RESEARCH ARTICLE

SAMRIDDHI Volume 14, Issue 4, 2022 Print ISSN: 2229-7111 Online ISSN: 2454-5767

Relative Entropy-based Energy Efficient Multi-hop Forwarding Routing Protocol for Wireless Sensor Network

Akant K. Raghuwanshi, Virendra S. Chaudhary

Department of Electronic and Communication Engineering, RKDF University, Bhopal, Madhya Pradesh, India

Abstract

Recently, cluster distribution and selection have been made in the most advanced protocols based on Low Power Adaptive Clustering Hierarchical Protocol (LEACH), followed by cluster head election. Conversely, it is not thought that data redundancy is sufficient. This study suggests a modified Leach approach incorporating relative information entropy (Leach-CIE). The relative entropy value of the data's probability distribution in the node's two subsequent periods is calculated in conjunction with the close entropy model. The information is then compared with a threshold value to determine the degree of redundancy. The network accomplishes reducing energy consumption by refusing to transmit redundancy data. Instead, a multi-hop routing (MFRCRE) considers its residual energy, the energy consumption of each node in the network during the data transmission phase, by optimizing node forwarding conditions and considering factors such as communication distance and energy consumption. It is proposed to balance the rate—account ratio. The results of the experiments demonstrate that this approach can successfully reduce and balance the energy consumption of network nodes to increase the life of the network.

Keywords: Wireless Sensor Network, LEACH, LEACH-CIE, MFRCRE, Optimise Node, Multi-Hop Routing, Relative Entropy, Probability Distribution.

SAMRIDDHI : A Journal of Physical Sciences, Engineering and Technology (2022); DOI: 10.18090/samriddhi.v14i04.02

Introduction

The wireless sensor network (WSN) has limited node energy and redundant data traits when the cost and application value are considered.[1] The energy consumption of each network node must be as low as possible in order to extend the WSN working cycle. The energy consumption rates of each node must be balanced in order for WSN to provide comprehensive and efficient monitoring of the targeted region. The traditional low-energy adaptive clustering hierarchy (LEACH) was cited in Literature[2] study from 2017 in that year. They clarified how the clustering concept was initially represented in the protocol. The energy consumption of network nodes is significantly reduced via network-layer data transfer. However, as the application system is upgraded, the protocol can no longer satisfy requirements, leading to the emergence of numerous new, enhanced LEACH protocols as needed.

Literature[3] added a cost function in the cluster head election strategy to maintain the energy balance of the nodes, and those closer to the sink and with higher residual energy were selected first. Still, the paper did not consider the total energy in the forwarding process. The energy consumption of multi-hop transmission may be higher than that of a single-hop direct message, resulting in higher overall network energy consumption. Literature[4] designed an efficient cluster head election strategy based on the particle swarm optimization algorithm. Although the network energy consumption is reduced compared to the random election cluster head protocol, this paper only optimizes the energy consumption within the cluster. The problem is that the energy consumption between sets and routing levels has not been discussed too much and needs to be further optimised. Literature[5] proposed a high-efficiency bee clustering WSN protocol based on the improved artificial bee colony meta-

Corresponding Author: Akant K. Raghuwanshi, Department of Electronic and Communication Engineering, RKDF University, Bhopal, Madhya Pradesh, India, e-mail: akantthakur7@gmail.com

How to cite this article: Raghuwanshi, A.K., Chaudhary, V.S. (2022). Relative Entropy-based Energy Efficient Multi-hop Forwarding Routing Protocol for Wireless Sensor Network. *SAMRIDDHI : A Journal of Physical Sciences, Engineering and Technology,* 14(4), 6-16.

Source of support: Nil

Conflict of interest: None

heuristic algorithm, which can obtain the optimal cluster head in WSN. The method is used to balance the network energy consumption, but there is no other design of the node energy consumption balance strategy at the routing level. Literature[6] believed that intermittent dormancy of nodes could save energy well, so they designed a node-adaptive sleep-wake scheduling algorithm. Still, the paper only reduced the energy consumption of a single node to reduce the overall energy consumption of the network. There is not much consideration from the perspective of balancing network energy consumption. Literature[7] used the bidirectional selection of standard and cluster head nodes to optimise cluster head election. They proposed a multi-hop data distribution method that comprehensively considered data transmission volume and transmission distance in the data transmission stage. Although this method can reduce network energy consumption to a certain extent, this method does not consider the remaining energy of nodes, which is not conducive to the balance of energy consumption of nodes. Literature[8] used the K-means algorithm and data regression strategy to achieve reasonable clustering for the network with uneven distribution of nodes. The paper achieves energy consumption balance between clusters by uniform clustering for the energy consumption at the routing level. Literature[9] uses the cuckoo algorithm to find the optimal path but ignores clustering and data redundancy problems. Based on the LEACH protocol, Literature[10] set the node clustering factor as the basis for the node to be elected as the cluster head and then designed a variable data transmission period within the cluster to achieve the energy balance of the network nodes. Still, this strategy's scope of use is limited, and it is challenging to implement the network with a fixed data collection period. Literature[11] uses data compression sensing to reduce the amount of data transmission as much as possible to save energy while accurately acquiring and analyzing data. The paper reduces the amount of data transmission through data fusion but does not discuss data transmission—energy consumption balance. Literature[12] first gridded the monitoring area, used the double-chain genetic algorithm to design a reasonable path, and used the convenient and movable sink node to follow the path to collect the entire data, reducing the communication distance of other nodes and saving the network. Energy consumption, but this strategy is challenging to implement for a fixed sink node network.

Because the energy consumption of WSN nodes is mainly reflected in data transmission, most of the traditional low-power WSN algorithm research and improved protocols based on LEACH are carried out from the transmission path, cluster head election and clustering strategy, etc., while ignoring the WSN data is redundant according to the characteristics of data redundancy in WSN.

Improved LEACH Protocol

LEACH Protocol

The core idea of the LEACH protocol is that the cluster head is elected regularly, and the remaining nodes select the cluster to which they belong based on the principle of the smallest distance. The cluster head collects the data of all cluster nodes and then forwards it to the sink node. The cluster head is randomly elected and held in turn. Each node in the network needs to generate a random number in the (0,1) interval and compare the random number with the threshold T (n).

$$T(n) = \begin{cases} \frac{p}{1-p\times\left(r mod \frac{1}{p}\right)}, n \in G_r \\ 0, n \notin G_r \end{cases} \quad (1)$$

Among them, p represents the percentage of cluster heads in the network, n represents the number of rounds that have been successfully completed, n represents the node number, and n Gr denotes that node n is a member of the set Gr. The number of games returned in the most recent r mod(1/p) round is known as Gr. If the random number produced by the nodes in Gr is smaller than the threshold T (n), the set of nodes that have not already been elected as cluster heads within a round will be chosen.

The following are flaws in the LEACH protocol: (1) The LEACH protocol uses single-hop direct routing, which is detrimental to the energy balance of the nodes in the overall network since long-distance communication takes more energy. (2) The energy-saving effect of the LEACH protocol needs to be further improved because it does not perform duplicate operations on the data of the wireless sensor network.

