**

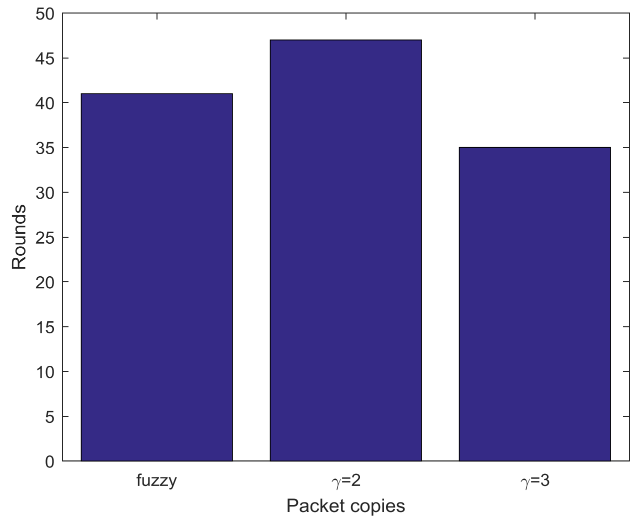


FIGURE 21. **Network lifetime comparison (in rounds).**

less energy cost for the configurable γ based on fuzzy logic compared with fixed $\gamma = 3$ under the guarantee of the request network transmission reliability $\epsilon = 80\%$. Fig. 21 shows the network lifetime by rounds under the cases. From the figure, we can see that the network lifetime is improved by 17.14% compared with fixed $\gamma = 3$. The above results show the effectiveness of the proposed scheme in extending the network lifetime under the guarantee of the request network transmission reliability.

VI. CONCLUSION

In this paper, the problem of energy efficient data gathering with in-network aggregation is studied in consideration of guaranteeing the network transmission reliability for WSN. A novel ring and fuzzy rule-based data aggregation scheme is proposed in the paper to improve the energy efficiency under the assurance of the request transmission reliability.

In the proposed scheme, the variable number of sending packet copies γ is inferenced by a fuzzy logic system to achieve an efficient data aggregation. It is adaptively tuned according to the request network transmission reliability and nodes distance to the sink. To prove its effectiveness, theoretical analysis and simulation are implemented. However, the fuzzy output space and membership functions influence on performance are worth further research.

REFERENCES

[1] J. M. Perkel, "The Internet of Things comes to the lab," *Nature*, vol. 542, no. 7639, pp. 125–126, 2017.

[2] P. Mach and Z. Becvar, "Mobile edge computing: A survey on architecture and computation offloading," *IEEE Commun. Surveys Tuts.*, vol. 19, no. 3, pp. 1628–1656, 3rd Quart., 2017.

[3] X. Liu *et al.*, "Construction of large-scale low-cost delivery infrastructure using vehicular networks," *IEEE Access*, vol. 6, pp. 21482–21497, 2018.

[4] Y. Kim, S. Atchley, G. R. Vallée, S. Lee, and G. M. Shipman, "Optimizing end-to-end big data transfers over terabits network infrastructure," *IEEE Trans. Parallel Distrib. Syst.*, vol. 28, no. 1, pp. 188–201, Jan. 2017.

[5] T. G. Rodrigues, K. Suto, H. Nishiyama, and N. Kato, "Hybrid method for minimizing service delay in edge cloud computing through VM migration and transmission power control," *IEEE Trans. Comput.*, vol. 66, no. 5, pp. 810–819, May 1, 2017.

[6] S. Singh, N. Saxena, A. Roy, and H. Kim, "A survey on 5G network technologies from social perspective," *IETE Tech. Rev.*, vol. 34, no. 1, pp. 30–39, 2017.

[7] S. Sarkar, S. Chatterjee, and S. Misra, "Assessment of the suitability of fog computing in the context of Internet of Things," *IEEE Trans. Cloud Comput.*, vol. 6, no. 1, pp. 46–59, Jan./Mar. 2015.

