

Integrated data aggregation with fault-tolerance and lifetime energy-aware adaptive routing in coffee plantations using WSNs

Roshan Zameer Ahmed¹, Sravani K.², Shilpa S. Chaudhari³, S. Sethu Selvi⁴, S. L. Gangadharaiah⁵

^{1,2,4,5}Department of Electronics and Communication Engineering, Bengaluru, India

³Department of Computer Science and Engineering, Bengaluru, India

^{1,2,3,4,5}M. S. Ramaiah Institute of Technology, Bengaluru, India

Article Info

Article history:

Received Jul 6, 2021

Revised Aug 20, 2021

Accepted Aug 23, 2021

Keywords:

Aggregation

Energy-efficiency

Fault-tolerance

Routing

ABSTRACT

The pest namely coffee white stem borer (CWSB) has harmed the economic progress of many emerging countries as a result of arabica coffee's agricultural products. The boring activity causes the stem to shrink, fade in color, and acquire translucent margins across the stem. The pest multiplier can be controlled by capturing the location with the utilization of a wireless sensor networks (WSNs) and blocking its exit point at the user end. In this work, we propose an integrated data aggregation with fault-tolerance and lifetime energy-aware adaptive routing (IDALAR) approach to transfer the sensed pest location data. The efficient packet format and statistical models based routing between clusterheads (CHs) and base station (BS) is proposed considering the availability of resources such as message overhead, algorithmic complexity, residual energy, and control overhead are all used to calculate its performance.

This is an open access article under the [CC BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.



Corresponding Author:

Roshan Zameer Ahmed

Department of Electronics and Communication Engineering

M. S. Ramaiah Institute of Technology, Bengaluru-560054, India

Email: roshanahmed31@gmail.com

1. INTRODUCTION

Coffee production is an important asset in the economic growth of many developing nations. Among the essential crops available, coffee positions the most helpful farming item all through the world. Coffea arabica and coffea robusta are two varieties of coffee cultivated on a large commercial scale [1]. Arabica coffee is most preferred in comparison to robusta coffee, due to the quality taste, and aroma. In countries such as Vietnam, Sri Lanka, Nepal, Java, Burma, China, and Thailand, the xylotrichus quadripes, also known as coffee white stem borer (CWSB), is the most harmful pest to arabica coffee plants. It was reported in India in the early 1830's [2]. Robusta coffee is not affected by the influence of CWSB as its stem and primary branches exhibit symptomatic ridges not giving room for the larval activity, due to which females do not prefer laying eggs. The female borer lays 50-100 eggs on the stem of an arabica plant to start the CWSB process. The grub is made up of these larvae, which are 1.26 mm long and 0.48 mm wide, and breed in 10-12 days. By boring in various directions, the grub creates galleries within the fundamental stem and vital branches without intersecting one another, collecting food and water supplies from the stem. To close the entry points, they cover the exterior layer of the stem with the remaining boring material. The boring activity is continued for a year. After a month, the pupa turns into a fully grown beetle, which emerges from the stem with a

void hole [3]. The grub's boring operation results in the creation of cavities varying in size from 4 mm to 1 inch. By trapping the pest's region and blocking its path, the pest's reproduction is also delayed. Integrated pest management (IPM) involving cultural, mechanical, and chemical methods had been practiced to manage CWSB activity but fell short in managing the crop losses due to destructive pest activity [4].

Wireless sensor networks (WSNs) is an emerging technology with one of its major applications used in agriculture for monitoring various parameters. WSNs based pest identification help farmers to monitor crops effectively, eradicate destructive pests and prevent disease spread at early stages [5]. The proposed work makes use of ultrasonic active sensors (UAS) that can be used to prove the existence of grubs in the cavity region. The reflected sound emitted by boring is measured at a distance of 10-15 meters at a frequency of 40.780 KHz. The reflected acoustic wave with the frequency shift is examined by the UAS receiver, which triggers an alert. The sensor nodes on the identification of pest-related data, communicate it to the BS by the CHs. These processes such as clustering, data aggregation, and routing help the information reach the BS in an energy-efficient manner maintaining network lifetime, scalability, and fault-tolerance. The entire coffee field scenario was simulated on QualNet 5.2 network simulator [6].

The remaining part of the paper is set out as follows. Section 2 explores the related works. The proposed fault-tolerance and lifetime energy-aware adaptive routing (IDALAR) system, which includes cluster-based data aggregation for sensed data transfer to base station (BS) while considering fault tolerance, is discussed in Section 3. In Section 4, network simulations using QualNet 5.2 are used to evaluate the proposed architecture. Finally, Section 5 concludes the work done.