The modified centralized low-energy adaptive clustering hierarchy protocol, or LEACH-CM (modified centralized low-energy adaptive clustering hierarchy protocol), is an advancement over LEACH in that it not only optimizes the cluster head election strategy but also ensures that the cluster head is more significant from the residual energy than the network average as a whole. The cluster head nodes (CHs) are where the point is formed and where the network's total energy is stored. In addition, LEACHCM dynamically alters the data transmission path of select nodes to achieve a more Excellent energy-saving effect. The number of live nodes is alive multiplied by the cluster head ratio p. To confirm the efficacy of this approach, the experimental component will pass a comparison with LEACH-CM.

Energy Consumption Model

According to the literature,[14] the node data transmission model is as follows. Sensor sending and transmission (l the unit is a bit) the energy consumed by the data:

$$E_{Tx}(l,d) = \begin{cases} lE_{elec} + l\varepsilon_{fs}d^2, d < d_0 \\ lE_{elec} + l\varepsilon_{fs}d^4, d \geq d_0 \end{cases} \quad (2)$$

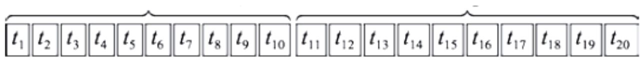

Figure 1: 10 sampling cycles are one round Data Transmission cycle.

Similar to long-distance freight commodities, loading and unloading fees are also required in addition to transportation fees. Equation (2) consists of two parts: the energy consumption of the transmission circuit and the transmission energy consumption. Elec represents the energy consumption of the unit bit data sent by the sensing circuit, and its size is related to encoding, modulation, filtering, etc. During the transmission process, according to the transmission distance d and the threshold $d_0\left(d_0=\sqrt{\frac{\varepsilon_{fs}}{\varepsilon_{amp}}}\right)$ relationship, subdivide the model into free-space models (transmission energy consumption vs distance is proportional to the square of). In the multipath attenuation model (transmission energy consumption versus distance is proportional to the fourth power), the energy consumption parameters under the two models are ε_{fs} and ε_{amp}. Since d is greater than 0, the energy consumption under both models varies with the distance increases and monotonically increases, compared to the free-space model, the multi-channel The energy consumption under the radial attenuation model is more, so for better energy saving, reduce the number of the transmission distance is particularly important. The number of l (unit is a bit) received by the sensor According to the energy consumed:

$$E_{Rx} = lE_{elec} \tag{3}$$

Similar to the energy consumption of the signal transmission circuit, the node is receiving the energy consumed by the data is only related to the size of the data packet, so the formula (3) The amount consumed by the sensor node to receive l (bit) data The energy is lE_{elec}.

Relative Entropy

Since the sensor collects the data value at a certain time, the relative entropy calculation method of discrete variables is used when calculating the relative entropy of periodic data. Suppose there are two probability distribution functions, respectively, f(x)= {f (1), f (2), ⋯} and g(y)= {g (1), g (2), ⋯},and

$$\begin{cases} \sum\limits_{i=1} f(i) = 1, f(i) \in [0,1] \\ \sum\limits_{j=1} g(j) = 1, g(j) \in [0,1] \end{cases}$$

Then the relative entropy of the two probability distributions is:

$$KL(f||g) = \sum f(i) lb \frac{f(i)}{g(i)} \tag{4}$$

The value of the relative entropy KL(f||g) can be used to describe the difference between the two probability distribution functions f(x) and g(y). In general, the greater the difference between the probability distributions, the farther the ratio of f (i) to g(i) is from 1, and the farther the value of KL(f||g) is from 0. In the extreme case, when the two probabilities are When the distributions are exactly the same, KL(f||g) = 0.

De-redundancy Method

The data set used in this paper is the environmental data collected by Intel Berkeley Research Lab in a real environment using multiple heterogeneous sensors for a continuous month, including data information such as date, time, sensor number, temperature, humidity, etc. This paper uses temperature value data t_i., the node completes the collection of environmental data every 30 s, and sends data once every 10 data collection processes. Each data sending period is represented by Ti, as shown in Figure 1.

The calculation method and de-redundancy method of node period data probability distribution are as follows:

- First, the current data transmission cycle *Ti* of the sensor node and the data column vectors X and Y of the previous cycle T_{i-1} are reconstituted into a new 10 × 2-dimensional vector x = [X, Y].
- Traverse the maximum value max x and the minimum value min x of the elements in the vector x, find the data span interval as (max x, min x), and divide the interval into duan segments (duan = 10 in this article), as follows shown in which:

min x . . . . max x

Wid =(max x - min x)/duan, the number of times the data of T_i and cycles appear in each segment are:
N1 = [n10,n11,n12,n13,n14,n15,n16,n17,n18,n19]´
N2 = [n20,n21,n22,n23,n24,n25,n26,n27,n28,n29]´

Then the probability distributions of T_i period and T_{i-1} period data are:

$$P = \frac{N1}{10}, Q = \frac{N2}{10}$$

- Calculate the relative entropy value KL(P||Q) of the data distribution between the current cycle T_i and the previous cycle of the sensor node according to formula (5). The smaller the relative entropy value, the greater the probability distribution of the data in the two cycles. near

$$KL(P||Q) = \sum_{j=1}^{duan} P_j \times lb \frac{P_j}{Q_j} \tag{5}$$

According to the practical application design of this article, the following will briefly analyze whether Q_j is 0:

* ∃j ∈ {1, 2, 3, 4, 5, 6, 7, 8, 9, 10} such that $Q_j = 0$, then $\frac{P_j}{Q_i} = \begin{cases} 0, P_j = 0 \\ Inf, P_i \neq 0 \end{cases}$ When = Inf, KL(P||Q) = Inf, which means that there is no data in the previous cycle, but there is in the current cycle, so there is a big difference in the data of the two cycles.

* ∀j ∈ {1, 2, 3, 4, 5, 6, 7, 8, 9, 10}, all have $Q_j \neq 0$, then this $Q_i \in$ {110,120,130,140,150,160,170,180,190,1} and $\sum_{j=1}^{10} Q_j = 1$, Hence Q_j = 1/10 for all j ∈ {1,2,3,4,5,6,7,8,9,10}. Similarly, P_j ∈{0,110,120,130,140,150,160,170,180,190,1}, and Q_j = 1/10, then this $\frac{P_j}{Q_j}$ ∈ {0,1,2,3,4,5,6,7,8,9,10}. Therefore, it can be concluded that $lb\frac{P_j}{Q_j}$ is non-negative, and the value of KL(P||Q) increases as the ratio of P_j to Q_j increases. That

is, the greater the difference between the two periods of data, the greater the value of KL(P||Q).

From this conclusion, the relative entropy value KL(P||Q) of the data probability distribution P of the current period T_i of the node and the data probability distribution Q of the previous period T_{i-1} is closer to 0, indicating that the correlation between the two periods of data is higher. The stronger it is, the higher the data redundancy in two cycles is, so it is necessary to set a reasonable threshold Do as the standard for measuring and judging data redundancy, and only when KL(P||Q) ≥Do will node data be sent out.

