

Research Article

Node Classification Based on Functionality in Energy-Efficient and Reliable Wireless Sensor Networks

Ning Sun,¹ Youngbuk Cho,² and Sangho Lee²

¹ Department of Computer Science, Chungbuk National University, 52 Naesudong-ro, Cheongju 361-763, Republic of Korea

² Department of Software Engineering, Chungbuk National University, 52 Naesudong-ro, Cheongju 361-763, Republic of Korea

Correspondence should be addressed to Sangho Lee, shlee@cbnu.ac.kr

Received 27 May 2012; Revised 24 September 2012; Accepted 30 November 2012

Academic Editor: Regina B. Araujo

Copyright © 2012 Ning Sun et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Energy efficiency and reliability are two important factors when we design a wireless sensor network (WSN). In this paper, we proposed an energy-efficient and reliable hierarchical mechanism for WSNs. We utilized clustering structure to decrease energy consumption. Because cluster head plays an important role in WSN, we designed a substitute of cluster head (SCH) in order to guarantee the data propagation reliability. Based on different functionalities of sensor nodes we creatively classify sensor nodes into five categories: branch node (BN), cluster head (CH), substitute of cluster head (SCH), intermedium node (IN), and normal node. By adopting different MAC and network strategies for different node categories, energy consumption can be greatly reduced and data reliability can be significantly increased in the proposed WSN mechanism. The simulation results showed that this proposed mechanism realized the objective in aspects of energy efficiency and reliability.

1. Introduction

Wireless sensor networks consist of hundreds to thousands of sensor nodes. These sensor nodes are usually homogenous and one or more resource-rich base stations exist in WSNs which is the medium between WSNs and the outside information systems. The sensor nodes collect interest data through sensing the environment parameters such as temperature, wind speed, and humidity, then propagate the sensed data to the base station for further processing. Nowadays, WSN technology has been considered mature enough for real-life application such as control and automation, environment monitoring, healthcare, security and surveillance, smart home and building, vehicle tracking and detection, and precision agriculture.

Due to the small size and low price, most sensor nodes have constraints in processors, memory size, communications, and energy capability. Among these limitations, energy saving is always a remarkable issue and has got much attention for designing the WSN protocol. Amounts of protocols in MAC and network layers have been proposed with the purpose of saving energy [1, 2]. Meanwhile, strict energy constraints restrict long and reliable performance of sensor nodes. Hardware often cannot work regularly due

to energy depletion. At the same time, sensor networks are often deployed in harsh and hazardous environments, which affect normal operation of sensor nodes. The wireless radios of sensor nodes are severely affected by these environment factors. Communication faults often occur in WSNs due to the interferences. Therefore, reliability is another highlighted challenge for the real-life applications in WSNs. It is necessary for the WSN system to maintain high fault tolerant capability.

Many literatures [2–5] have shown that hierarchical structure is an efficient means to save energy in WSNs. Hierarchical structure can greatly reduce the amount of direct communications between sensor nodes and base station. It is concluded [3] that the energy consumption of cluster head is much higher than normal nodes due to the extra works. By electing cluster heads in rotation the energy consumption can be balanced among nodes and the total energy consumption can be significantly reduced such that network lifetime is prolonged. In this paper we utilized the clustering hierarchical structure.

Besides the hierarchical structure, energy-efficient MAC protocols [1] also achieve energy savings by controlling the radio wake/sleep and arranging reasonable access schedule

to reduce energy waste. MAC protocols for WSNs can be categorized into two main groups: schedule-based and contention-based protocols. In our proposed mechanism, for different category of nodes and in different phases we utilized different MAC protocols flexibly to reduce the energy consumption. By turning off radio of sensor nodes when they are idle in WSNs, our proposed mechanism maximized energy saving.

Although cluster head plays an important role in hierarchical WSNs, most literatures did not consider the reliability of cluster head. Cluster head has to keep awake in the whole process, from cluster formation, data collection, and data aggregation to data propagation. It needs to receive/transmit data all the time and execute complex operations such as TDMA schedule generation and data aggregation. Extra works make it more easily energy exhausted. Long time wakeup also makes it possible to have hardware or communication malfunction. Once cluster head fails, the cluster working has to halt and whole cluster's data would be lost. So the reliability of cluster head is very important. In this paper, we proposed an approach in which a substitute of cluster head (SCH) is selected and works in the sudden failure of cluster head. For substitute of cluster head, it needs more energy consumption than other noncluster head nodes. But by sacrificing consumption in one node, the reliability of whole cluster could be increased. Moreover, by selecting SCH in rotation, the average energy consumption is insignificant and does not affect the whole system.

Finding multiple node-disjoint paths also provides fault tolerance in routing. The system can switch from an unavailable path with broken links to alternative candidate paths. Major solutions [6, 7] for fault-tolerant and reliable data propagation of WSNs are based on the multipath routing paradigm, which provides each sensor node with alternative paths to destination. It solved the problem that wireless communication is not stable and provided fault tolerance for data transmission. In the phase of aggregated data transmission from cluster heads to the base station, we perform multipath discovery algorithm for increasing reliability.

Based on the above strategies, we classify sensor nodes into five categories according to their different roles in WSNs. They are branch node (BN), cluster head (CH), substitute of cluster head (SCH), inter-medium node (IN), and normal node. For each category, it executes alternative network and MAC protocol and operates specific functions, such as schedule generation, data aggregation, energy alarm, radio control, route discovery, and data transfer. All the design standards for each category are based on considerations of energy efficiency and reliability.

There are some original and progressive contributions in this paper: firstly, we proposed an energy-efficient and reliable WSN mechanism, which adopted multiple strategies including the hierarchical clustering structure, multipath routing discovery protocol, energy-efficient and flexible MAC protocol, and substitute of cluster head. Secondly, we classify sensor nodes into five categories according to their functionalities in WSNs. For each category, we considered extra works that are responsible for and designed flexible

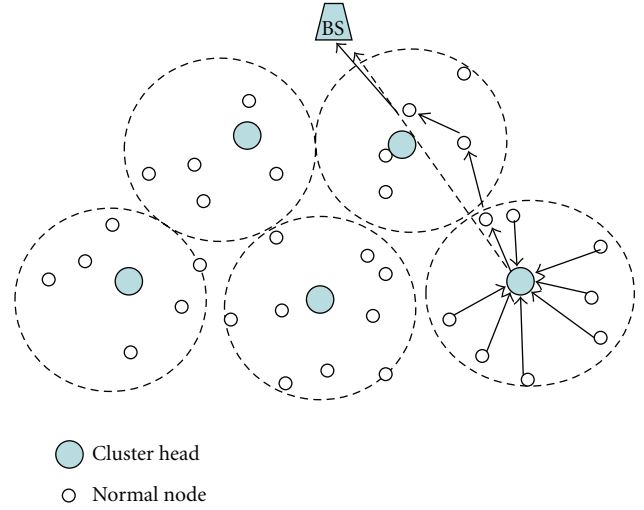


FIGURE 1: Clustering structure in WSNs.

