EE-TLT: Energy-Efficient Routing Protocol Using Two-Level Tree-based Clustering in Wireless Sensor Network

Nguyen Duy Tan and Van-Hau Nguyen

Abstract—When designing routing protocols for wireless sensor networks, the principal challenge is to prolong the network's lifespan by effectively using the limited battery energy of the sensor nodes. To address this issue, we propose an energy-efficient routing protocol employing a two-level tree-based clustering (called EE-TLT) approach to stabilize and efficiently use the sensor node's energy. In EE-TLT, the regional network is logically divided into clusters, with the number of nodes balanced in each cluster. Within each cluster, the nodes are again separated into polygons and the data is transmitted only via short links using a two-level routing tree, which is composed of one or more minimum spanning trees based on the Kruskal algorithm with a sub-cluster head (sub-CH) node serving as the root and a two-level tree linking sub-CHs at different polygons and the base station (BS). To determine the cluster head or relay cluster head node in each polygon or sector respectively, EE-TLT considers the energy residual and distance among candidate nodes and the BS. Furthermore, EE-TLT selects the optimal data transmission stage length in each round, significantly increasing the number of data packets that the BS receives. Our experimental results demonstrate that EE-TLT not only further balances the energy consumption among sensors but also improves the ratio of data packets accepted by BS and energy efficiency compared to the LEACH-VA, PEGCP, and STDC by approximately 25%, 15%, and 10%, respectively, in both homogeneous and heterogeneous networks. The code and the simulation results of EE-TLT may be found at https://tinyurl.com/ee-tlt-wsn.

Index Terms—Data fusion, energy balancing, energy-efficient routing, heterogeneous wireless sensor networks, routing protocol, tree-based clustering.

I. INTRODUCTION

WIRELESS sensor networks (WSNs) involves numerous micro-sensor nodes that are placed in the natural environment to support various applications such as forest fire alarm systems [1], environmental monitoring [2], [3], health care systems [4], battlefield surveillance [5], smart homes, and intelligent transportation systems [6]. Due to their tiny size, cheap price, restricted processor abilities, memory, bandwidth, and resources, sensor nodes cannot replace or recharge their batteries in harsh areas where humans cannot access them [7]. Therefore, designing energy-efficient routing protocols for WSNs is crucial to stabilize energy consumption and increase the network's lifespan [8], [9]. The low-energy

adaptive clustering hierarchy (LEACH) protocol, proposed by Heinzelman *et al*. [10], is an example of a protocol that distributes nodes into clusters and designates a control node as the cluster head (CH) responsible for gathering data packets from non-CH nodes within the cluster. The CH then combines the data for enhanced security and reliable information and transmits it to the base station (BS) [11], [12]. The other nodes transmit their data packets to their respective CHs, allowing LEACH to reduce energy usage and prolong the network's lifespan. However, the CHs may deplete their batteries faster due to increased workload and communication over longer distances with the BS. As a result, the CH role should be randomly assigned to other nodes to achieve balanced energy dissipation in the network. Unlike LEACH, centralized LEACH (LEACH-C) [10], [13] uses the BS for CH selection and cluster formation, reducing control message overhead and enabling the identification of the maximum amount of clusters in each round using a block-based clustering approach. Furthermore, LEACH suffers from long-distance communication between nodes and their CHs and the BS, leading to faster battery depletion. LEACH-VA [14] improves upon LEACH by using Voronoi cells and ant colony optimization to optimize multi-hop communication in intra-clusters. Another approach, called power-efficient gathering in sensor information systems (PEGASIS) [15], constructs a long chain of nodes with a greedy algorithm and selects a node to be the CH for data fusion and transmission. However, PEGASIS suffers from the overload of the selected CH and long distances to the BS. The power efficiency grid-chain routing protocol (PEGCP) [16] aims to reduce energy consumption in WSNs by dividing devices of the network into virtual cells and using a chain communication scheme. PEGCP outperforms LEACH in terms of energy efficiency and network lifespan, but has some drawbacks. PEGCP does not ensure balanced energy consumption among smart sensor devices due to randomly deployed nodes and fixed cell division, and there may be long communications due to a single chain algorithm. Mittal *et al*. [17] suggested a tree-based clustering and threshold-sensitive routing approach that uses an enhanced flower pollination algorithm (EFPA) to connect sensor nodes into a routing tree with CH as the root. The experiment results show that tree-based routing protocols significantly reduce energy consumption compared to chain-based and block-based routing mechanism [18]–[20]. Overall, the energy efficiency of routing protocols is crucial for prolonging the lifespan of the network and balancing energy consumption in WSNs. To address this, we propose a two-

level tree-based clustering routing protocol called EE-TLT. Our contributions are as follows:

1) We distribute alive nodes into different sectors (clusters) based on their location in the monitoring zone to ensure an equal number of nodes in each cluster.