[8] Cisco. *Internet of Things Market Forecast*. Accessed: Aug. 20, 2018. [Online]. Available: http://postscapes.com/internet-of-things-market-size

[9] T. Qiu, R. Qiao, M. Han, A. K. Sangaiah, and I. Lee, "A lifetime-enhanced data collecting scheme for the Internet of Things," *IEEE Commun. Mag.*, vol. 55, no. 11, pp. 132–137, Nov. 2017.

[10] X. Liu, W. Liu, Y. Liu, H. Song, A. Liu, and X. Liu, "A trust and priority based code updated approach to guarantee security for vehicles network," *IEEE Access*, vol. 6, pp. 55780–55796, 2018.

[11] S. Rani, S. H. Ahmed, R. Talwar, and J. Malhotra, "Can sensors collect big data? An energy-efficient big data gathering algorithm for a WSN," *IEEE Trans. Ind. Informat.*, vol. 13, no. 4, pp. 1961–1968, Aug. 2017.

[12] V. K. Singh, V. K. Singh, and M. Kumar, "In-network data processing based on compressed sensing in WSN: A survey," *Wireless Pers. Commun.*, vol. 96, no. 2, pp. 2087–2124, 2017.

[13] T. Li, N. N. Xiong, J. Gao, H. Song, A. Liu, and C. Zhiping, "Reliable code disseminations through opportunistic communication in vehicular wireless networks," *IEEE Access*, vol. 6, pp. 55509–55527, 2018.

[14] C. Joo and N. B. Shroff, "On the delay performance of in-network aggregation in lossy wireless sensor networks," *IEEE/ACM Trans. Netw.*, vol. 22, no. 2, pp. 662–673, Apr. 2014.

[15] Y. Liu, A. Liu, and Z. Chen, "Analysis and improvement of send-and-wait automatic repeat-reQuest protocols for wireless sensor networks," *Wireless Pers. Commun.*, vol. 81, no. 3, pp. 923–959, 2015.

[16] Q. Ju and Y. Zhang, "Clustered data collection for Internet of batteryless things," *IEEE Internet Things J.*, vol. 4, no. 6, pp. 2275–2285, Dec. 2017.

[17] P. G. V. Naranjo, E. Baccarelli, and M. Scarpiniti, "Design and energy-efficient resource management of virtualized networked Fog architectures for the real-time support of IoT applications," *J. Supercomput.*, vol. 74, no. 6, pp. 2470–2507, 2018.

[18] S. Chen, K. Wang, C. Zhao, H. Zhang, and Y. Sun, "Accelerated distributed optimization design for reconstruction of big sensory data," *IEEE Internet Things J.*, vol. 4, no. 5, pp. 1716–1725, Oct. 2017.

[19] Y. Kawamoto, H. Nishiyama, N. Kato, Y. Shimizu, A. Takahara, and T. Jiang, "Effectively collecting data for the location-based authentication in Internet of Things," *IEEE Syst. J.*, vol. 11, no. 3, pp. 1403–1411, Sep. 2017.

[20] B. Pourghebleh and N. J. Navimipour, "Data aggregation mechanisms in the Internet of Things: A systematic review of the literature and recommendations for future research," *J. Netw. Comput. Appl.*, vol. 97, pp. 23–34, Nov. 2017.

[21] Q. Zhang, C. Zhu, L. T. Yang, Z. Chen, L. Zhao, and P. Li, "An incremental CFS algorithm for clustering large data in industrial Internet of Things," *IEEE Trans. Ind. Informat.*, vol. 13, no. 3, pp. 1193–1201, Jun. 2017.

[22] R. Mohanasundaram and P. S. Periasamy, "Clustering based optimal data storage strategy using hybrid swarm intelligence in WSN," *Wireless Pers. Commun.*, vol. 85, no. 3, pp. 1381–1397, 2015.

[23] D. Virmani, G. Mandal, N. Beniwal, and S. Talwar, "Dynamic data aggregation tree for data gathering in wireless sensor network," *Int. J. Innov. Technol. Exploring Eng.*, vol. 2, no. 3, pp. 226–230, 2013.

[24] A. Sinha and D. K. Lobiyal, "A multi-level strategy for energy efficient data aggregation in wireless sensor networks," *Wireless Pers. Commun.*, vol. 72, no. 2, pp. 1513–1531, 2013.