MAC protocol. This strategy guarantees more fairness in node energy consumption and prolongs the network lifetime as far as possible. Thirdly, the mechanism of substitute of cluster head (SCH) accounts for the situation of CH's sudden failure. Moreover, multipath discovery scheme can solve the problem when a path fails. These guarantee the reliability and robustness of the system. Fourthly, all the phases including cluster formation, data transmission within cluster, and multiple path transmission are in distributed manner; it guarantees the scalability of the network. Finally, we use two global parameters of total energy and node sum to improve the performances of both cluster head/SCH election and routing discovery. In our scheme, these two parameters are updated after every round and calculated by the base station. It is an efficient way to control the network quality.

The remainder of this paper is organized as follows. Section 2 presents the related works. Section 3 describes the classification of sensor nodes and explains the detailed algorithms of each category. Section 4 analyzes the simulation results and shows why this mechanism outperforms others. Finally, Section 5 draws the conclusion.

2. Related Works

2.1. Energy Saving Mechanisms

2.1.1. Hierarchical Structure. The aim of the hierarchical structure is to reduce the energy consumption by reducing the amount of direct communications. Clustering structure is the main form of hierarchical structures in WSNs.

The literature [2, 3] stated a survey of energy-efficient hierarchical cluster-based protocols in WSNs. As illustrated in Figure 1, the network is divided into a few clusters. Nodes are grouped into clusters with a cluster head responsible of routing from the cluster to the base station. In this model, data is firstly collected periodically by normal sensor nodes. Instead of directly transmitting data to the base station, cluster heads collect the data from the sensor nodes within their clusters and send an aggregated data to the base

station directly or through a multihop path (based on the communication radio range of CHs and scale of network).

Cluster heads consume relatively more energy than non-cluster head nodes because cluster heads have more loads to handle. Hence, cluster heads may run out of their energy faster than other sensor nodes.

LEACH [4] is the classical clustering hierarchical protocol for WSNs, which incorporates randomized rotation of the high-energy cluster head position among the sensors to avoid draining the energy of any one sensor in the network. In this way, the energy load of being a cluster head is evenly distributed among the nodes. The probability of cluster head election depends on the number of clusters k and the sum of the sensor nodes N :

$$P_i(t) = \begin{cases} \frac{k}{N - k * (r \bmod (N/K))} & : C_i(t) = 1. \\ 0. & \end{cases} \quad (1)$$

In (1), r is the round number and $C_i(t)$ depicts whether node i has been the cluster head in recent rounds.

LEACH enhances network lifetime and energy consumption compared with the flat algorithms. LEACH approach is based on the assumption that all nodes start with an equal amount of energy and cluster heads have uniform energy consumption. Obviously the latter is not realistic. One shortcoming of LEACH is that it does not consider the energy factor when electing cluster headers, whereas we designed a more efficient cluster head election algorithm. Secondly, LEACH does not work well in the dynamic network. It is because parameters (number of clusters, number of nodes) are programmed into the nodes a priori, but in reality these parameters continuously change with the depletion of nodes. Furthermore, LEACH is not scalable for large size of WSN. It is assumed that every node has enough power to communicate with others and BS directly, but it is not realistic for a large scale of WSN.

The subsequent LEACH-optimized protocols in literatures [5, 8, 9] mostly improve the capability of cluster head distribution, so that the energy consumption of the whole network is reduced and the system lifetime is prolonged. However, most of these hierarchical protocols only improved the algorithm of cluster formation and gave little consideration on aggregated data transmission. Since the aggregated data is important, the reliability of transmission should be guaranteed.

2.1.2. Energy-Efficient MAC Protocols. In wireless communications, energy waste mainly comes from the following sources: idle listening, data collision, overhear and control packet overhead, and so forth. Among them, idle listening causes nearly half of the energy waste in WSNs [1], so turning off the radio in unnecessary time is almost the most efficient way to save energy. Energy-efficient MAC protocols achieve energy savings by controlling the radio wake/sleep and arranging reasonable access schedule to reduce energy waste caused by these sources.

MAC-layer protocols can be categorized into two types: schedule-based and random access protocols. The former

is the protocol in which access to the channel is based on a schedule. Channel access is usually limited to one node at a time. Random access protocols avoid preallocation of resources to nodes, so they are more flexible than schedule-based protocols. However, this results in collision. The main objective of random access protocols is to minimize the occurrence of collisions.

TDMA (time-division multiple access) is a typical schedule-based MAC solution. It is achieved by dividing the radio frequency into time slots and then allocating unique time slots to each communicating node. Nodes take turns wake up and transmit/receive data. In LEACH protocol, it uses TDMA in data transmission within cluster. The time line of LEACH protocol is as shown in Figure 5(a). Most clustering WSN protocols use TDMA as the MAC protocol to manage the access of nodes within a cluster.

Although TDMA protocol can reduce energy consumption greatly, it limits the number of joint nodes and the constraints of time allocation so that it is short of scalability and flexibility. Contention-based protocols require no coordination among the nodes accessing the channel and provide more scalability and flexibility. CSMA series are the popular contention-based protocols. For WSNs, S-MAC [10] is proposed to reduce the energy efficiency by using a low duty cycle operation. It also provided schemes for collision avoidance, overhearing avoidance, and synchronization. But the energy saving may result in high latency and low throughput. In the same time, collisions cannot be avoided in content-based MAC protocols. In our proposed mechanism, we use both TDMA and content-based MAC protocols in different phases.

2.2. Reliability Mechanisms

2.2.1. Multipath Routing Algorithm. As described in [6], there has been a lot of research works in multipath routing for WSNs in recent years. Using multipath routing provides fault tolerance of node failure along single path and increases the network reliability. Constructing node disjoint paths has been considered the most reliable method in WSNs.

H-SPREAD [7] uses the branch-aware flooding method to find a set of node-disjoint paths. As the extension of the flooding-based beaconing protocol [11], H-SPREAD proposed an N -to-1 multipath discovery protocol. The multipath discovery procedure is presented in two phases, with each phase implementing one of the mechanisms. The mechanism used in the phase one, which is termed as a branch-aware flooding, takes advantage of the simple flooding technique. Without introducing additional routing messages, the mechanism is able to find a certain number of node-disjoint paths, depending on the density of the network topology. The mechanism used in the phase two, which is termed as a multipath extension of flooding, helps to exchange the node-disjoint paths found in the phase one among the nodes on different branches. At the cost of some more message exchanges, it is able to increase the number of paths found at each sensor node.

Figure 2 shows the algorithm of branch-aware flooding. We name the one-hop neighbors of the BS as branch

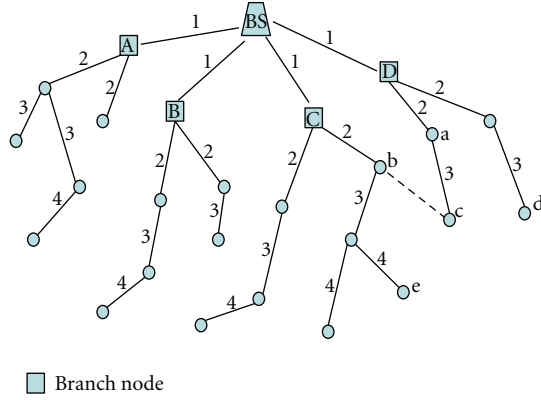


FIGURE 2: Branch-aware flooding algorithm.

node (BN). If a node receives a request message from the neighbor with different BN value, it can be regarded as the discovery of an alternative path. For example the square nodes (A, B, C, D) represent BNs. Node c firstly receives message from node a, then it constructs a primary path (A-a-c). And then c continuously receives messages from nodes b, d, e. When judging an alternative disjoint path, c will compare the BN of each sender. If BN of sender is the same with its, it means the new path is joint; otherwise, the new path is a disjoint one. For the paths with the same BN, c will compare the hop number. The less hop number path is selected. By this way, an alternative disjoint path (C-b-c) is constructed by c.