2) We select CH nodes in each cluster based on remaining energy and distance to the BS, determining which node becomes the sub-cluster head (sub-CH) or relay-cluster head (relay-CH).

3) We build a two-level tree for establishing data transmission paths from nodes to the BS in order to avoid long links by creating several minimum spanning trees (MSTs) for the first-level intra-polygon and the second-level inter-polygon communication.

4) We analyze the length of the data transmission stage to determine optimal durations for increasing the number of successfully delivered data packets.

5) Finally, we conduct extensive experiments and find that EE-TLT outperforms LEACH-VA, PEGCP, and STDC by approximately 25%, 15%, and 10%, respectively, in terms of energy efficiency in both homogeneous and heterogeneous networks.

To facilitate the organization and presentation of our findings, we have structured our work as follows: In Section II, we present an overview of related works. We then introduce the system model employed in our study in Section III and provide a detailed account of the EE-TLT in Section IV. The simulation results are thoroughly analyzed and evaluated in Section V. Finally, we conclude our study in Section VI, where we present our overall findings and conclusions.

II. RELATED WORKS

In hierarchical routing mechanisms, protocols can support both similar and different network models with multi-hop and single-hop communication modes for WSNs [21]–[23]. This section briefly outlines the previous research on improving energy efficiency in hierarchical routing protocols, which can be classified into three types: block-based, chain-based, and tree-based routing schemes [24].

A. Block-based Routing

Several protocols have been proposed to increase energy efficiency in WSNs through effective clustering routing protocols. These protocols include SEECP by Mittal *et al.* [25], which selects CHs based on the residual energy of sensing nodes and calculates the optimal radius around the BS to reduce transmission costs. Liu *et al.* [26] proposed an energy-efficient routing protocol based on a two-dimensional monitoring zone that is divided into cells, and a candidate node is selected in each cell to become the CH based on its energy level and the number of neighbors. These CHs collect and forward data to the BS. Liu *et al.* [27] also proposed an improved energy-efficient LEACH protocol (IEE-LEACH) that chooses CHs based on the network's average energy, the initiation energy of nodes, and the residual energy of alive nodes. Nodes nearer to the BS than CHs directly transmit

data to the BS. Tang *et al.* [28] proposed a reliable and energy-efficient routing protocol called DSEERA, which uses Dempster Shafer's evidence theory to fuse idleness degree, energy-dense factor, and transmission energy-efficient ratio to determine the optimal routing. However, the computation of the probability function in DS-based protocols has high complexity, which may not be practical for sensor nodes with limited computation capacity. Firdous *et al.* [29] proposed efficient clustering-based routing for energy management in the WSN-assisted Internet of things (PERC), in which the clustering algorithm according to K-means is used to distribute nodes into clusters. The selection CHs and main CH are based on the relative location and residual energy of sensing nodes. PECR decreases the energy used and extends the lifespan of the network. However, the single-hop communication intra-cluster does not guarantee the minimum energy consumption.

B. Chain-based Routing

Marhoon *et al.* [30] proposed the deterministic chain-based routing protocol (DCBRP), which is another clustering mechanism routing scheme commonly utilized to improve energy efficiency in WSNs. DCBRP includes three mechanisms: backbone formation, chain head node selection (CHS), and next-hop selection. The backbone formation mechanism divides the network area into clusters, and the CHS mechanism selects the CH for each cluster based on residual energy and distance to the BS. The next-hop selection mechanism connects nodes in the network into the chain for data transfer, choosing the appropriate next-hop to prevent connection failure. Zi and Chen *et al.* [31] proposed the branched-chain routing protocol (BranChain) to improve the PEGASIS protocol [15]. BranChain evaluates node weights based on energy residual levels and distance from the BS and selects the node with the highest weight to become the CH. Each node then finds the nearest node to connect to the branched-chain using the greedy algorithm. Aziz *et al.* [32] proposed an efficient energy routing scheme by combining a chain formation algorithm and data fusion, a data compression method that reduces the data packet size before forwarding it to the BS, thereby prolonging the network lifetime. Experimental results show that this scheme is more efficient than PEGASIS and ETCS protocols due to not using data compression. However, data compression before transmission increases the complexity of calculations in nodes. Overall, chain-based clustering and data fusion techniques can be used to improve energy-efficient routing in WSNs.