Multi-hop Forwarding Routing Improvements

Communication is the primary energy consumption path of nodes in WSN,[15] so a proper routing protocol needs to be designed to save energy.[16] Moreover, the communication distance of nodes is limited; to balance the communication energy consumption of long-distance nodes, it is necessary to design multi-hop forwarding routes further.

Geographical adaptive fidelity (GAF) is an energy-efficient routing algorithm in WSN.[17] It utilises the dense distribution of network nodes and reduces the number of redundant nodes in the network by setting redundant nodes to sleep. Network energy consumption, but the algorithm has a significant packet loss rate for the mobile node network. Literature[18] proposed a Cluster Routing Algorithm for Virtual Grid based on a virtual grid, CRVB); the algorithm first divides the network into virtual in each area, and the cluster head is selected according to the principle of maximum residual energy. The communication between sinks adopts multi-hop forwarding to reduce network energy consumption while balancing the network load. Literature[19] proposed a hybrid multi-hop routing algorithm for cellular meshes (improved hybrid clustering routing algorithm, IHCRA), which monitors. The area is divided into several regular hexagons, which are set based on factors such as angle and distance. Set the cost function to find the centre with the most minor cost from the neighbour nodes. The successor node acts as its next hop node, extending the network life, and has better results than traditional LEACH, GAF and CRVB protocols. A multi-hop forwarding route that considers its residual energy proposed in this paper (multi-hop forwarding routing considering its residual energy, MFRCRE) is divided into two parts: intra-cluster multi-hop forwarding and cluster Maximum step size transfer. Intra-cluster transmission adopts the idea of splitting into zeros. Turn high-energy long-distance single-hop transmission into low-energy multi-hop information. It reduces the network's overall energy consumption and balances each node rate's energy consumption. The inter-cluster message is based on the principle of high efficiency, and the maximum step size is used. Transmission until the data is transmitted to the sink node at a single hop distance (sensor within the farthest transmission distance of the node). This paper will be used in conjunction with the IHCRA method in the comparative simulation experiments, verifying the MFRCRE method's effectiveness.

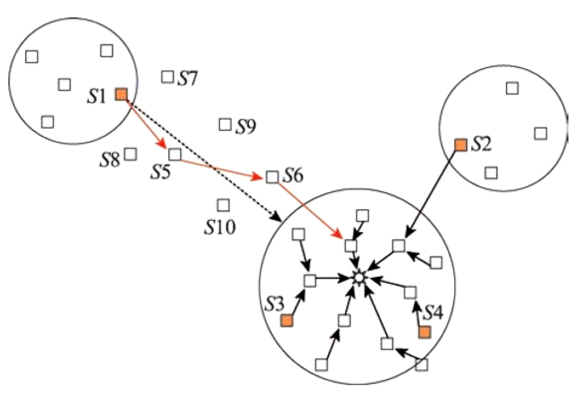

Figure 2: Schematic diagram of maximum step size transmission

Intra-cluster Multi-hop Forwarding

Usually, the cluster head and each node in the cluster is transmitted in a single hop within a distance, or although the source and destination nodes do not belong to the same clusters, but their distances are within a single hop distance, as shown in Figure 2.

Show: The circle in the Figure is a certain cluster in the area, and the S1 node is the distance within the cluster. The node with the farthest cluster head, the energy consumption of this node is larger, if it raises front failure, which may cause the upper left corner of the square monitoring area Loss of effective monitoring, to prevent this from happening, it is necessary to the direct communication between S1 and the cluster head is transformed into indirect communication, that is, finding a suitable the relay node S2 helps to forward data.

Assumption $\begin{cases} E_{fs} = lE_{elec} + l\varepsilon_{fs}d^2, d < d_0 \\ E_{amp} = lE_{elec} + l\varepsilon_{amp}d^4, d \geq d_0 \end{cases}$, because $d_0 = \sqrt{\frac{E_{fs}}{E_{amp}}}$, So $E_{fs} = E_{amp}$= when d = d_0 since both E_{fs} and E_{amp} are monotonic increasing function and E_{fs} is proportional to d^2, and E_{amp} is proportional to d^4 ratio, so for better energy saving, it needs to be larger than the long-distance under d_0 the single-hop straight hair is transformed into a multi-hop jump at a shorter distance, and at the same time, to balance the energy consumption rate between nodes, this paper establishes the energy consumption ratio ($\frac{Es_ic}{Es_i}$: turn the energy consumption of the development process is divided by the remaining energy), and the lower energy consumption is satisfied. On the premise, it is necessary to meet a lower energy consumption ratio to avoid excessive energy consumption and premature failure.

- Distance formula (guarantee lower energy consumption) Distance formula: when the source node S1 and the relay node S2 and the target nodes together form an obtuse triangle, and the connection between S1 and the target node constitutes an obtuse triangle. When forming the longest side of an obtuse triangle, there must be $d_1^2 + d_2^2 < d^2$, , according to the energy consumption model of the sensor node shows that the

Table 1: Experimental parameter settings

Parameter	*Experimental settings*
Square network monitoring area, area/m2	100×100
Sink node coordinates/m	(150,51)
Initial energy/J	0.11
/(J/bit)	50×
/(J/(bit m2))	10×
/(J/(bit m4))	0.0013×
Energy consumption of data fusion EDA/(J/bit)	5×
Sensor packet size/bit	4000
cluster head ratio p	0.21
The farthest transmission distance of a sensor node in one hop is dr/m	91

Table 2: 30 groups rERR and rAVG values under LEACH-CIE method when Do = 0.05

Index	*Value*										*Mean*
	0.264	0.27	0.275	0.271	0.269	0.262	0.271	0.268	0.262	0.269	
Average weight	0.264	0.271	0.263	0.266	0.273	0.267	0.26	0.261	0.264	0.273	0.267
	0.268	0.271	0.271	0.26	0.27	0.261	0.263	0.265	0.267	0.269	
	0.273	0.279	0.276	0.277	0.283	0.284	0.278	0.273	0.281	0.283	
rERR	0.287	0.273	0.279	0.284	0.28	0.286	0.282	0.279	0.279	0.278	0.28
	0.281	0.279	0.28	0.283	0.281	0.282	0.284	0.287	0.279	0.283	

energy consumption of the sensor node is related to the distance is proportional to the square of the space, according to which "long-distance, high-power single-hop direct"transmission" is converted to "close-range, low-power multi-hop transmission". Therefore d, d1, d2 must satisfy the following distance formula:

$$\begin{cases} d < d_0 \begin{cases} d^2 - d_1^2 - d_2^2 \geq 2\frac{E_{elec}}{\varepsilon_{fs}}, d_1 < d_0, d_2 < d_0 \\ \quad Single\ jump\ straight\ hair, other \end{cases} \\ d \geq d_0 \begin{cases} \varepsilon_{amp} d^4 - \varepsilon_{fs}(d_1^2 + d_2^2) \geq 2E_{elec}, d_1 < d_0, d_2 < d_0 \\ \varepsilon_{amp}(d^4 - d_2^4) - \varepsilon_{fs} d_1^2 \geq 2E_{elec}, d_1 < d_0, d_2 \geq d_0 \\ \varepsilon_{amp}(d^4 - d_1^4) - \varepsilon_{fs} d_2^2 \geq 2E_{elec}, d_1 \geq d_0, d_2 < d_0 \\ \varepsilon_{amp}(d^4 - d_1^4 - d_2^4) - \varepsilon_{fs} d_1^2 \geq 2\frac{E_{elec}}{\varepsilon_{fs}}, d_1 \geq d_0, d_2 \geq d_0 \end{cases} \end{cases} \tag{6}$$