In our proposed path discovery algorithm, similar routing protocol is used. Instead of store alternative paths before data propagation, we utilized a more flexible path discovery algorithm.

Branch node and nodes nearer, the base station relay more packets and more actively participate in communication. As a result, these nodes expand more energy and are more failure prone due to energy depletion. How to conserve energy in these nodes is an important issue, for if these nodes are exhausted, all paths to the base station are broken. In previous multipath routing protocols for WSNs [6, 7, 12, 13], the problem of how to avoid extra energy consumption in neighbors of base station has not been discussed, whereas in our proposed mechanism this problem was considered.

2.2.2. Clustering Fault Tolerant Mechanism. In [14, 15], fault tolerance of cluster-based WSNs is discussed.

CPEQ in [14] is a cluster-based routing protocol. It takes account of the reliability of data transmission within clusters. It assumes that data propagation within clusters is multihop. It implements a path repair mechanism within clusters. It did not account for the situation that cluster head failed. The literature [15] uses the CHs coordination mechanism for fault tolerance. CHs exchange status message after every cycle to detect the failure of CH. If failure is detected, the nodes with the fault CH are assigned to near cluster in next round automatically. This mechanism needs CHs exchange a lot of messages resulting in much overhead. Moreover, it causes node isolation problem after reassignment.

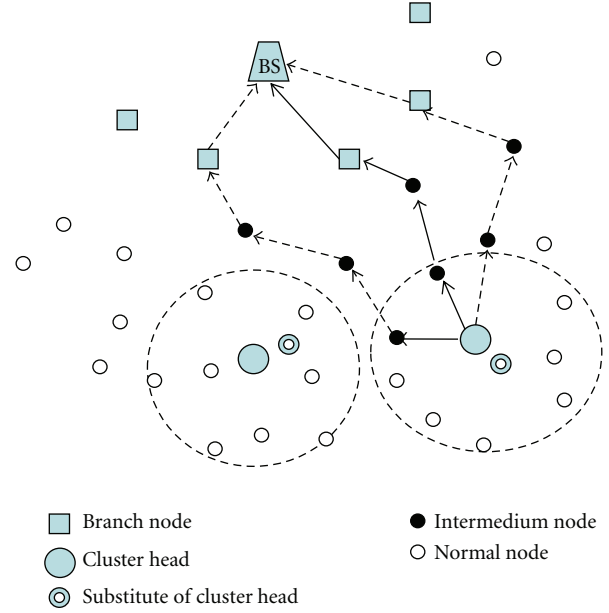


FIGURE 3: Node classification mechanism.

3. Node Classification Mechanism

3.1. Network Model. We assume this proposed mechanism is for a large scale and dense wireless sensor network. One or more base stations exist as the gateway between WSN and the outside information system. Except base station, all sensor nodes have similar initial energy and capabilities (sensing/processing/communication). Nodes need not change their radio range in our system. Communications between nodes are bidirectional and symmetric. All nodes are able to be aware of their residual energy. Each node has a unique identifier to be marked in the network. Within a cluster the communication from normal nodes to CH is one hop, whereas the communication from cluster heads to base station is multihop. This approach is more suitable for the applications in which sensed data is collected periodically. Finally, synchronization is necessary among sensor nodes due to the demand of MAC protocol.

3.2. Energy-Efficient and Reliable Mechanism. As in [13], this protocol mainly consists of three phase. The first phase is the initialization phase, in which each node retrieves the information about the neighbors and the network. The second phase is the cluster formation phase, in which clusters are formed in a distributed manner. The final phase is the data propagation phase, which includes two steps: sensed data is transmitted within clusters, and aggregated data is transmitted from cluster heads to base station. Our scheme utilized hybrid MAC protocols for different phases. The entire WSN is displayed in Figure 3, where we can see different categories of sensor nodes.

3.2.1. Phase One: Initial Broadcasting. This phase starts from the BS broadcasting a request message. The format of this message is $\{INI, RID, BID, SID, E_{re}, H_{count}, E_{to}, N, T_1\}$, where

TABLE 1: Structure of neighbor information table.

Neighbor ID	IsParent	BID	Residual energy	Hop count	RSSI
-------------	----------	-----	-----------------	-----------	------

INI indicates that the type of message is initialization; RID is the round identifier, which is generated by the base station; BID is the identifier of the branch, that is, identifier of the branch node; SID is the identifier of the sender node; E_{re} is the residual energy of the sender node; H_{count} is the hop count from the sender to the base station; E_{to} depicts the total energy of all nodes, and it is calculated by the base station. Finally N is the sum of nodes after last round. Here E_{to} and N are prepared for the selection of cluster heads. T_1 depicts the timestamp for the ending of phase 1 and the beginning of the next phase.

Hop by hop, node receives the broadcast message, updates information of message, and rebroadcasts it. Upon receiving the first message, node marks the sender node as its parent. After this phase, every node decides whether it is a Branch node, stores the parameters of total energy E_{to} and total number of nodes N for the next phase, and constructs the table of neighborhood information, which will be used to discover multiple paths in the data propagation phase. This mechanism takes advantage of the broadcast nature of the wireless communication. Each node rebroadcasts once and only once. The table of neighbor information is as shown in Table 1.

In Table 1, node stores all the received neighbors' information, including neighbor node ID, whether it is the parent node, branch ID of this neighbor, residual energy, hop count to BS, and the received signal strength indicator (RSSI) which is proportional to the distance between them. In this phase, branch aware flooding algorithm is used. In this process even some messages were lost due to some node failures or link conflicts, and nodes could obtain the network parameters from other neighbors and construct the approximate impress about the neighbors. All sensor nodes keep awake until this phase ends, that is, time T_1 .

3.2.2. Phase Two: Cluster Formation. After previous phase, every node has got the information about the total energy E_{to} and sum of nodes N . After time T_1 , phase two begins and each node begins to decide whether to be a cluster head in a distributed manner.

Once the node has selected itself to be cluster head, it broadcasts an advertisement message (ADV) using CSMA MAC protocol. This message is a small message containing the node's ID, cluster's ID, and an ADV header. Nodes decide which cluster they should belong to by choosing the cluster head that requires the minimum communication energy, based on the received signal strength of the advertisement messages from cluster heads. After each node has decided to which cluster it belongs, it transmits a join-request message (JOIN-REQ) back to the chosen cluster head also using CSMA MAC protocol. This message consists of the node's ID, the joining cluster head's ID, the residual energy, and the hop count to BS of the node.