2. RELATED WORKS

The section discusses the existing works on cluster-based data aggregation, fault tolerance, and energy-adaptive routing in comparison to the proposed IDALAR technique. The AVR-INJECT tool, described in [7], automates fault injection and analysis on WSN nodes. The program use software to simulate the insertion of hardware flaws into assembly code, such as bit flips. The categorization of defects for adding the model and rating performance has not been addressed. In [8], Hierarchical cluster-based data aggregation with multiple sinks is proposed, which reduces redundancy while boosting energy efficiency. The virtual cluster head (VCH) was chosen because of the relevance of the CHs and to enhance their fault tolerance. The flexibility given by dynamically created groups in each cluster represents a benefit as compared to the opposing aggregation approaches. The study in [9] proposed a new strategy that re-configures the healthy nodes and reconstructs the aggregation tree while sidelining the problematic nodes. Different inquiries were made at the source using a conventional querying algorithm, leading the data gradients inside the network to show up to the details. The authors proposed the minimum capacitated overid (MCO) spanning tree in [10], the technique takes into account bit complexity as well as the simple implementation of the esau-williams algorithm. The data collected by the child node is sent to the parent node, which subsequently sends it to the sink. When the scheme fails, it's usually because the child node is unable to communicate with the parent node. In [11], the adaptive cluster determination technique selects CHs based on the node's distance to the BS and the node's leftover resources. The fuzzy logic-based ranking table is used to select the cluster-head in [12]. It utilizes multi-tier clustering to reduce energy consumption while increasing network lifespan in WSNs. The cluster formation is categorized as spontaneous, adaptive, or self-clustering by low energy adaptive clustering hierarchy (LEACH) for clustered aggregation routing in WSNs [13]. The algorithm was defined in two phases: set-up and reliability. Cluster formation and CH assortment are included in the setup process, and the information transmission is scheduled in the reliable phase. Nodes can only transmit information in their respective time slot with perfect synchronization required to be maintained. Data routing in-network aggregation (DRINA) aim at establishing data synchronization among the nodes [14]. The information is sent by a node, solely after awaiting information from its neighbors. Since there are likely to be more aggregation points, the time it takes for aggregators to collect data from another node should be reduced. To maintain balanced energy consumption, each node selects the next-hop node in the opposite direction of the BS in distributed energy balance routing (DEBR). This unnecessarily results in increased delay for transmitting data towards the BS. If the selected node has no next-hop node within its range, data packets cannot enter the BS [15]. The fault-tolerant clustering (FC) [16] works on two-phase of detecting and recovering faulty CHs without re-clustering or shutting down the system. The method implemented resulted in enormous energy consumption. The work proposed in [17] is the graph flow modeling by creating virtual CHs to reduce energy consumption and to efficiently tolerate the

CHs failures. The faculty CHs and the rest failure nodes were not been differentiated in terms of their transmission loads which resulted in unbalanced energy consumption. Fast local clustering (FLOC), which divides multi-hop WSNs into non-overlapping clusters of nearly equal size, is discussed in [18]. With clustering localization, it achieves fault-local self-adjustment. In [19], the authors present a message authentication code method that employs a secret key made out of biological features represented by deoxyribonucleic acid and the output of a blum shub random number generator, as well as a unique hash technique. The authors of [20] suggested a low-weight authentication system that employs implicit certificates to allow mutual authentication and key agreement across IoT devices using implicit certificates. Lavish Kansal *et al.* [21] proposed analyzing high-level modulations on an orthogonal frequency division multiplexing (OFDM) system. The effects of additive white gaussian noise (AWGN) and rayleigh channels on BER for high data rates have been shown using AWGN and Rayleigh channels for study. The comparative parameter analysis is shown in Table 1.

Table 1. Comparative parameter analysis

Proposed Work	Aggregation Ratio (AR)	End-To-End Delay (ETED)	Message Overhead (MO)	Control Overhead (CO)	Algorithmic Complexity (AC)	Clustering Time (CT)
CBDAMS [8]	Low	Average	Average	Average	$O(n^2)$	Average
SRA [9]	High	Low	Average	Average	$O(\log n)$	Average
MCO [10]	Average	Average	High	High	$O(n \log n)$	Average
DRINA [14]	High	Average	Average	Average	$O(\sqrt{n})$	Average
DEBR [15]	Average	Average	High	High	$O(n^2)$	Low
FC [16]	Average	Average	High	High	$O(n \log n)$	Average
IDALAR	High	Low	Low	Low	$O(n)$	Low

3. PROPOSED IDALAR SYSTEM

The proposed IDALAR discussed in this section aims to effectively send aggregated sensed pest location data to the BS through clustering considering network lifetime and a multipath-based fault tolerance approach. The energy and hop distance aware CH selection for routing of aggregated data and identification of faulty nodes in fault tolerance is performed by employing statistical models of network lifetime, hop distance, and residual energy. Our contributions to this work are as follows; (1) design and develop the clustering technique with CH selection based on node status, energy, and hop distance for routing of aggregated data in WSNs of uniformly spaced UAS, (2) design and develop sensor location data polling at CH with data aggression technique, (3) design and develop network lifetime aware routing scheme considering node/CH residual energy, and hop distance from each CH to BS to minimize the delay, (4) defining route monitoring of the established route and potential route adaptation when connected with similar CHs. The proposed routing support fault tolerance for established routes using the Weibull distribution-based multipath routing technique, (5) simulation analysis for performance comparison of the proposed IDALAR system in terms of message overhead, algorithmic complexity, control overhead, and energy consumption.