Among them, when d $<d_0$, if $d_1 \geq d_0 >$ d or $d_2 \geq d_0 >$ d, that means the distance forwarded by the relay node is longer than that of the single-hop direct transmission. Longer and consumes more energy. At this time, the source node S1 will directly communicate with the target node point-to-point without forwarding by the relay node S2.

- Condition of energy consumption ratio (guarantee energy consumption balance) The way to control energy balance should not be to control the energy consumption of nodes. The value is the same, and the node energy consumption ratio should be held to be as the same as possible. Energy consumption is the ratio of the node's current communication energy consumption to the current remaining energy. The energy consumption ratio condition transforms a single-hop direct transmission with a high energy consumption ratio into a low energy consumption ratio of multi-hop forwarding.

$$\frac{Esc}{Es} \geq \frac{Es_i c}{Es_i} \begin{cases} d < d_0, d_2 < d_0 \frac{lE_{elec} + l\varepsilon_{fs} d^2}{Es} \geq \frac{2lE_{elec} + l\varepsilon_{fs} d_2^2}{Es_i} \\ d \geq d_0 \begin{cases} \frac{lE_{elec} + l\varepsilon_{fs} d^4}{Es} \geq \frac{2lE_{elec} + l\varepsilon_{fs} d_2^2}{Es_i}, d_2 < d_0 \\ \frac{lE_{elec} + l\varepsilon_{fs} d^4}{Es} \geq \frac{2lE_{elec} + l\varepsilon_{amp} d_2^2}{Es_i}, d_2 \geq d_0 \end{cases} \end{cases} \tag{7}$$

Among them, Esc is the energy consumed by the source node in the single-hop direct transmission mode, Es is the remaining energy before the source node data is sent, and *$Es_i c$* is the specific candidate relay node s_i of the distance formula forwarding the data of the source node The energy value consumed by the relay node *Es_i*, indicates that the candidate relay node forwards the remaining energy before only the nodes that satisfy equations (6) and (7) can be selected as a relay node to perform forwarding tasks.

Maximum Step Size Transmission Between Clusters

In the random distribution of nodes, the nodes closer to the sink are the point that plays the role of a bridge, not only transmitting their data but also responsible for transferring Sending data from distant nodes so that the energy consumption will be relatively higher. Therefore, the multi-hop forward routing (MFRCRE) method considers its residual energy when the distance between the source node and the target node is greater than the sensor's single

Table 3: Performance of the three methods under the same environment and dataset

Method	*The first cycle in which a failed node occurs*										*Average*
	1	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	
LEACH	21	13	22	24	25	22	22	17	19	24	20.9
LEACH-CM	30	33	27	32	36	33	32	29	25	34	31.1
LEACH-CIE	35	28	26	34	40	29	33	39	36	38	33.8

hop. In the transmission distance (dr), to better balance the energy consumption and at the same time reduce the Low propagation delay (the more transmission times, the higher the packet loss rate, and even the energy consumption will be higher [20]), setting such nodes to transmit with the maximum stride to Sink is the centre, dr is the radius of the circular area, as shown in Figure 3. The maximum step size transfer steps are as follows:

- The cluster head whose distance from Sink is greater than dr will search for the node closest to Sink as its next-hop node within its largest communication radius to realize the transmission with the most significant step size;
- If the next hop node in step (1) is still outside the centre circle area, repeat step (1) until the data reaches the centre circle area;
- The data is transmitted to the sink using the intra-cluster multi-hop forwarding rule in the central circle area.

Experimental Test And Analysis

To verify the method's effectiveness in this paper, the following simulation experiments are carried out in MATLAB R2014b. The monitoring location is set to be a square island area with an S = 100 m × 100 m; the sensor node has no external power supply, and the sink node is located outside the square area. The coordinates are (150 and 50), and the remaining 50 ordinary sensor nodes are randomly distributed in the square area. The data adopts Intel Berkeley Lab. (http://db.lcs.mit.edu/labdata/labdata.html) Deploy the natural environment dataset collected by wireless sensor network nodes to monitor environmental changes in the area, predict fires and test the network life cycle. The following assumptions are made for the wireless sensor network set up in this paper.[21]

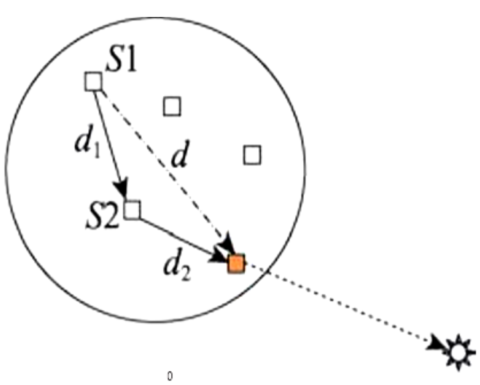

Figure 3: Multiple jumps forwarding within cluster

- Each sensor node has the capabilities of environmental perception, data collection, data fusion processing, computing, wireless communication, etc. The environment will not cause damage to the sensor node and will not affect the regular operation of the node.
- Once the position of the sensor node in the network is determined, it will not change, and each node knows its position coordinates; the battery power of each node and the initial energy is not zero but limited.
- The energy consumption of sensor nodes mainly occurs in the transmission, reception, transmission and fusion of data, and the energy consumed by sensor environment monitoring and calculation is ignored. The experimental parameters are set as shown in Table 1.

The calculation model of the experimental performance comparison index:

- The number of failed nodes in the entire network

$$rDEAD = \sum_{r=1}^{r_{max}} dead(r) \tag{8}$$

Among them, dead(r) is the number of newly added failed nodes in each round period, and the sum of the number of failed nodes in all rounds is superimposed to obtain the total number of failed nodes in the network. The trend of the comparison index over time can be used to characterize the speed of the network inactivation. When the network's overall number of failed nodes is equal to the total number of nodes, the network is completely deactivated.

- Average residual energy of the network

$$rAVG = avg(200) = \frac{1}{n}\left(\sum_{i=1}^{n} E_{199i} - e_{200i}\right) \tag{9}$$

Among them, e_{ri} is the energy consumption value of node Si in the rth round period, E_{ri} is the remaining energy value of the node Si after completing the rth round of the data transmission period, and avg(r) is the net energy at the end of the rth round period. The network average residual energy comparison index can be used to compare the energy-saving effect of different methods. The faster the index parameter decreases with time, the higher the network energy consumption.