The cluster head node sets up a TDMA schedule and broadcasts this schedule to the nodes in the cluster. In the same time, according to the parameters of residual energy and hop count and the RSSI, CH selects a node as a SCH. SCH declaration is broadcasted in the same message with the TDMA schedule. Node matching the SCH_ID will mark itself as a SCH.

In this phase, CSMA MAC protocol is utilized to avoid communication collisions. To conserve energy, branch nodes do not join any cluster, so they turn into asleep in the whole phase. Other sensor nodes keep awake during cluster formation.

3.2.3. Phase Three: Data Transmission. This phase mainly consists of two steps: at first the data propagation within a cluster, then the data propagation from the CHs to the BS, which is multihop transmission along multipaths.

In the data transmission within clusters, nodes comply with the TDMA MAC protocol. Nodes send their data to the cluster head at most once per frame during their allocated transmission slot. Once the cluster head receives all the data, it performs data aggregation to enhance the common signal and reduce the uncorrelated noise among the signals. In our proposed mechanism, cluster heads execute a spatial correlation on collected data. In our approach, after every round, BS needs to know the whole residual energy of all nodes and the sum of nodes alive. Therefore, in this step, besides the data aggregation, the cluster head also does the following computations:

$$N_{cluster} = \sum_{i=1}^n i, \quad (2)$$

$$E_{cluster} = \sum_{i=1}^n E_i - N_{cluster} * (E_{TX} + E_{RV}) - E_{fusion}.$$

$N_{cluster}$ is the sum of nodes in this cluster, while $E_{cluster}$ is the whole residual energy value of all nodes in this cluster, which accounts for the energy consumption during the data propagation. Finally, the cluster head sends the data message, which consists of the message type (DATA), aggregated data, $N_{cluster}$, and $E_{cluster}$.

During the above process, if the cluster head fails, for example, running out of energy, all the sensed data within the cluster will be lost. It is why we select a substitute of cluster head. Even if CH cannot work normally, SCH will take over the works of CH to guarantee the reliability of data propagation.

In this step, CHs and SCHs keep awake during the whole process, whereas other nodes only wake up in its allocated time slot.

The second step is the aggregated data transmission from the cluster heads to the BS. Since the BS may be far away, this is a multihop and high-energy transmission. Before data transmission there is a process that we call "INs Broadcasting" as shown in Figure 4. In this process, CHs (or SCHs) firstly select appropriate next hop nodes based on the information of neighbor table and broadcast a NEXT_IDs message. This message not only includes the

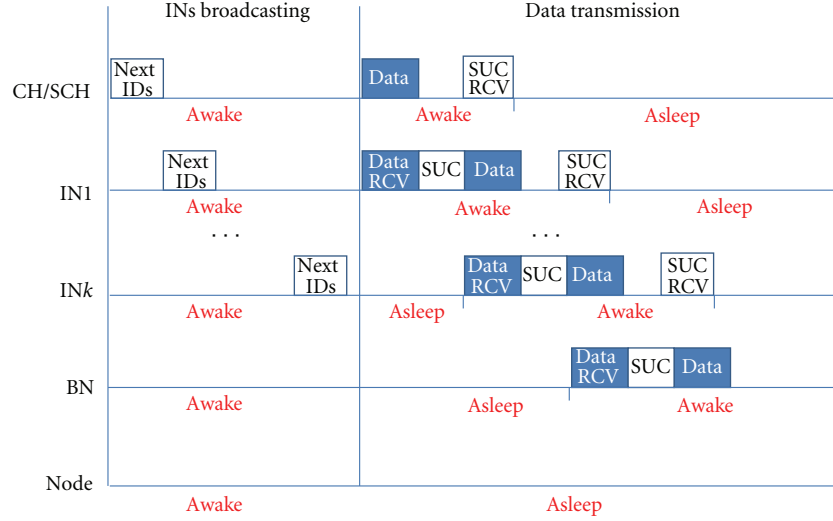


FIGURE 4: Timeline for each category during data propagation from CHs to BS.

next hop nodes' ID, but also includes hop count from CH to BS along each next hop node. The nodes which receive NEXT_IDS message from CHs/SCHs will mark themselves as IN1 select their next hop nodes, and rebroadcast NEXT_IDS messages. NEXT_IDS messages recurrently broadcast until they are received by branch nodes. By this way, sensor nodes in WSNs can be distinguished between intermedium nodes and none-intermedium nodes.

In the data transmission from CHs to BS, only CHs/SCHs, INs, and BNs keep awake and take part in the data transmission, whereas other nodes turn into sleep until the process ends. As shown in Figure 4, after every data transmission, the sender node will wait for the SUC message from the recipient. If the SUC message is received, the sender will turn into sleep until the end of data transmission; otherwise, if the sender cannot receive the SUC message in a limited period of time, the sender will resend the data to another next hop node and the data transmission process will repeat. Furthermore, in last INs broadcasting process, each IN knows the hop count of the path it resides, so it can estimate the waiting time T_{INi} according to (3):

$$T_{INi} = k \left[(H_{\text{path}} - H_{INi}) \times \Delta T \right]. \quad (3)$$

In (3), H_{path} is the whole hop count of the path, whereas H_{INi} is the hop count from INi to BS. ΔT is the approximate time for data transmission between two one-hop nodes, which includes the time for data transmission and receiving, plus the time for SUC message transmission and receiving. Coefficient k is larger than 1, which accounts for delay of data transmission. During period of T_{INi} , INs keep asleep and do not wake up until the end of T_{INi} . Using this flexible MAC protocol, all nodes turn into sleep in their unnecessary time, so energy consumption in this phase could be decreased greatly.

There are several disjoint paths for data delivery from CHs to BS. As illustrated in Figure 3, the solid arrow path is

the primary path from CH to BS, whereas the dashed arrow paths are alternative paths as the backup of primary path.

Ultimately, the sensed data and the parameters of the network (sum of nodes N , whole residual energy E_{to}) should be retrieved by the BS. Then the BS initiates another round.

3.3. Node Classification. Based on the functionality in WSN, we defined five classifications for sensor nodes. It is remarkable that some category may exist only in a specific phase not the entire process.

3.3.1. Branch Node (BN). The definition of the BN is the one-hop neighbors of the base station. According to the branch aware flooding algorithm [7], each BN represents one branch. As shown in Figure 2, nodes A, B, C, and D are BNs. In phase 1, upon receiving the first message, the node marks the sender node as its parent and the path in which parent node resides is regarded the primary path. If a node receives an initial broadcasting message with different BN value from the value of its parent node, it can be regarded as the discovery of an alternative path.

According to the simulation result in [7], the nearer to the BS, the node has more burdens on data transmission. The BN acts a critical role in the network, because once it is exhausted the whole branch is separated and the downstream paths are correspondingly failed. In order to conserve energy, in our proposed mechanism the BN does not join the cluster formation and data sense. It only acts as a router in the network. Furthermore, BNs keep asleep in phase 2, by this way BNs conserve much energy. If its energy is below a limitation value, it should announce that it abandons the role of branch node and transforms to a normal node.