C. Tree-based Routing

Energy-efficient routing is crucial for the longevity of WSNs; hitherto, various routing protocols have been developed to achieve this goal, including tree-based clustering. Kim *et al.* [33] proposed a tree-based clustering scheme called TBC, where each node in a cluster constructs a tree based on distance information from member nodes to their CH. They also developed a real-time data collection scheme based on a MST constructed by the CHs [18] named STDC, which has been developed based on a distributed clustering process in heterogeneous wireless sensor networks [18]. This scheme

utilizes a MST constructed by the CHs and employs HELLO messages, similar to the LEACH protocol, to facilitate the clustering process. Notably, STDC enhances the selection of CHs by incorporating both initial energy and residual energy of the sensor nodes into a probability model. Furthermore, STDC employs a multi-hop communication approach for intra-cluster data transmission, employing multiple spanning trees to mitigate the energy consumption associated with long links. Despite these improvements, the distributed clustering process in STDC necessitates additional energy due to the dissemination of numerous HELLO messages and places an increased workload on the CHs during the construction of spanning trees. Moreover, the determination of an optimal number of clusters in STDC proves challenging due to the arbitrary CH selection based on statistical probability models. Additionally, the inter-cluster data forwarding from the CHs to the BS occurs through single-hop communication over long links, resulting in significant energy consumption.

Karunanithy *et al.* [34] presented a cluster-tree based energy efficiency data aggregating protocol for industrial automation applications using WSNs and IoT called CTEEDG. In CTEEDG, CH nodes is selected based on the Fuzzy logic model with the remaining energy, the number of neighbors, and the average distance that is considered input parameters for the fuzzy logic inference system. The data transmission routes from the clusters toward the BS are based on establishing tree topology, as the result, CTEEDG preserves energy and lengthens the network lifespan. However, the intra-cluster communication links are still long away due to using single-hop transmission architecture, so the performance achievement is not high (see more (8)). Osamy *et al.* [35] proposed a cluster-tree routing algorithm for data aggregation (CTRSDG) that considers intra and inter-distance ratios, remaining energy, and the distance from candidate nodes to the BS. CTRS-DG uses compressive sensing data techniques for CH data fusion and builds a backbone tree for multi-hop routing. CTRS-DG outperforms CREEP [36] in the matter of energy consuming and network lifetime. Zhang *et al.* [37] introduced a hybrid tree-based and clustering routing protocol (HTC-RDC) that builds a tree structure for multi-hop communication between sensing nodes and the BS, improving network longevity. Anzola *et al.* [38] proposed a multi-hop routing algorithm called H-kdtree, which employs a k-d tree algorithm for routing to reduce communication costs. H-kdtree outperforms LEACH and LEACH-C protocols in terms of network lifespan, packet delivery ratio, and throughput. However, the weakness of this proposed protocol is that it does not compare with other multi-hop communication tree-based clustering algorithms, while LEACH and LEACH-C are based on single-hop transmission clustering schemes.

III. THE SYSTEM MODELS

This section addresses some sensor networks and energy-consumed models, which are fundamental for designing our EE-TLT routing protocol.

A. Network Model

We use both homogeneous and heterogeneous sensor network models that involve N micro-sensor nodes, where the deployment of network area has A square meter and only one BS device, which is far away from the nodes. In a homogeneous network setup, all micro-sensor nodes are the same energy initialization and other characteristics. In the heterogeneous network setup, it is assumed that N micro-sensor nodes have various starting energy levels that are utilized with three kinds of nodes with diverse energy levels: *normal*, *intermediate*, and *advanced*. Let M_1, M_2 denote the parentage of the entire N nodes of *intermediate* and *advanced* sensing nodes. Consequently, we have:

$$N_A = N \times M_2, N_I = N \times M_1, \text{ and } N_N = N - (N_I + N_A), \tag{1}$$