- Data error

$$err(i) = \frac{1}{10}\sqrt{\sum_{k=0}^{9}(x_k - y_k)^2} \tag{10}$$

$$rERR = \frac{\sum_{r=1}^{r_{max}}\left(\sum_{i=1}^{n} err(i)\right)}{n \times r_{max}} = \frac{\sum_{r=1}^{r_{max}}\left(\sum_{i=1}^{n}\sqrt{\sum_{k=0}^{9}(x_k-y_k)^2}\right)}{n \times r_{max}} \tag{11}$$

Where err(i) refers to the data error value of node i in each cycle, because this paper assumes that the node performs 10 data collection processes in each cycle, the denominator under this calculation model is 10, rERR is the data value obtained by the sink node after the network runs for cycles and the actual value of the n sensor nodes. The average error of the collected data. Experiments must be carried out within the allowable error range to be effective.

- The mean square error of the residual energy of the network

$$D_e = \sqrt{\frac{\sum_{i=1}^{n_{alive}}\left(E_i - \frac{\sum_{j=1}^{n_{alive}} E_j}{n_{alive}}\right)^2}{n_{alive}}} \quad (12)$$

Among them, n_{alive} represents the number of surviving nodes in the network in the current cycle, and E_i represents the remaining energy of node Si in the current process; the smaller the mean square deviation of the remaining power of the network, the stronger the energy balance capability of the network.

Relative Entropy

Since the data value of the sensor node is non-negative, the relative entropy value is non-negative, and the closer KL(P||Q) is to 0, the higher the data redundancy. In terms of the design of the relative entropy, threshold Do, due to the extensive Do design, too much data is judged to be redundant and discarded, resulting in too large errors in the data obtained by the sink node. According to preliminary experiments, Do<0.5 is limited. The influence of Do changes on the error, set Do as the step size of 0.01, increase from 0.01 to 0.5, and calculate the network under the LEACH-CIE (improved LEACH protocol combining relative information entropy) method. After running for 200 cycles under different Do conditions, the nodeaverage residual energy (rAVG) and the average error of the network rERR) change. The graph of the relationship shown in Figure 4 is obtained as shown in Figure 4, as the relative entropy threshold, Do, increases continuously.

The average residual energy (rAVG) of the network is growing, and the average error (rERR) is also growing. Therefore, to better solve the energy consumption, it is necessary to make Do as large as possible; however, to reduce errors, it needs to be as small as possible. In the practical application assumed in this paper, to monitor and warn the environment better and longer, it is necessary to set a reasonable error range and choose the optimal energy-saving strategy. Combined with the literature Xu Chi *et al.*,[23] the system test accuracy is ± 0.3°C, and the error of the alarm temperature measurement value of the train on-board equipment set by Yin Keqiang et al.[24] does not exceed 0.4°C, and the PCDA (precision configurable the average absolute error of the mean query of the data aggregation algorithm) method is 0.136 6~0.303 0. The error is much smaller than that of the B-based method. Therefore, this paper sets the allowable data error of WSN to 0.300. As can be seen from Figure 4, taking Do = 0.05 can Satisfy the ideal error condition. To further verify that Do = 0.05 is the excellent set value, now keep Do unchanged and carry out 30 groups of experiments based on the LEACH-CIE method for 200 cycles of network operation.

Each experiment's positions of sensor nodes are randomly generated, and 30 groups are obtained. The values of the data average error rERR and the network average residual energy rAVG are shown in Table 2. The data in Table 2 shows that the errors of the 30 experiments are all within the ideal range. According to Figure 4 and Table 2, after determining Do = 0.05, in the average case, the network under the LEACH-CIE method still has an ideal energy saving after running for 200 cycles, and the average data error value of 30 experiments is 0.280 and less than the error threshold of 0.300. In a similar application, Literature[20] designed an abnormal data-driven data aggregation method (DA-ADD) in WSN based on anomalous data. The adopted strategy is to calculate the absolute value of the difference between the two data, the data more significant than the threshold value is judged as the shock data, and only the shock data is transmitted. In addition, when the cataclysm data is generated, an algorithm with a complexity of n2 is used to calculate the support degree between nodes to further determine the cataclysm data's validity. Next, the DA-ADD method is compared with the LEACH-CIE way (setting Do = 0.05) in terms of error and energy consumption. Figures 5 and 6 are the comparison experiments of 30 groups of average error and average residual energy between the DA-ADD and the LEACH-CIE methods, respectively.

It represents the average error of the network, and the ordinate in Figure 6 depicts the average residual energy of the network. The x-axis of Figure 7 illustrates the average error, the y-axis represents the number of cycles, and the z-axis

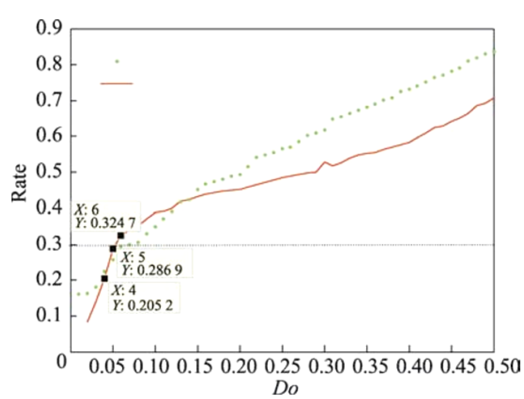

Figure 4: Graph of mean error and average remaining energy of LEACH-CIE varying with Do.