In phase 1, if node i received initial broadcasting message directly from BS, it will decide whether it should be a branch node according to its residual energy:

$$E_{re} \geq E_B = k(M * (E_{TX} + E_{RC})). \quad (4)$$

E_{TX} is the energy consumption in data transmission, whereas E_{RC} is in data reception. M is proportional to the number of messages. k is the coefficient. If the condition is satisfied, node i marks itself as the BN. By this strategy, we limited branch nodes to the nodes with more residual energy avoiding early energy exhaustion of some nodes.

3.3.2. Cluster Head (CH). In our approach, cluster heads are selected in distribution based on the residual energy, the density, and the distance to base station. CH is in charge of data receive, TDMA schedule generation, data process, data aggregation, and data transmission. The energy consumption of CH is much quicker than normal nodes. So an evenly CH selection algorithm is necessary.

In our paper, CH is selected in a distributed manner. Node i generates a random number P between (1, 0). Then node i calculates the probability $P_i(t)$ as (5). If $P < P_i(t)$, node i will declare itself a CH:

$$P_i(t) = \min \left\{ k * \left[\alpha \frac{E_i}{E_{to}} + \beta \frac{N_{neig}}{N} + \gamma \frac{1}{H_{toBS}^2} \right], 1 \right\}. \quad (5)$$

E_i is the residual energy of node i . N_{neig} is the sum of neighbors of node i , which can be retrieved from the neighborhood table. N is the total number of all sensor nodes, whereas E_{to} is the total energy of all nodes. H_{toBS} is the hop number to BS. k is the optimal number of clusters, which can be calculated as about 5% of N . α , β , and γ are the coefficients, whose values are in the range (0, 1).

Compared with LEACH, the CH selection algorithm considers the residual energy, the density (number of neighbors), and the distance (hop count) to BS of each node, instead of the simple randomized rotation. So the cluster election algorithm has more efficiency. In here two global parameters are used to control the quality of clustering, which are the total number of existing nodes N and the total amount of energy of all nodes E_{to} . These two parameters are updated after one round and reflect the status of network.

3.3.3. Substitute of Cluster Head (SCH). Substitute of cluster head (SCH) is defined as the substitute when CH suddenly fails. Because of the important role in WSNs, the reliability of cluster head is very important. In this paper, we proposed a method where a substitute of cluster head is selected and works when cluster head has failure. For this substitute of cluster head (SCH), it needs more energy consumption than other noncluster head nodes. But by sacrificing energy consumption in one node, the reliability of whole cluster can be increased. Moreover, by selecting SCH in rotation, the average energy consumption is insignificant and does not affect the whole system. This strategy guarantees that the data could be transmitted correctly even if the cluster head fails. This could improve the reliability and fault tolerance of the system.

The strategy of SCH includes two critical parts: the first is SCH selection algorithm, and the second is the MAC protocol designed for SCH.

SCH selection work is done by CH. After CH is selected, CH broadcasts ADV message, then nodes send their

information including their ID, residual energy, and hop count to BS in JOIN-REQ messages to the cluster head. Based on these parameters and received signal strengthen indicator (RSSI), CH finally appoints a node as a substitute who has more residual energy, less hop count, and nearer to CH. SCH appointment is broadcast with the TDMA schedule message:

$$ID(SCH) = ID \left(\max \left\{ k_1 E_i + k_2 RSSI_i + \frac{k_3}{H_i} \right\} \right) \quad (1 \leq i \leq n). \quad (6)$$

Equation (6) shows the principle to select the SCH. A CH receives JOIN-REQ messages, which include the residual energy and hop count to BS of all cluster member nodes. Through calculation as in (6), the node which has the maximal composite value of residual energy, hop count, and RSSI is selected as the SCH.

During the process of data transmission within cluster, if the cluster head is in any urgent situation, for example, physical fault or communication fault, at this time the data transmission or data aggregation process is carrying on, then all the sensed data within the cluster will be lost. It is why we select a substitute node. In this process, all nodes within the cluster comply with the TDMA schedule except SCH. SCH keeps awake during all the process. It receives data as the CH does. After the TDMA schedule expires, different with LEACH, there is a specific time slot for CH to transmit an RCV_SUC to SCH. In this timeout, if SCH receives RCV_SUC, it empties the memory and waits for the next phase; otherwise it means that there are some problems with CH, so SCH takes over the works of data aggregation and aggregated data transmission. By simply inserting a timeslot, the TDMA schedule need not be altered at all. It is a very reliable and flexible fault tolerant scheme. The difference of our MAC protocol with LEACH is illustrated in Figure 5.

3.3.4. Intermedium Node (IN). The data transmission between CHs and BS is multihop communication. In this process, we assume that only nodes in the delivery paths need to be awake. We name these nodes intermedium nodes (INs). IN is specially defined for the process of data transmission from CHs/SCH to BS. By this way, the idle listening and overhearing of nodes not in the routes from CHs/SCHs to BS are avoided, achieving significant energy saving.

Before transmitting data, CH/SCH performs a route discovery phase beforehand. After this phase, CH/SCH determines its next hop nodes and broadcasts a message. The next hop nodes receive the message, mark themselves as Ins, and keep awake in the following data transmission process, whereas other nodes go to sleep until the data transmission ends. The MAC protocol for this process has been discussed in phase 3.

The detailed next hop node discovery algorithm is described in Algorithm 1. At first the CH (or SCH) checks its neighborhood table as shown in Table 1. The node marked with "Parent" is the next hop along the primary path from the cluster head. Then CH continuously looks for the neighbors with different BID value from its parent. After comparing the energy and hop count, the CH chooses

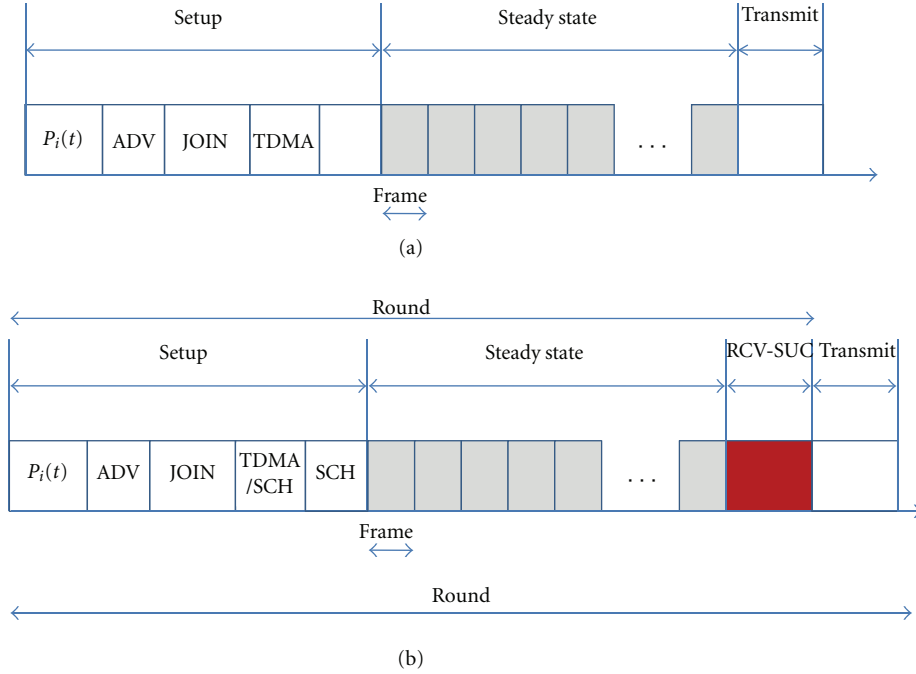


FIGURE 5: (a) Timeline of LEACH. (b) Timeline of the proposed protocol.

the next hop nodes. Other INs execute the same algorithm recursively to discover the next hop nodes. This route discovery algorithm generates multiple disjoint paths to improve the reliability of data delivery.