3.1. Network model

In the coffee field scenario, the nodes are uniformly distributed in a specific order to monitor 10-12 coffee stems for CWSB identification. The field area is clustered so that it covers all of the nodes with the fewest clusters possible. The proposed IDALAR scheme takes into account two types of nodes: sensor nodes and CH nodes, which act in a hybrid model of the event-driven or time-driven concept. Detected CWSB event is immediately and efficiently transmitted using event-driven model through their respective CHs [22] where received signal strength-based distance used to find multi-hop routes to BS node [23]. A time-driven approach is utilized in absence of CWSB event results to indicate sensor node aliveness statistics. Hop distance, network lifespan, node energy, and multipath-based fault tolerance are used to determine the best route for transmitting an event or time-driven data. The node lifespan N_{LT_i} is determined by the time of battery level from the start of the first data transmission until the node's battery is drained. If E_{res} is the residual energy of node i at time t , then the lifespan of the node is $N_{LT_i} = \inf(t | E_{res} = 0)$.

The time interval between the start of data transmission at a node and the Infimum/minimum time from the set of the time at which the remaining battery level at the node reaches zero reflects the network lifespan of a node. To examine the lower bound value within the set of time values from when the node is idle, we utilize the \inf function. The sensor network lifespan $Z_{net(max)}$ is the time period between the

commencement of data transmission and the death of the first node i among the set of nodes as $Z_{net}(max) = \inf(N_{LT_i} | i = 1, 2, 3, \dots) | E_{res_i} = 0$. When a node's residual energy reaches zero, the node is said to be dead. The BS node is assumed to have learned the network topology through topology discovery [24] and to be able to screen the energy usage and residual energy of each node.

3.2. Energy utilization and residual energy of node

Both free-space and multipath fading channels are covered by the [25] radio model for energy utilization. If the distance is less than the threshold value d_0 , the free space ((f_s)) model is used; otherwise, the multipath (mp) model is used. Let E_{elec} denote the energy acquired by the circuit, which is based on signal filtering and spreading, and ϵ_{fs} , ϵ_{mp} , respectively, signify the energy required in free space and multipath.

$$E_T(l, d) = \begin{cases} lE_{elec} + l\epsilon_{fs}d^2 & \text{for } d < d_0 \\ lE_{elec} + l\epsilon_{mp}d^4 & \text{for } d \geq d_0 \end{cases} \quad (1)$$

As shown in (1) is used to determine the amount of energy required by the radio to transmit an 1-bit message across a distance d . The amount of energy required by the radio to receive an 1-bit message is calculated as $E_R(l) = lE_{elec}$. The residual energy present at each node is calculated using (2), which is then used to determine the network lifespan.

$$E_{res}\% = \frac{E_{res_i} - E_{th}}{E_{total}} * 100 \quad (2)$$

The threshold energy, denoted by the symbol E_{th} , is the minimal energy below which the network cannot operate. The residual energy is calculated by subtracting the available energy from the sum of Equations 1 and $E_R(l)$. E_{total} is the initial total energy, which is lowered with each data transmission and reception.

3.3. Cluster-based data aggregation with fault tolerance distribution

The data aggregation model depicted in Figure 1(a) is discussed in this section. A cluster-head database, BS database, WSNs initialization, and several modules, including; (a) clustering, (b) sensor data polling, (c) sensor information collection by cluster-head, (d) aggregation at cluster-head, and (e) data assortment by BS that operates in iterations, are all defined by the proposed technique. WSNs are initialized during clustering, and the data aggregation measure is done during iteration.

Once node deployment is complete, the clustering of nodes with the CH selection is established and maintained at each node. The cluster head is chosen based on the node state, remaining energy, and hop distance. As depicted by the hybrid model of WSNs, the CHs perform data aggregation and send the obtained data to the BS. The CH and the BS are in charge of the cluster's details. These modules communicate with one another through packet handshaking. Each sensor node has its unique identification number, which is recorded as SN_{ID} in these packets. A unique identification number CH_{ID} is assigned to the sensor node that serves as the cluster head. Both the SN_{ID} and the CH_{ID} are numeric and are allocated in order, starting at 0. The format of a peer packet comprises the following elements: sort, sender, recipient, and information to be sent.

Clustering and CH Selection Technique: The scenario is separated into non-overlapped and equal size clusters according to the sensor node's solid disc clustering property (S_{DC}) in FLOC [18]. The two-band wireless radio model is used for indicating inner band range for 1-hop distance node from the CH while outer band range for 2-hop distance nodes from the CH. Every deployed node maintains track of its status, cluster-id, cluster members, cluster-head data, distance to the BS, residual energy, and an invitation from another CH to join its cluster. To increase network competency, another CH should be chosen to relieve the existing CH of aggregation duties.