where N_A, N_I, and N_N are the corresponding amounts of *intermediate*, *advanced*, and *normal* sensing nodes [39], [40]. If E_0 denotes the initialization energy level of the *normal* node, then $E_0(1 + \alpha)$ and $E_0(1 + \beta)$ are the initiation energy of each *intermediate* and *advanced* node, respectively, where α and β are the factors of energy that is greater than the *normal* ones. Consequently, the total initialization energy of entire nodes within the network is as follows [41]:

$$E_{init} = E_0(N_N + N_I(1 + \alpha) + N_A(1 + \beta)) \tag{2}$$

In general, if N sensor nodes are dispersed uniformly in A square meter zone, the probability distribution density function called $\rho(x, y)$ is expressed as below [10]:

$$\rho(x, y) = 1/(A^2/nc), \tag{3}$$

where nc be the number of the clusters (sectors), the average area size of a single sensor node is:

$$A_{node} = A^2/N \tag{4}$$

The average distance of a node to the nearest neighbor is:

$$d_{toNB}^2 = A/\sqrt{N} \tag{5}$$

The average size of the occupied area of each cluster is approximated as below [10]:

$$A_{cluster} = A^2/nc \tag{6}$$

Assuming that the position of CH is in the center of the sector, the maximum distance from CH to the farthest node in each cluster is expressed as below: [10]

$$d_{cluster-max} = A/\sqrt{nc \times \pi} \tag{7}$$

Experiment upon simulation, it is assumed that overall nodes are randomly dispersed over a two-dimensional sensor field to always observe the surroundings and periodically send collected data to a BS device, which is defined as not limited to energy sources.
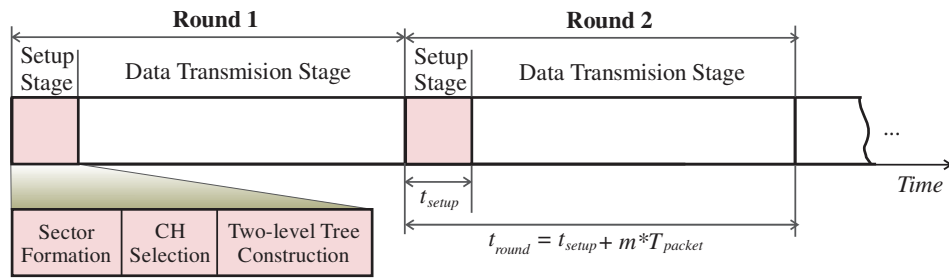
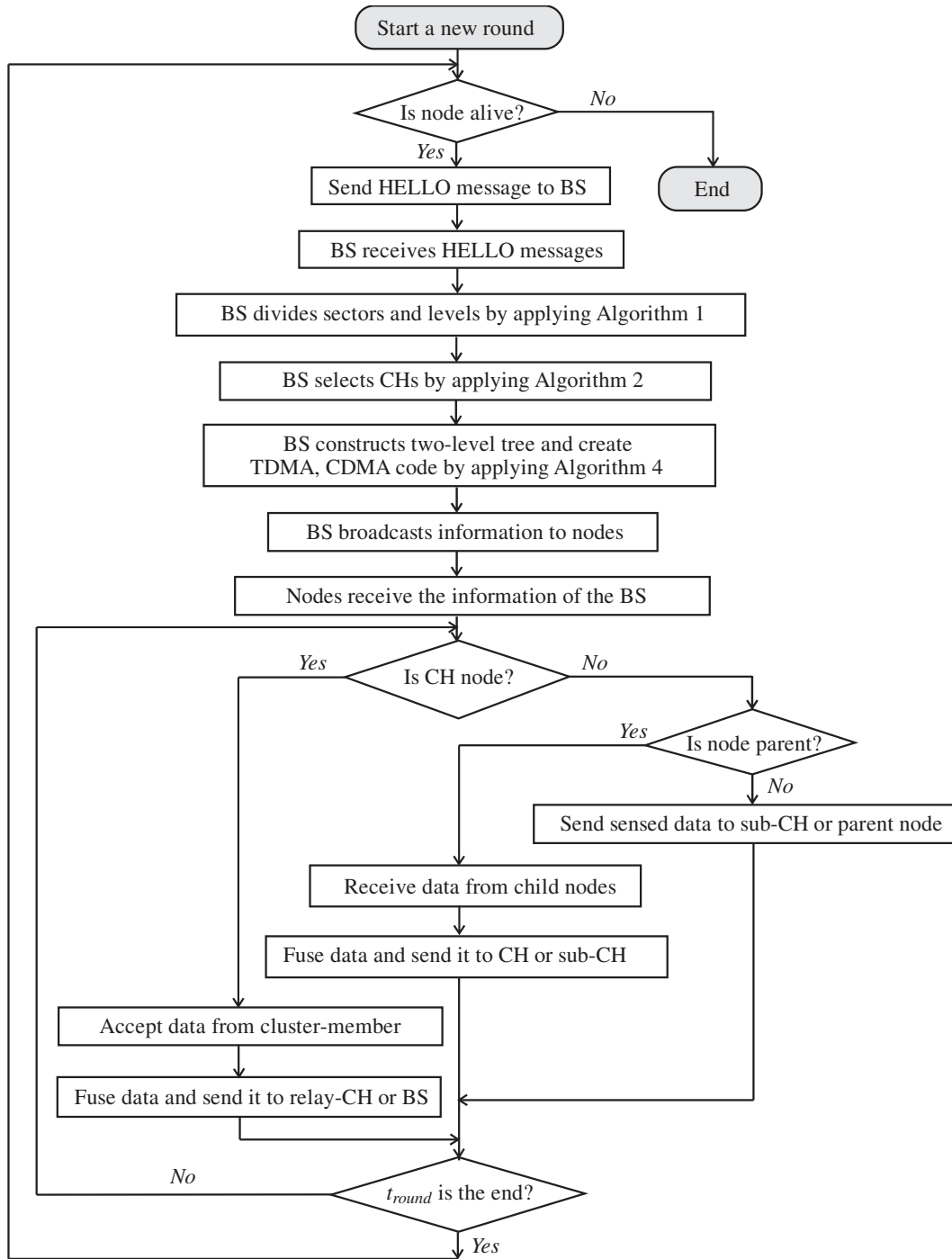Fig. 1. The division of EE-TLT protocol into rounds.