3.3.5. Other Nodes. The other nodes are in charge of basic works such as initialization, data sensing, cluster formation, and data transmission (within cluster). Normal nodes have minimal energy consumption than other categories.

4. Performance Analysis

In this section, we evaluated the performance of our proposed mechanism based on two factors: energy consumption and reliability. At first, we observed the energy dissipation amount of each category in one round and analyzed the cause in details. Then we analyzed the energy savings by using strategies of branch node energy conservation and flexible INs MAC protocol. We also compared the proposed mechanism with previous clustering based works [4, 5] on the performance of network lifetime, which is a synthesis factor to evaluate the superiority of wireless sensor networks. In the next we set different node failure rates and observed the average packet delivery ratios, which reflect the success rate of packet transmissions to the BS. We compared the situations of with/without SCHs and multipath routing algorithm, respectively. WSN reliability can be measured through this result.

4.1. Experiment Setup. We have carried out a set of experiments using NS-2 simulator. In the simulation, we used the same radio energy dissipation model as in [4]. The values of parameters used for simulation are as shown in Table 2.

TABLE 2: Simulation parameters.

Parameters	Value
Topology size	1000 m × 1000 m
Number of nodes	1000
Transmission range	50 m
Sensing range	40 m
Initial energy	5 J
Packet size	4000 bits
Location of BS	(600, 160)
Bandwidth	1 Mb/s
Electronics energy	50 nJ/bit
Free space ϵ_{fs}	10 pJ/bit/m ²
Multipath fading ϵ_{mp}	0.0013 pJ/bit/m ⁴
Average data aggregation energy	5 nJ/bit/signal
Percentage of CHs (%)	3, 4, 5

In the following experiments, we set that the percentage of cluster head number 3%, 4%, and 5%, respectively. We evaluated the energy dissipation of different categories of nodes in one round and the reliability with two kinds of strategies. Every result shown is the average of 100~200 experiments, and the confidence interval used in the simulation is 95%.

4.2. Energy Dissipation. The first experiment we do is to obtain the average energy dissipation of different category of nodes in one round. Here we divided a round into four phases as illustrated in Figure 6, which is little different from the phases we divided in Section 3.2. We can observe that in each phase the energy dissipation of each kind of node


```

(1) Define:
(2)  $n_i$ : sensor node  $i$ 
(3)  $NB_i$ : the set of neighbors of  $n_i$ 
(4)  $NEXT_i$ : the set of next hop nodes of  $n_i$ 
(5) IN: Inter-medium Node
(6) Begin
(7) For each  $j \in NB_i$  do
(8)   If  $n_j \cdot IsParent == True$  then
(9)      $n_j \rightarrow NEXT_i$  //Primary path is found
(10)   Else
(11)      $IsAlterPath = True$ 
(12)     For each  $k \in NEXT_i$ 
(13)       If  $n_j \cdot BID == n_k \cdot BID$  then
(14)          $IsAlterPath = False$ 
(15)         If  $n_j \cdot H_{count} < n_k \cdot H_{count} \ \&\& \ n_j \cdot E_{re} > E_0$  then
(16)            $n_j \rightarrow NEXT_i$ 
(17)            $n_k$  is removed from  $NEXT_i$ 
(18)         End If
(19)       End If
(20)     End for
(21)     If  $IsAlterPath = True$  then
(22)        $n_j \rightarrow NEXT_i$  // Alternative path is found
(23)     End If
(24)   End If
(25) End For
(26)  $n_j$  broadcast  $NEXT_i$ 
(27) For each  $k \in NEXT_i$  do
(28)   Receive  $NEXT_i$ 
(29)    $IsIN = True$ 
(30)   Execute 7
(31) End For
(32) End

```

ALGORITHM 1

is different. It is due to adopting specific network and MAC protocol for different category in each phase. Now we will analyze the reasons to cause difference of energy dissipation for different categories.

The first phase in a round is the initial broadcasting. In this phase, all sensor nodes keep awake and broadcast/receive messages to collect information of whole network and the neighbors. Energy dissipation in here is the same for each category of nodes. As shown in Figure 6, lines representing varied category of nodes coincide in this phase.

The second phase in a round is the cluster formation phase. In this process, BNs go to sleep and do not join any cluster formation activities in order to conserve energy. CHs consume more energy than other nodes due to their high workload.

Third phase is the data transmission within clusters. BNs continue sleeping, so their energy consumption is 0. Because in this process TDMA MAC protocol is used, normal nodes only wake up in their allocated time slot to transmit sensed data. So the average energy dissipation of normal nodes is very low. Whereas CHs and SCHs keep awake during the whole phase and continuously receive data, this makes their energy consumption increased rapidly.

The last phase is the data transmission from CHs/SCHs to BS. In this phase we utilized a flexible MAC protocol as we

have stated in Section 3.2.3. Nodes are divided into INs and normal nodes. Only nodes participating in data transmission keep awake, while other nodes go to sleep after declaring they are not INs. In this phase CHs (or SCHs), INs, and BNs all wake up during their work time and turn to sleep after work and before the packets arrive. It is shown that BNs consumed maximal energy because nodes nearer to BS have high workload to transfer data.

Overall, CHs greatly have more energy consumption than other nodes in a round. This problem could be solved by electing CHs in rotation. It is suitable to the SCHs too.

It needs to mention that the timeslot for each phase marked in horizontal axis is not the exact one in a round. In reality, the period of data transmission within cluster occupied the most time in a round. In the experiment, we set the percentage of cluster heads is 4%. Because the trend is same for other percentages, it is not necessary to do the alternative experiments.

In this experiment, we can find that BNs greatly conserved energy by limiting their participations only in some specific phases (other time they should be in sleep mode). The amount of energy dissipation of BNs was reduced as little as the normal nodes. We compared mechanisms with BN sleep strategy and without BN strategy in different percentage of CHs. In Figure 7(a), the result shows that the

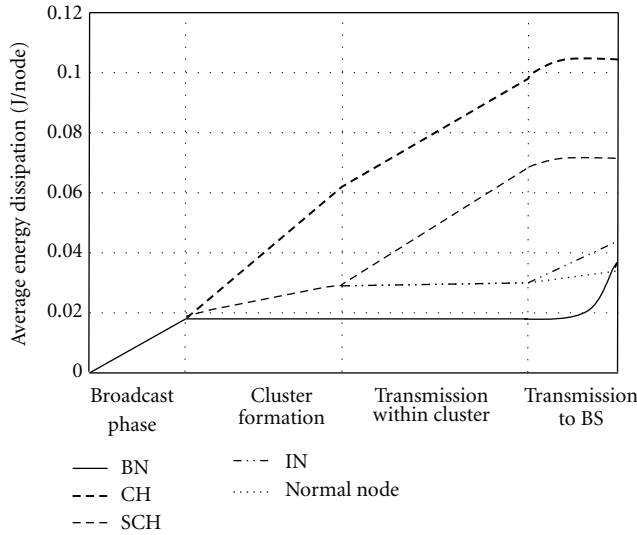


FIGURE 6: Average dissipated energy of each category in a round time.

average energy dissipation of BNs is decreased about 30%–50% through utilizing BN sleep strategy.