The state of each node changes during the clustering process, as illustrated in Figure 1(b), and given as follows; (1) IDLE state-which implies the node is not connected to any cluster, (2) ReqCL state-a node demands a CH position if none are available within its communication range, (3) CH state-which means the node is carrying out the CH obligations, (4) IB state-the node is a member of the CH's inner band within its reach, (5) OB state-the node is a member of the CH's outer band within its range.

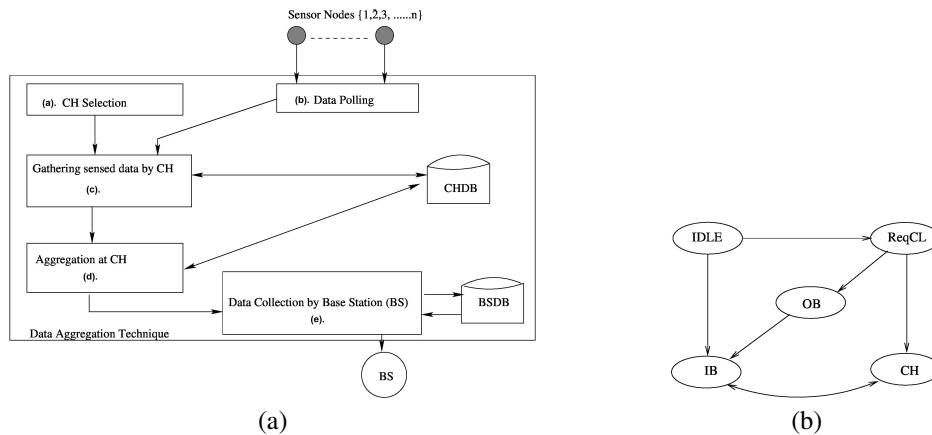


Figure 1. These figures are; (a) data aggregation approach: logical perspective, (b) node's states over the course of its lifespan

Before communication may occur, a node must be manually assigned to the CH or IB state. During communication, nodes in the CH state, which is the CH, spend more energy than nodes in the IB state. When the available energy goes below or equals the limit, changes in the state of the nodes start the clustering process. Various arbitrary timers are utilized throughout the clustering phase to execute activities such as transmitting heartbeats, time-out of a node while waiting for a condition, and so on.

Sensor data aggregation technique: The CH keeps track of each cluster member's data, including its status. The likelihood of existence or absence of the pest event is indicated using polling as 0 or 1 to the CH by sensor nodes. The 2-bit record at a node indicates its state and action along with SN_{ID} , ClusterHD, cluster member list, distance to BS, residual energy, an invitation from another CH, and current status. The MSB value of 1 in 2-bit record indicates existence of pest event on at least one cluster member while 0 indicates pest absence in cluster. The LSB value of 1 in 2-bit record indicates that the action is taken while 0 indicates action not taken by the node. The detected event packet data received by respective CH CH_{ID} extracts SN_{ID} from and searches its entry in the already existing information during this period.

Algorithm 1 At Cluster Head: Level of Status Collection

```

1: Input: ( $S[1], S[2], \dots, S[n]$ )
2: Output: Depth of Level
3: Begin
4:  $i=0$  and  $j=1$ ;
5: while ( $i < n$ ) do
6:   Get input data as  $status == S[i]$ ;
7:   Aggregated the data at  $d_j$ ;
8:   Store aggregated data as  $status$ ;
9:   if ( $S[i] = 1$ ) then
10:    store  $S[i]$  sensor id in sensor-list;
11:   end if
12:    $i++$ ; and  $j++$ ;
13: end while
14: End

```

If it exists, CH does nothing otherwise SN_{ID} status is changed to 1, indicating the presence of the CWSP pest, which is delivered to the BS. If more than one member informs pest existence, the duplicated information is ignored by CH for better energy usage using kolmogorov's zero-one rule, as illustrated in Algorithm 1. Every CH that is an intermediary node on the path constructed towards BS receives aggregated data from sensor list S without any duplicated information. The depth or the number of intermediate nodes i are indicated as d_i . The BS keeps track of each CH's data set as well as the status of each cluster member. In the

action taken field, the status of each affected cluster member is updated. Algorithm 2 depicts the BS node's aggregated data collecting method. The BS consolidates the *ACK* packet and transfers it to each CH of received data aggregation packets by including a list of the CH in the *ACK* packet. After resolving the recognized pest in the predetermined region of the affected stem, the BS administrator sends the action taken packet to the appropriate CH with the cluster members. Finally, identical changes were made to the BS database.