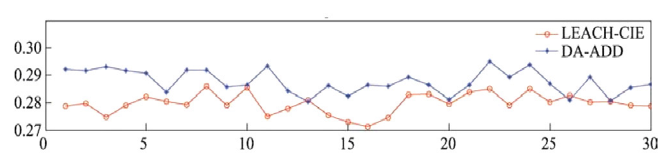

Figure 5: Comparison of average error between DA-ADD and LEACH-CIE

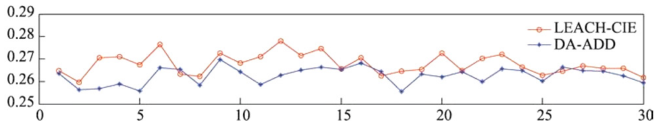

Figure 6: Comparison of average residual energy between DA-ADD and LEACH-CIE

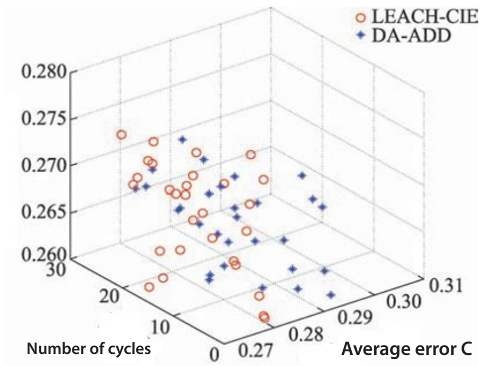

Figure 7: Comparative experiment of DA-ADD and LEACH-CIE

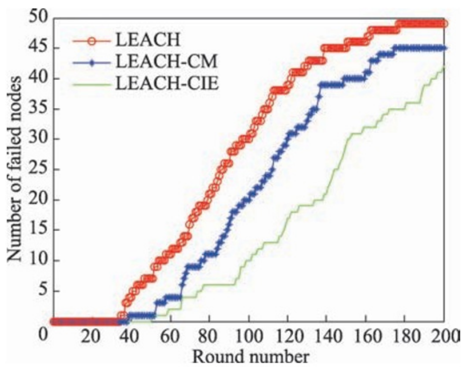

Figure 8: Comparison of the number of failed nodes

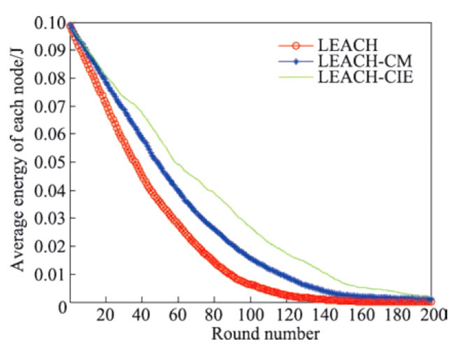

Figure 9: Curve of average value of remaining energy in network.

represents the average residual energy. Figure 7 compares the error and energy consumption between DA-ADD and LEACH-CIE from a three-dimensional perspective. It is not difficult to see from it. This is because average; the LEACH-CIE method has higher average residual energy and lower average error.

Since the DA-ADD method uses the comparison between individual data, the absolute value of the difference reflects the abnormality of the data. In contrast, the LEACH-CIE method in this paper uses the relative entropy value of the probability distribution of each 10 data in adjacent periods to reflect the abnormality of the data. Reflecting the exception of data, both methods achieve the purpose of energy saving through limited transmission (only abnormal data is transmitted). The DA-ADD method dramatically reduces the amount of data transmission in the process of de-redundancy but also significantly increases. The data error obtained by the sink node is reduced. In the 30 sets of simulation experiments, the average error value of the LEACH-CIE method is 0.280 4, the average residual energy value is 0.268 4, while the average error of the DA-ADD process is 0.286 5, the average residual energy value is 0.264 1, combined with Figure 5, Figure 6 and Figure 7, it is not difficult to see that the LEACH-CIE method has higher average residual energy and lower average error than the method in the literature,[24] so it is reasonable to set Do=0.05. Subsequent experiments will Continue this parameter setting value.

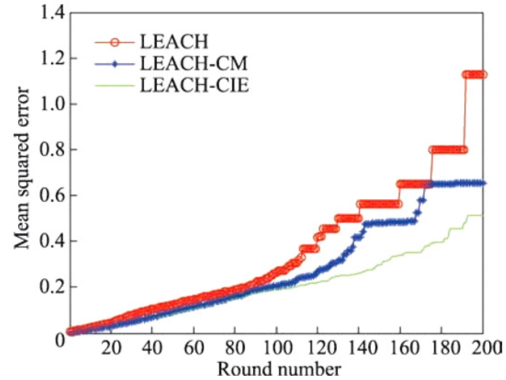

Figure 10: Comparison of mean squared errors of network residual energy

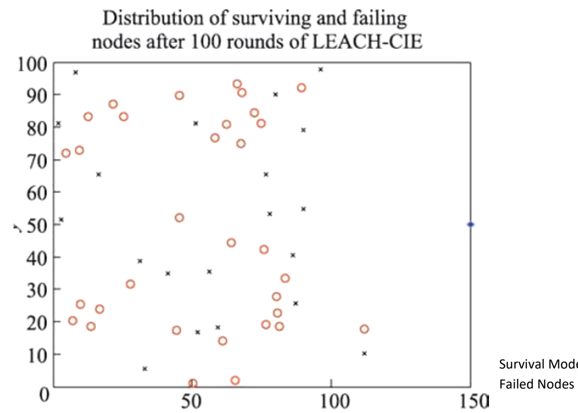

Figure 11: Distribution of two types of nodes after 100 rounds of network operation under LEACH-CIE protocol

Comparative Experiment

Comparison of different performance indicators

- Comparison of the number of failed nodes in the network after the end of each cycle. As shown in Figure 8: (1) With

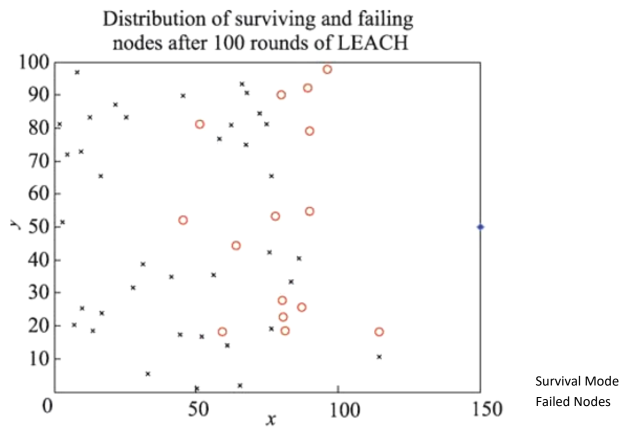

Figure 12 : Distribution of two types of nodes after 100 rounds of network operation under LEACH and LEACH-CM protocols

the continuous growth of the network operation cycle, the number of sensor node failures under the LEACH, LEACH-CM and LEACH-CIE methods are increasing. However, the LEACH-CIE method proposed in this paper is better than other methods. In the two methods, the number of failed nodes increases more slowly; (2) LEACH-CIE appears as the first failed node at the latest. Table 3 records the performance of the three methods under the same environment and dataset.Line 10 times, each time the first cycle of the failed node occurs. Table 3 Number of cycles failed nodes first appear in 3 methods

From the data in Table 3, we can calculate the average number of cycles of the first failure node in the network under the LEACH, LEACH-CM and LEACH-CIE methods, which are 19.9, 30.1 and 32.8, respectively. On average, the way in this paper has the latest failure node. Therefore, the number of cycles is in line with extending the network life cycle.