In Figure 7(b) we compared the average energy consumption of all sensor nodes in two mechanisms, respectively, with flexible MAC protocol and without. Flexible MAC protocol means the power control strategy we designed for respective phase. In phases of cluster formation and data transmission within clusters, BNs turn off radio and keep in sleep mode; in phase of data transmission from CHs to BS, only INs/BNs keep awake in their active time. The energy conservation amount by using this strategy reaches 25%–45%. It is noteworthy that, with the increase of percentage of CHs, the amount of energy saving declines. It is due to the number of paths of multihop data transmission from CHs to BS increases, correspondingly the number of joint INs increases.

In order to show the superiority of the performance on energy saving, we compared our proposed protocol with the clustering based protocols, LEACH and HEED. It should be noticed that the routing protocols of data forwarding from CHs to BSs are different in each work. In LEACH, the authors assume the CHs can communicate with the BS directly after data collection and aggregation, while in HEED, the authors utilized the TinyOS beaconing approach, which constructs a breadth-first spanning tree rooted at the BS. It is obvious that the former does not fit for the large-scale WSNs for sensor nodes cannot communicate with the BS directly when the distance increases, whereas the latter does not consider for the reliability and fault tolerance of data transmission. The strategies for improving reliability of data transmission also cause extra energy consumption because they require extensive message exchange. But by adopting more optimal CH selection algorithm and flexible MAC protocol, we also effectively decrease the energy consumption. We compared the network lifetime for three protocols when the number of nodes is 500 and 1000, respectively. The results are shown in Figure 8. The round number for the first and last node dies

in each protocol are illustrated, which indicates the network lifetime. From Figure 8 we can observe that our proposed protocol extends lifetime above 140% than LEACH and 50% than HEED. It is because LEACH randomly selects cluster heads and HEED selects cluster heads based on the residual energy of nodes, whereas our proposed algorithm considers the residual energy, number of neighbors (density), and the depth to BS (distance to BS), so that the final cluster heads are well distributed across the network. The flexible MAC protocol also improves the capability of energy saving by turning off node radio when they are idle in the network, since it is well known that idle listening almost causes half of power wastes in wireless sensor networks.

4.3. Data Delivery Reliability. In this section we analyzed the reliability of data transmission when using different reliable strategies. As illustrated in Figure 9, with the increasing of node failure rate in WSNs, the average packet delivery ratio decreases correspondingly. Here, the reliability is represented by the probability that a message is successfully delivered, which is calculated as the total number of effective information received at the BS over the total number of effective information initiated from all the sensing nodes. Since we used data aggregation approach, the number of delivered messages cannot be used as the measure parameter. We prefer to use the number of effective information as the measure parameter of data reliability. The data delivery consists of two steps: delivery within cluster and delivery to the BS. The SCH strategy is designed responsible for the failure of data delivery within cluster while multi path routing strategy is designed for the fault tolerance during data delivery from CH to BS.

The faults in WSNs include two types of impact: the node failure and the link failure. In our simulation, we only considered the situation of node failure. It is because that the SCH mechanism is our original and creative strategy, which should be evaluated with high priority. The node failure is random, which includes the situation of either the normal nodes or the CHs (SCHs) fail. The situation of both CH and SCH in the same cluster that have failure simultaneously is also possible in this case. So this fault model is suitable for our research. In previous works [6, 7, 12, 14], the impact of multipath routing protocols has been fully discussed.

As shown in Figure 9, in the mechanism using both SCH and multipath routing mechanisms, the decrease of data delivery ratio is not significant; whereas the mechanism without SCH or multipath routing algorithm decreases very severely. The mechanism only with multipath discovery algorithm performs between the formers. It proved that both SCH approach and multipath routing algorithm are effective in improving the reliability of data delivery.

5. Conclusion

In the paper, we proposed a cluster-based energy-efficient and reliable mechanism for WSNs, which integrates the advantages of clustering hierarchical structure and fault tolerant systems. Based on the functionalities in WSNs, sensor nodes are classified into five categories. Each category

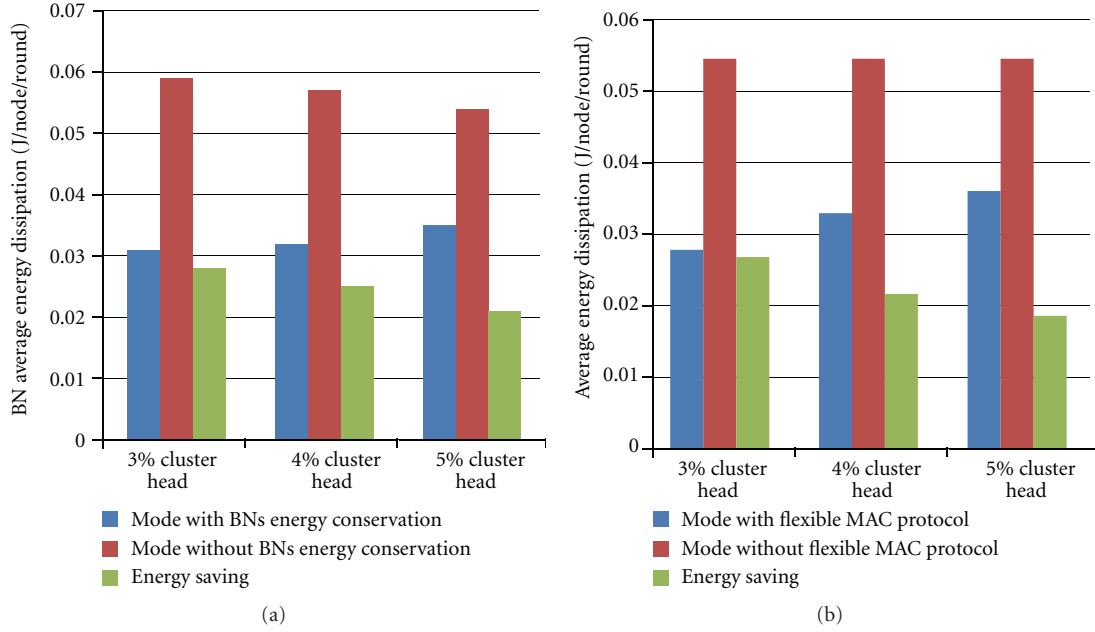


FIGURE 7: (a) Comparison of average energy dissipation of BNs with/without BN energy conservation strategy. (b) Comparison of node average energy dissipation with/without flexible MAC protocol.