[25] P. Zhang, G. Xiao, and H.-P. Tan, "Clustering algorithms for maximizing the lifetime of wireless sensor networks with energy-harvesting sensors," *Comput. Netw.*, vol. 57, no. 14, pp. 2689–2704, Oct. 2013.

[26] P. G. V. Naranjo, M. Shojafar, H. Mostafaei, Z. Pooranian, and E. Baccarelli, "P-SEP: A prolong stable election routing algorithm for energy-limited heterogeneous fog-supported wireless sensor networks," *J. Supercomput.*, vol. 73, no. 2, pp. 733–755, 2017.

[27] A. E. A. A. Abdulla, H. Nishiyama, J. Yang, N. Ansari, and N. Kato, "HYMN: A novel hybrid multi-hop routing algorithm to improve the longevity of WSNs," *IEEE Trans. Wireless Commun.*, vol. 11, no. 7, pp. 2531–2541, Jul. 2012.

[28] Y. Liu, Y. Zhu, L. M. Ni, and G. Xue, "A reliability-oriented transmission service in wireless sensor networks," *IEEE Trans. Parall. Distrib. Syst.*, vol. 22, no. 12, pp. 2100–2107, Dec. 2011.

[29] Z. Rosberg, R. P. Liu, T. Le Dinh, Y. F. Dong, and S. Jha, "Statistical reliability for energy efficient data transport in wireless sensor networks," *Wireless. Netw.*, vol. 16, no. 7, pp. 1913–1927, 2010.

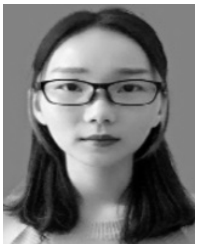
[30] P. Gupta and P. R. Kumar, "The capacity of wireless networks," *IEEE Trans. Inf. Theory*, vol. 46, no. 2, pp. 388–404, Mar. 2000.

[31] V. Mhatre and C. Rosenberg, "Design guidelines for wireless sensor networks: Communication, clustering and aggregation," *Ad Hoc Netw.*, vol. 2, no. 1, pp. 45–63, Jan. 2004.

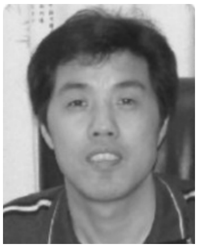
[32] J. Zhang, J. Long, G. Zhao, and H. Zhang, "Minimized delay with reliability guaranteed by using variable width tiered structure routing in WSNs," *Int. J. Distrib. Sensor Netw.*, vol. 2015, Jan. 2015, Art. no. 4.

JINHUAN ZHANG received the Ph.D. degree in communication and information systems from the Wuhan University of Technology in 2009. Since 2009, she has been an Assistant Professor with the School of Information Science and Engineering, Central South University. Her research interest includes wireless sensor networks, resource management, and IoT.

PENG HU received the Ph.D. degree in electrical engineering from Queen's University, Canada, in 2013. He is currently a Professor with the School of Information and Communications Technology, Seneca College of Applied Arts and Technology, Canada, and a Licensed Professional Engineer (P.Eng.), ON, Canada. His research interests include industrial Internet of Things, 5G, and low-power wireless networks. He served on the organizing and technical boards/committees of some industry consortia and international conferences, including AllSeen Alliance, DASH7, IEEE PIMRC'17, and IEEE AINA'15. He serves as a member on the IEEE Sensors Standards committee and serves as an Associate Editor for the *Canadian Journal of Electrical and Computer Engineering*.



FANG XIE is currently pursuing the master's degree with the School of Information Science and Engineering, Central South University, China. Her research interest is crowd sensing networks and wireless sensor networks.



JUN LONG received the M.Sc. and Ph.D. degrees from Central South University, China. He was a Visiting Professor with Columbia University from 2012 to 2013. He is currently a Professor with the School of Information Science and Engineering, Central South University.

AN HE is currently pursuing the Ph.D. degree with the School of Information Science and Engineering, Central South University, China. Her research interest is wireless sensor networks.