Algorithm 2 Aggregated Data Collection at BS

```

1: Input: BS database;
2: Begin
3: Schedule event to receive data aggregation packet;
4: Extract cluster-head list from data aggregation packet
5: for (Each cluster-head ) do
6:   Extract corresponding affected sensor-list;
7:   Update BS database for affected sensor status;
8:   Add cluster-head on  $CH_{ACK}$  list for sending ACK;
9: end for
10: Frame ACK packet using  $CH_{ACK}$  list;
11: Send ACK packet;
12: while (ActionTaken = 1) do
13:   Add sensor for each cluster-head in actionTaken packet;
14: end while
15: if actionTaken packet then
16:   Send actionTaken packet;
17:   Update BS database for action taken nodes;
18: end if
19: End

```

Fault-tolerance approach: The fault model of WSNs is a result of sensor node or link malfunction. Sensor nodes are prone to loss due to hardware failure or disruption caused by an external event, as well as battery power depletion. The WSNs fault model takes into account faulty CHs using a reliability function that follows the weibull distribution, resulting in a variety of failure patterns. Hardware component malfunction, energy depletion, and permanent link failure are all considerations taken into account when designing the model. The CH can fail at any time during the network setup process due to energy depletion or damage. The fault can be found in the following two scenarios; (a) each intermediate node indicated as $ComCH(S_i)$ broadcasts the help message to the neighbor cluster head indicated as CH_G when the intermediate node does not receive any data acknowledgment, (b) when the CH failure is verified by its cluster members S_i , each S_i broadcast the help message to the neighbor cluster members indicated as S_j within their transmitting range. The number of neighbors who receive this message $NodeSet_{Received-HELP}$ are equal to $\forall S_j \in Neighbor S_i$, and the number of CHs who respond to the message help within the range $CHSet_{Received-HELP}$ are equal to $\forall CH_G \in ComCH(S_i)$.

If the node receives a response from other CHs, it remains a member of the covered set CO_{set} , otherwise, it becomes a member of the uncovered set $UnCO_{set}$. Covered nodes CO_{set} are a group of nodes that have at least one CH in their communication range, while uncovered nodes $UnCO_{set}$ do not. The nodes S_i receive responses from different nodes, which causes the backup collection of nodes to be reset. Despite the existence of several routes, the nodes S_j join a CH based on the cost value, as show in (3). The next-hop CH with the most residual energy and the smallest separation from the sender CH and the BS is chosen by the cost function. The fault-tolerance model assumes the cost value factor and its state as it is run by the nodes to track down the defective CH.

$$CH_{cost}(CH_i, S_j) = \frac{E_{res}(CH_i)}{dis(S_j, CH_i) \times dis(CH_i, CH_{m+1})} \quad (3)$$

3.4. Lifetime energy based adaptive routing

The transmission produced by the network's nodes, a lot of energy was wasted. The proposed counter-based approach in a fixed topology with clustering uses the concept of inner-band (1-hop distance with neighbor

nodes) and outer-band (2-hop distance with border nodes) range to select cluster members within a homogeneous network structure. The CH established routes to the BS node by broadcasting route request (RREQ) packets to neighbor nodes in the 1-hop region, which then traverses with the 2-hop range nodes within the cluster, and finally with the remaining nodes of the other cluster. The nodes receiving the RREQ packet will have RREQ from 2-3 neighbor nodes when traversing it. Just one RREQ from the same *Seq-no* should be considered in this situation, along with T_s from the neighbor node. As neighbor nodes send RREQs, the packet field data is gathered and saved. The buffer capacity of the node involved in route establishment must be less than the threshold ($P_{bc} < P_T$), which is one of the requirements. If the state is not fulfilled, the node will start losing packets that are sent to it. As BS receives the RREQ at the destination, it will accept them all at the same time T_S from the same CH. The T_S variable is initially set to a default value x , but when the RREQ packet traverses the network, the T_S variable is decremented ($T_S - 1$) before it reaches the BS node. The packet expires if the T_S is already present in the loop as it reaches a value of zero. The RREP packet is generated at the BS node based on the RREQ packet gathered, which is given by $Si=Simax$, and the path to the source.

4. SIMULATION MODEL AND RESULT ANALYSIS

To help explain and evaluate the methodology's performance and expertise, the QualNet 5.2 Network Simulator was used to model IDALAR. Three models make up the simulation environment: network, traffic, and propagation. The network is made up of " N " nodes, " r " is the transmitting range of each node, and β is the free space propagation constant used. The simulation results obtained with the proposed IDALAR provide the following results, which are analyzed; (1) message overhead (MO): The total number of messages sent and received over a certain period, (2) algorithmic complexity (AC): The total number of operations performed throughout the clustering and data aggregation process, (3) energy consumption (EC): The average amount of energy spent at the CH during the aggregation process, (4) control overhead (CO): The total amount of control packets required to aggregate the data at the CH.

4.1. Analysis of MO

At various points during simulation time, the message overhead in IDALAR is compared to FC and DEBR. As show in Figure 2(a), aggregation at CH in IDALAR results in a decrease in message overhead as compared to FC and DEBR. The reduction in MO while using IDALAR is due to the use of an aggregation approach on the received message packet from each sensor node during CWSB pest detection at CH before delivering it to the BS. In FC and DEBR, message packets are used extensively to deliver data of interest to its destination node as quickly as possible, resulting in increased message traffic and overhead.