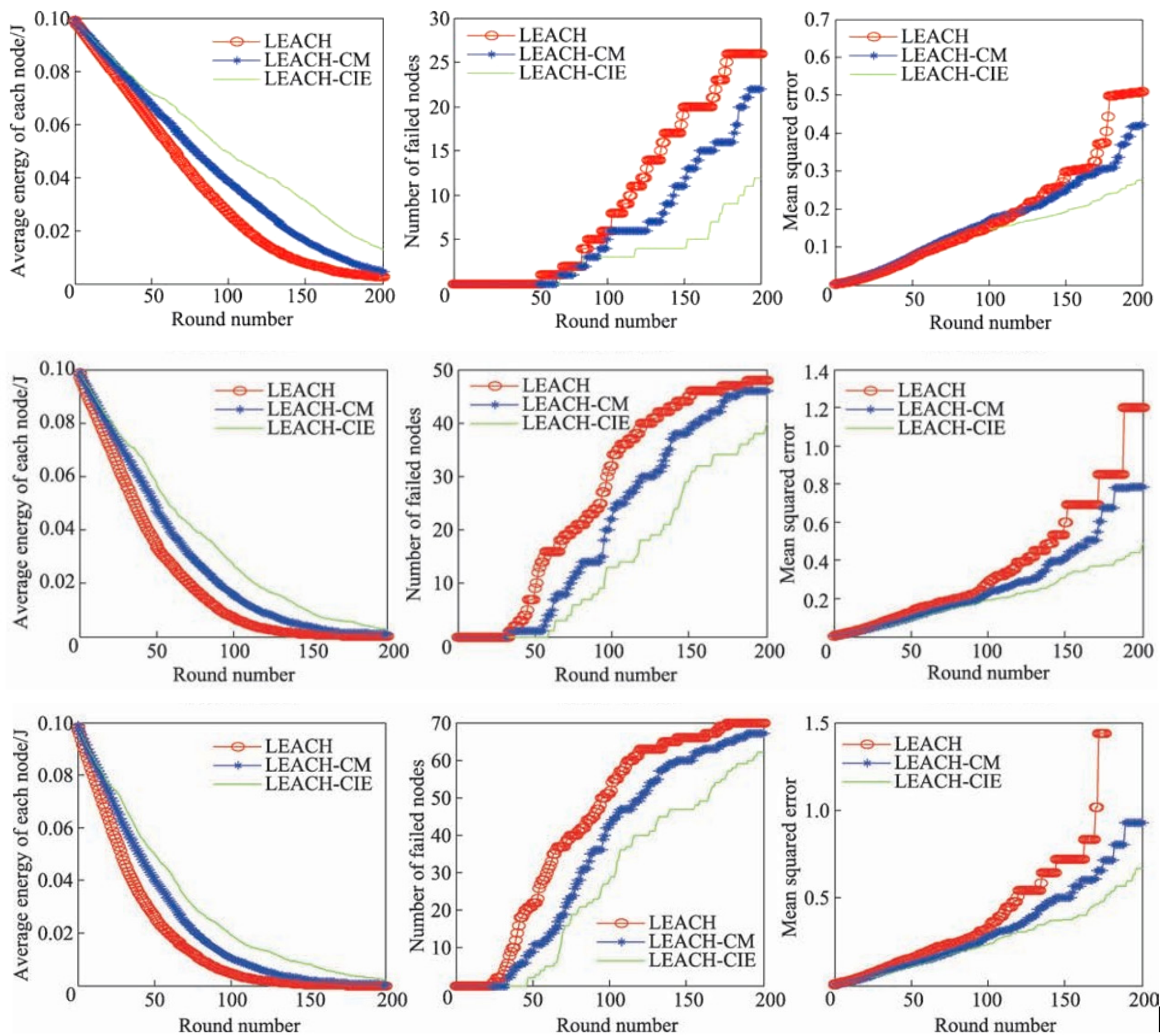

Fig.13: Comparison of 3 methods under different number of nodes

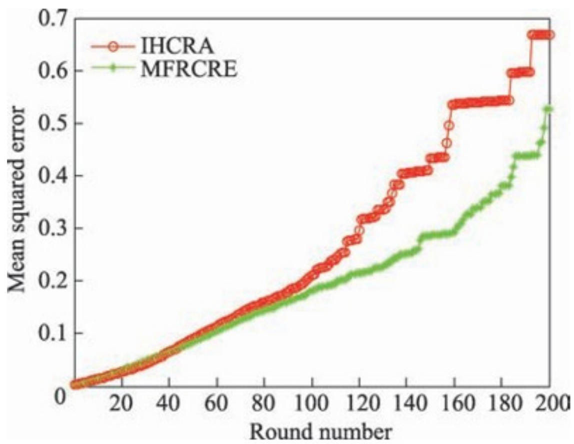

Fig.14: Multi-hop forwarding routing comparison

- After each round of data transmission, the average value and mean square error of the remaining energy of the network are compared.

It can be seen in Figure 9:

(1)Under the same conditions, LEACH-CIE has a slower energy consumption rate, so the energy saving effect is more significant;

(2)When the average residual energy of the network is as low as close to 0, it means that the network has been paralyzed. It can be seen from the Figure that LEACH-CIE has a longer network survival time. It can be seen from Figure 10 that in the same period, the method LEACH-CIE in this paper has a smaller mean square error value of residual energy in the network, so the energy consumption rate of nodes is more balanced.

- Distribution of "surviving" nodes in the network

The rectangles in Figures 11 and 12 are both 150×100 rectanglesshape area, the abscissa and ordinate in Figure 11 and Figure 12 represent the section.

The coordinate value of the point in the network, where the coordinates of the sink node are (150, 50). Figure 11 shows the distribution of surviving and failed nodes in the network after running for 100 cycles under the method LEACH-CIE in this paper. As seen from the Figure, compared with LEACH and LEACH-CM, LEACH-CIE has more nodes surviving in the same period, and the distribution is more uniform, so the energy consumption of network nodes under LEACH-CIE is lower and more equal.

In Figure 12, "." represents the node still alive, and "×" represents the failed node. It can be seen that after the networks under the LEACH protocol and the LEACH-CM protocol run for 100 cycles, the distribution of surviving nodes is uneven, and the nodes farther away from the sink will fail first.

The distance formula finds the candidate relay nodes in this paper. On the premise that the total energy consumption of forwarding is lower than the energy consumption of direct transmission of the source node S1, the network's total energy consumption is guaranteed to be lower. Then the energy consumption of the relay S2 is certified by the energy consumption ratio condition. Furthermore, the rate is not higher than S1 to avoid premature failure of relay nodes due to too fast energy consumption. Through these two points, the purpose of reducing the network's overall energy consumption and balancing each node's energy consumption rate is achieved.

Comparison of changing node density

To better verify the method's robustness in this paper, when the area S = 100 × 100 remains unchanged, the number of nodes is increased from 30 to 70 with a step size of 20 for comparative experiments, and Figure 13 is obtained.

The first column of Figure 13, there is a comparison of the average network remaining energy values, the second column is a comparison of the number of network failure nodes, and the third column is a comparison of residual energy means square error. As seen from the Figure, in the case of different nodes, the network under the LEACH-CIE method still has a longer life cycle than the network under LEACH and LEACH-CM, and the energy consumption of the network nodes is lower and more balanced.

Comparison of Multi-hop Forwarding Routes

Based on the mean square error of the residual energy in the network, the multi-hop forwarding routing method (MFRCRE), which takes into account the residual energy of the network, is compared with the IHCRA routing algorithm, and the results are shown in Figure 14.

Figure 14 compares the MFRCRE algorithm with the IHCRA algorithm under the same network arrangement, the same clustering structure, and the same network data. It can be seen from the Figure that the multi-hop routing algorithm that takes into account its residual energy has lower residual energy mean square error than IHCRA, which means that in the same period, the network node energy consumption under the multi-hop routing algorithm that takes into account its residual energy is higher balanced.