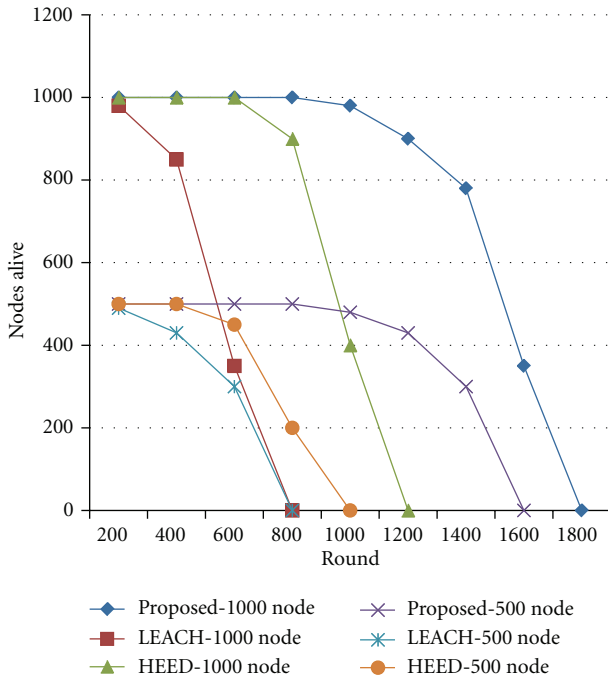


FIGURE 8: Comparison of network lifetime.

executes specific network and MAC protocol and processing operations. Using this flexible strategy, our proposed mechanism provides maximal reliability, energy efficiency, and scalability.

It is worth mentioning that some strategies for reliability and fault tolerance cause extra energy consumption whereas strategies for energy efficiency also impact the performance of reliability. For example, the flexible MAC

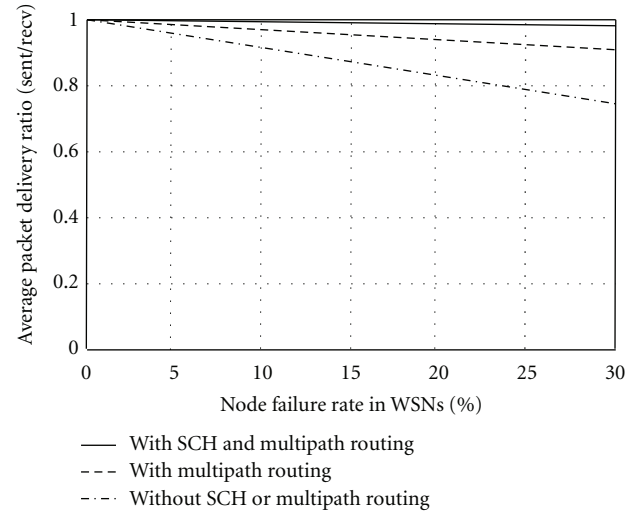


FIGURE 9: Comparison of average packet delivery ratio.

protocol decreases energy consumption greatly, meanwhile it is possible that nodes cannot wake up timely to participate network due to the power control algorithm. The tradeoff between the energy efficiency and reliability is a critical issue in the future. As future works the proposed mechanism will be evaluated in a real environment.

References

- [1] M. Al Ameen, S. M. R. Islam, and K. Kwak, "Energy saving mechanisms for MAC protocols in wireless sensor networks," *International Journal of Distributed Sensor Networks*, vol. 2010, Article ID 163413, 16 pages, 2010.

- [2] Shio Kumar Singh, M. P. Singh, and D. K. Singh, "A survey of energy-efficient hierarchical cluster-based routing in wireless sensor networks," *International Journal of Advanced Networking and Applications*, vol. 2, no. 2, pp. 570–580, 2010.
- [3] A. A. Abbasi and M. Younis, "A survey on clustering algorithms for wireless sensor networks," *Computer Communications*, vol. 30, no. 14–15, pp. 2826–2841, 2007.
- [4] W. B. Heinzelman, A. P. Chandrakasan, and H. Balakrishnan, "An application-specific protocol architecture for wireless microsensor networks," *IEEE Transactions on Wireless Communications*, vol. 1, no. 4, pp. 660–670, 2002.
- [5] O. Younis and S. Fahmy, "HEED: a hybrid, energy-efficient, distributed clustering approach for ad hoc sensor networks," *IEEE Transactions on Mobile Computing*, vol. 3, no. 4, pp. 366–379, 2004.
- [6] Y. Challal, A. Ouadjaout, N. Lasla, M. Bagaa, and A. Hadjidj, "Secure and efficient disjoint multipath construction for fault tolerant routing in wireless sensor networks," *Journal of Network and Computer Applications*, vol. 34, no. 4, pp. 1380–1397, 2011.
- [7] W. Lou and Y. Kwon, "H-SPREAD: a hybrid multipath scheme for secure and reliable data collection in wireless sensor networks," *IEEE Transactions on Vehicular Technology*, vol. 55, no. 4, pp. 1320–1330, 2006.
- [8] Y. Jing, X. Mai, X. Jinfu, X. Baoguo, and H. Lu, "A cluster-based multipath delivery scheme for wireless sensor networks," in *Proceedings of the 2nd IEEE International Conference on Broadband Network and Multimedia Technology (IEEE IC-BNMT'09)*, pp. 286–291, Beijing, China, October 2009.
- [9] S. M. Xiong, L. M. Wang, and J. Y. Wu, "Energy-efficient hierarchical topology control in wireless sensor networks using time slots," in *Proceedings of the 7th International Conference on Machine Learning and Cybernetics (ICMLC'08)*, pp. 33–39, chn, July 2008.
- [10] W. Ye, J. Heidemann, and D. Estrin, "An energy-efficient MAC protocol for wireless sensor networks," in *Proceedings of the 21st Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM'02)*, vol. 3, pp. 1567–1576, 2002.
- [11] C. Karlof and D. Wagner, "Secure routing in wireless sensor networks: attacks and countermeasures," *Ad Hoc Networks*, vol. 1, no. 2–3, pp. 293–315, 2003.
- [12] H. Huang, Y. Xu, Y. E. Sun, and L. Huang, "Cluster-based load balancing multi-path routing protocol in wireless sensor networks," in *Proceedings of the 7th World Congress on Intelligent Control and Automation (WCICA'08)*, pp. 6686–6691, Chongqing, China, June 2008.
- [13] N. Sun, Y. B. Cho, and S. H. Lee, "A distributed energy efficient and reliable routing protocol for wireless sensor networks," in *Proceedings of the 14th IEEE International Conference on Computational Science and Engineering*, pp. 273–278, Liaoning, China, August 2011.
- [14] A. Boukerche, R. W. N. Pazzi, and R. B. Araujo, "Fault-tolerant wireless sensor network routing protocols for the supervision of context-aware physical environments," *Journal of Parallel and Distributed Computing*, vol. 66, no. 4, pp. 586–599, 2006.
- [15] G. Gupta and M. Younis, "Fault-tolerant clustering of wireless sensor networks," in *Proceedings of the IEEE Conference on Wireless Communications and Networking*, vol. 3, pp. 1579–1584, New Orleans, La, USA, March 2003.