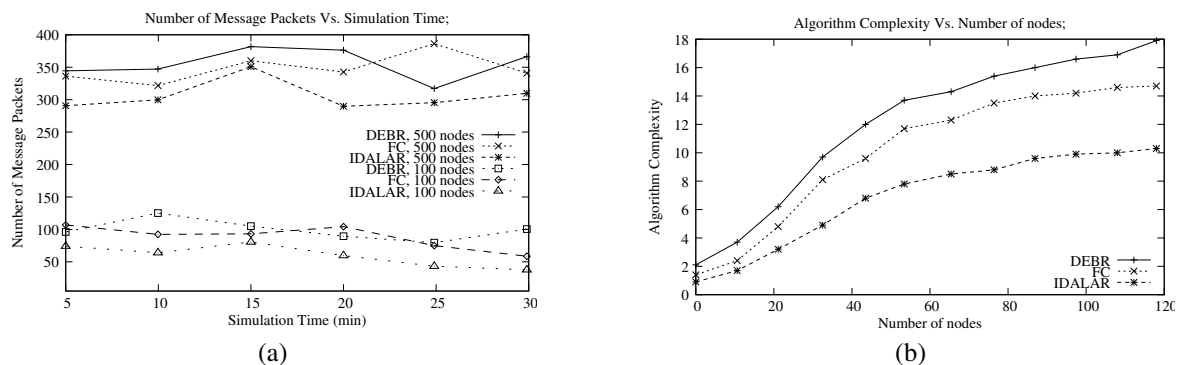


Figure 2. These figures are; (a) no. of message packets v/s simulation time, (b) algorithm complexity v/s no. of nodes

4.2. Analysis of AC

Big O notation is used to calculate the algorithm's complexity, which is defined as the total number of operations executed. As shown in Figure 2(b), it is computed by taking N as the number of nodes $N=1,2,3,\dots,n$ and (s) as the number of steps completed. The complexity of the proposed algorithm rises linearly $O(n)$. This is because the algorithmic complexity increases as the number of nodes in a scenario grow since more

processes must be executed sequentially. As a result, the complexity of algorithms varies. FC has a $O(n \log n)$ complexity, whereas DEBR has a $O(n^2)$ complexity.

4.3. Analysis of EC

The CHs residual energy for DEBR, FC, and IDALAR is shown in Figure 3. Even if the remaining CHs contain a lot of energy, the residual energy of the CHs lined up adjacent to the BS is practically similar for the data routed. IDALAR’s CHs residual energy is larger than DEBR and FC’s because its energy consumption follows the Weibull distribution, which preserves uniformity over time.

4.4. Analysis of CO

Control overhead grows as the number of nodes increases, as shown in Figure 4(a). Because IDALAR functions in both the inner and outer bands, the number of control packets required is limited to those required to establish a connection within that band. Because of the cluster’s multi-hop connection, the FC and DEBR require extra control packets. As a result, as compared to IDALAR, the number of control packets required is greater, resulting in higher control overhead. At different simulation time instants, the CO is significantly lower than the FC and DEBR, as shown in Figure 4(b). Because IDALAR uses Kolmogorov’s zero-one rule to remove duplicate packets and send them without interruption, fewer control packets are produced. In addition to queuing, FC utilizes signal strength evaluation and matching with a sequence number to detect duplicate packets, whereas DEBR uses signal strength assessment and matching with a sequence number. All of these methods need the nodes and the CH to exchange additional control packets.