Conclusions

The method in this paper combines the information entropy theory. It uses the relative entropy model to calculate the relative entropy value of the data distribution of the adjacent two node cycles as the basis for judging the data redundancy. The operation of blocking transmission can achieve the purpose of de-redundancy and energy saving of WSN data. In the data transmission process, the node communication distance and energy consumption ratio factors are comprehensively considered, and the node forwarding conditions are set: distance condition and energy consumption ratio condition. Only when the two needs are met simultaneously will the node perform the forwarding task, which y balances the energy consumption of each node in the network and avoids the rapid failure of nodes in some areas. By changing the distribution density of nodes and conducting multiple sets of simulation experiments, compared with the LEACH and LEACH-CM protocols, the network under the LEACH-CIE method in this paper has a longer life cycle, and the node energy consumption is more balanced.

Reference

[1] X. Cui, X. Zhang and Y. Shang, "Energy-saving Strategies of Wireless Sensor Networks," 2007 International Symposium on Microwave, Antenna, Propagation and EMC Technologies for Wireless Communications, 2007, pp. 178-181, doi: 10.1109/MAPE.2007.4393575.

[2] D. Ma et al., "FAT-WSN: A Non-Destructive and Secure Aggregation Strategy for Energy Saving in WSN," 2019 IEEE Globecom Workshops (GC Wkshps), 2019, pp. 1-6, doi: 10.1109/GCWkshps45667.2019.9024492.

[3] A. Dekdouk, "Energy efficient oriented mobile learning that emphasizes energy saving awareness education," 2013 ACS International Conference on Computer Systems and Applications (AICCSA), 2013, pp. 1-4, doi: 10.1109/AICCSA.2013.6616496.

[4] X. Zhong, Y. Tan, Q. Ding and L. Zhang, "Integrative management and energy-saving strategy in WSN," 2014 World Congress on Computer Applications and Information Systems (WCCAIS), 2014, pp. 1-7, doi: 10.1109/WCCAIS.2014.6916580.

[5] A. Mondal and S. Bhattacharya, "Energy Efficient and Effective Control Strategy of HVAC System in Large Shopping Complex," 2014 3rd International Conference on Eco-friendly Computing and Communication Systems, 2014, pp. 116-120, doi: 10.1109/Eco-friendly.2014.99.

[6] Z. Zhang, R. Luo and W. Fu, "Energy Saving Algorithm of Wireless Network Nodes in Cluster," 2020 International Wireless Communications and Mobile Computing (IWCMC), 2020, pp. 394-397, doi: 10.1109/IWCMC48107.2020.9148250.

[7] S. H. Singh, K. R. Verma and P. Rajpoot, "Partition Based Strategic Node Placement and Efficient Communication Method for WSN," 2018 3rd IEEE International Conference on Recent Trends in Electronics, Information & Communication Technology (RTEICT), 2018, pp. 1807-1812, doi: 10.1109/RTEICT42901.2018.9012336.

[8] N. A. M. Alduais, J. Abdullah and A. Jamil, "RDCM: An Efficient Real-Time Data Collection Model for IoT/WSN Edge With Multivariate Sensors," in IEEE Access, vol. 7, pp. 89063-89082, 2019, doi: 10.1109/ACCESS.2019.2926209.

[9] C. Jing, T. Gu and L. Chang, "ESCAL: An Energy-Saving Clustering Algorithm Based on LEACH," 2008 IEEE International Symposium on Knowledge Acquisition and Modeling Workshop, 2008, pp. 403-406, doi: 10.1109/KAMW.2008.4810509.

[10] Xiangyi Li, Guixia Kang, Xidong Zhang and Dongyan Huang, "An energy-efficient cooperative MIMO strategy for Wireless Sensor Networks with intra-body channel," 2012 International Symposium on Communications and Information Technologies (ISCIT), 2012, pp. 679-684, doi: 10.1109/ISCIT.2012.6380987.

[11] H. Bizagwira, J. Toussaint and M. Misson, "Multi-channel routing protocol for dynamic WSN," 2016 Wireless Days (WD), 2016, pp. 1-3, doi: 10.1109/WD.2016.7461512.

[12] W. Feng et al., "Joint Energy-Saving Scheduling and Secure Routing for Critical Event Reporting in Wireless Sensor Networks," in IEEE Access, vol. 8, pp. 53281-53292, 2020, doi: 10.1109/ACCESS.2020.2981115.

[13] A. M. I. Alkuhlani and S. B. Thorat, "Enhanced location privacy and energy saving technique for sensors in Internet of Things domain," 2016 International Conference on Global Trends in Signal Processing, Information Computing and Communication (ICGTSPICC), 2016, pp. 122-125, doi: 10.1109/ICGTSPICC.2016.7955282.

[14] H. Ben Fradj, R. Anane and R. Bouallegue, "Energy Consumption for Opportunistic Routing Algorithms in WSN," 2018 IEEE 32nd International Conference on Advanced Information Networking and Applications (AINA), 2018, pp. 259-265, doi: 10.1109/AINA.2018.00048.

[15] D. Pescaru, C. Istin, D. Curiac and A. Doboli, "Energy Saving Strategy for Video-based Wireless Sensor Networks Under Field Coverage Preservation," 2008 IEEE International Conference on Automation, Quality and Testing, Robotics, 2008, pp. 289-294, doi: 10.1109/AQTR.2008.4588754.

[16] M. L. Laouira, A. Abdelli, J. B. Othman and H. Kim, "An Efficient WSN Based Solution for Border Surveillance," in IEEE Transactions on Sustainable Computing, vol. 6, no. 1, pp. 54-65, 1 Jan.-March 2021, doi: 10.1109/TSUSC.2019.2904855.

[17] X. -S. Yi, P. -J. Jiang, X. -W. Wang and S. -C. Zhang, "Survey of Energy-Saving Protocols in Wireless Sensor Networks," 2011 First International Conference on Robot, Vision and Signal Processing, 2011, pp. 208-211, doi: 10.1109/RVSP.2011.43.

[18] S. Misra and R. Kumar, "A literature survey on various clustering approaches in wireless sensor network," 2016 2nd International Conference on Communication Control and Intelligent Systems (CCIS), 2016, pp. 18-22, doi: 10.1109/CCIntelS.2016.7878192.

[19] Jia Meng, Hao Wu, Hengliang Tang, Lan Dong and Hua Cui, "Research on energy efficiency in cooperative communication of WSNs," 2012 IEEE 14th International Conference on Communication Technology, 2012, pp. 485-490, doi: 10.1109/ICCT.2012.6511267.

[20] Z. Hu, Y. Sun and Y. Ji, "A Dynamic Spectrum Access Strategy Based on Real-Time Usability in Cognitive Radio Sensor Networks," 2011 Seventh International Conference on Mobile Ad-hoc and Sensor Networks, 2011, pp. 318-322, doi: 10.1109/MSN.2011.3.