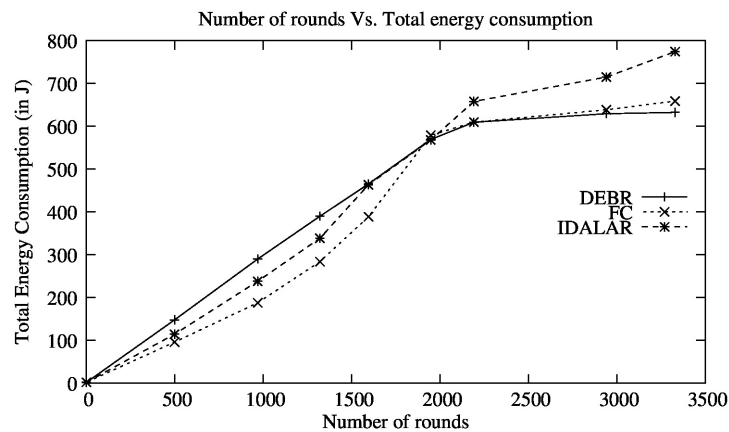


Figure 3. No. of rounds versus total energy consumption

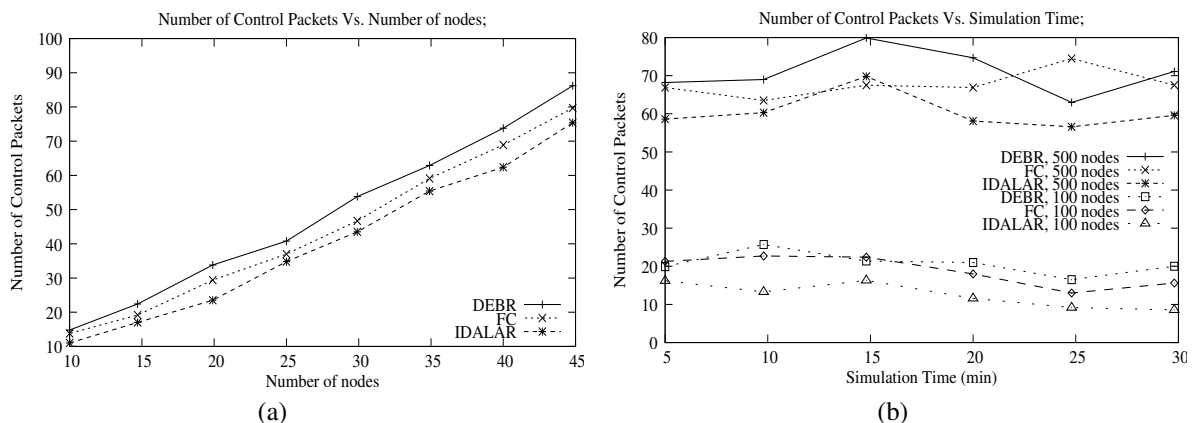


Figure 4. These figures are; no. of control packets v/s (a) no. of nodes, (b) simulation time

5. CONCLUSION

This paper proposed an integrated data aggregation with fault-tolerance and lifetime energy-aware adaptive routing approach to locate the CWSB pest within the coffee stem using WSNs. Clustering, CH selection, and data aggregation are all taken into consideration as part of the scheme's fault-tolerance approach. The aggregated data are provided to the BS for further processing by establishing a route using the lifetime energy-based adaptive routing protocol. When compared to DEBR and FC, the proposed IDALAR minimizes message overhead, control overhead, algorithmic complexity, and energy consumption.

REFERENCES

- [1] Venkatesha M. G., and Dinesh A. S., "The Coffee White Stem Borer *Xylotrechus Quadripes* (Coleoptera: Cerambycidae): Bioecology, Status and Management," *International Journal on Tropical Insect Science*, vol. 32, no. 4, pp. 177-188, 2012, doi: 10.1017/S1742758412000331.
- [2] Le Pelley R. H., "Pests of Coffee," *Longmans, Green and Co. Ltd.*, pp. 1-590, 1968.
- [3] Santosh P., Sreenath H. L., Vinod Kumar P. K., Seetharama H. G., Jayarama, "New record of Coffee White Stem borer *Xylotrechus Quadripes* Chevrolat on *Psilanthus Bengalensis* in India," *Journal of Plantation Crops*, vol. 39, no. 3, pp. 396-398, 2011.
- [4] Ahmed, R. Z., Biradar, R. C., and Chaudhari, S. S., "Cluster-Based Data Aggregation for Pest Identification in Coffee Plantations using Wireless Sensor Networks," *Computers and Electrical Engineering*, Elsevier, vol. 56, pp. 591-607, 2016, doi: 10.1016/j.compeleceng.2016.05.001.
- [5] Sarika Datir, and Sanjeev Wagh, "Monitoring and detection of agriculture disease using WSNs," *International Journal of Computer Applications*, vol. 8, pp. 887-902, 2014, doi: 10.5120/15193-3573.
- [6] Roshan Zameer Ahmed, and Rajashekhar C. Biradar, "Redundancy aware data aggregation for pest control in coffee plantation using wireless sensor networks," In *2nd International Conference on Signal Processing and Integrated Networks (SPIN)*, 2015, pp. 984-989, doi: 10.1109/SPIN.2015.7095271.
- [7] Cinque, M., Cotroneo, D., Di Martino, C., Russo, S., and Testa A., "AVR-INJECT: A tool for injecting faults in Wireless Sensor Nodes," In *IEEE International Symposium on Parallel and Distributed Processing*, 2009, pp. 1-8, doi: 10.1109/IPDPS.2009.5160907.
- [8] Tanushree M., Roopa H., and Vani V., "Hierarchical Cluster based Data Aggregation with Fault Tolerance to extend Network Lifetime," In *Second International Conference on Inventive Research in Computing Applications (ICIRCA-2020)*, 2020 pp. 785-789, doi: 10.1109/ICIRCA48905.2020.9183105.
- [9] Beneyaz A. Begum, and Satyanarayana V. Nandury, "Self-Reconfiguration Aggregation Scheduling to Recover from Node and Interference Faults in WSNs," In *2017 International Conference on Wireless Communication, Signal Processing and Networking (WiSPNET)*, 2017, pp. 1512-1517, doi: 10.1109/WiSPNET.2017.8300014.
- [10] Mustafa A., Can Umut Ileri, and Orhan Dagdeviren, "An Energy-Efficient Capacitated Minimum Spanning Tree Algorithm for Topology Control in Wireless Sensor Networks," In *25th Signal Processing and Communications Applications Conference (SIU)*, 2017, pp. 1-4, doi: 10.1109/SIU.2017.7960603.
- [11] Neeti Jain, and Prakriti Trivedi, "An adaptive sectoring and cluster head selection based multi-hop routing algorithm for WSN," In *2012 Nirma University International Conference on Engineering (NUICONE)*, 2012, pp. 1-6, doi: 10.1109/NUICONE.2012.6493208.
- [12] Wan Isnii Sofiah Bt Wan Din, and Saadiah Yahya, "Ranking Table for Cluster Head Selection of Wireless Sensor Networks," In *2012 International Conference on Advanced Computer Science Applications and Technologies (ACSAT)*, 2012, pp. 100-104, doi: 10.1109/ACSAT.2012.81.
- [13] Abdul Rehman Khan, Nitin Rakesh, Abhay Bansal, and Dev Kumar Chaudhary, "Comparative study of WSN Protocols (LEACH, PEGASIS and TEEN)," In *Third International Conference on Image Information Processing (ICIIP)*, 2015, pp. 422-427, doi: 10.1109/ICIIP.2015.7414810.
- [14] Villas, L. A., Boukerche, A., Ramos, H. S., De Oliveira, H. A. F., de Araujo, R. B., and Loureiro, A. A. F., "DRINA: A Lightweight and Reliable Routing Approach for In-Network Aggregation in Wireless Sensor Networks," *IEEE Transactions on Computers*, vol. 62, no. 4, 2013, doi: 10.1109/TC.2012.31.
- [15] Chang-Soo Ok, Seokcheon Lee, Prasenjit Mitra, and Soundar Kumara, "Distributed energy balanced routing for wireless sensor networks," *Computers and Industrial Engineering*, vol. 57, no. 1, pp. 125-135, 2009, doi: 10.1016/j.cie.2009.01.013.
- [16] Gupta G., and Younis M., "Fault-tolerant clustering of wireless sensor networks," In *2003 IEEE Wireless Communications and Networking (WCNC)*, 2003, pp. 1579-1584, doi: 10.1109/WCNC.2003.1200622.
- [17] Jenn-Wei Lin, Pethuru Raj Chelliah, Meng-Chieh Hsu, and Jia-Xin Hou, "Efficient Fault-Tolerant Routing in IoT Wireless Sensor Networks Based on Bipartite-Flow Graph Modeling," *IEEE Access*, vol. 7, pp. 14022-14034, 2019, doi: 10.1109/ACCESS.2019.2894002.

- [18] M. Demirbas, A. Arora, V. Mittal, and V. Kulathumani, "A Fault-Local Self-Stabilizing Clustering Service for Wireless Ad Hoc Networks," *IEEE Transactions on Parallel and Distributed Systems*, vol. 17, no. 9, pp. 912-922, 2006, doi: 10.1109/TPDS.2006.113.
- [19] Gurpreet Kour Sodhi *et al.*, "Preserving Authenticity and Integrity of Distributed Networks through Novel Message Authentication Code," *Indonesian Journal of Electrical Engineering and Computer Science (IJECS)*, vol. 12, no. 3, pp. 1297-1304, 2018, doi: 10.11591/ijeecs.v12.i3.pp1297-1304.
- [20] Valmiki Siddhartha, Gurjot Singh Gaba, and Lavish Kansal, "A Lightweight Authentication Protocol using Implicit Certificates for Securing IoT Systems," *Procedia Computer Science*, vol. 167, pp. 85-96, 2020, doi: 10.1016/j.procs.2020.03.185.
- [21] Lavish Kansal, Ankush Kansal, and Kulbir Singh, "Analysis of Different High Level Modulation Techniques for OFDM System," *International Journal of VLSI and Signal Processing Applications*, vol. 1, no. 2, pp. 102-107, 2011.
- [22] Xu, J., Liu, W., Lang, F., Zhang, Y., and Wang, C., "Distance Measurement Model Based on RSSI in WSN," *Wireless Sensor Network*, vol. 2, no. 8, pp. 606-611, 2010, doi: 10.4236/wsn.2010.28072.
- [23] Sungju Lee *et al.*, "An Energy-Efficient Distributed Unequal Clustering Protocol for Wireless Sensor Networks," *International Journal of Electronics and Communication Engineering*, vol. 2, pp. 2715-2719, 2008.
- [24] A. A. Nezhad, D. Makrakis, and A. Miri, "Anonymous topology discovery for multihop wireless sensor networks," In *3rd ACM Workshop on QoS and Security for Wireless and Mobile Networks*, 2007, pp. 78-85, doi: 10.1145/1298239.1298254.
- [25] Chandrakasan A. P., Heinzelman W. B., and Balakrishnan H., "An application specific protocol architecture for wireless microsensor networks," *IEEE Transactions on Wireless Communications*, vol. 1, no. 4, pp. 660-670, 2002, doi: 10.1109/TWC.2002.804190.