Fig. 2. Flowchart for the operation of EE-TLT protocol in a single round.
*Time division multiple access (TDMA), *Code division multiple access (CDMA).

B. Energy Consumption Model

We employ the energy dissipation model, which is similar to [32], [42] for consuming energy in wireless communication. Whenever a node transfers a packet containing q-bit data among two micro-sensor nodes with distance d, the energy used by the radio is calculated as follows:

$$E_{Tx}(q,d) = \begin{cases} q(E_{elec} + E_{friis}d^2) & , \text{ if } d < d_0 \\ q(E_{elec} + E_{tworay}d^4) & , \text{ if } d \geq d_0 \end{cases} \quad (8)$$

Here, E_{elec} is energy-consumed for the transmitter or receiver electric circuits; E_{friis} and E_{tworay} are the portions of energy amplification needed for the transceiver circuit according to the free space (d^2) or two-ray ground (d^4) model that depend on the distance d, and the threshold value d_0 is the crossover distance utilized in [43] in our simulation case:

$$d_0 = \sqrt{\frac{E_{friis}}{E_{tworay}}} \quad (9)$$

To receive a data packet containing q-bit, the transceiver circuit consumes energy as follows [43]:

$$E_{Rx}(q) = q \times E_{elec} \quad (10)$$

IV. DESCRIPTION OF EE-TLT

In this section, we present a novel two-level tree routing protocol called EE-TLT, which leverages the advantages of LEACH-centralized and tree-based routing methods [14], [18], [20]. The proposed protocol, as depicted in Fig. 1, operates in multiple rounds, each comprising two stages: the setup stage and the data transmission stage. During the setup stage, three steps are followed: (1) Sector formation, where the sensing field is logically divided into clusters and polygons; (2) CH selection, which involves choosing a CH or relay-CH for each sector and a sub-CH for each polygon based on factors such as energy residual and node-to-BS distance; (3) two-level tree formation, wherein routing trees are constructed using modern algorithms, potentially incorporating one or more MSTs to enable efficient data transmission within the network. The subsequent data gathering and transmission stage requires active nodes to continuously collect and transmit data to the BS while minimizing energy consumption throughout the round. Fig. 2 illustrates the operation diagram of EE-TLT in a single round, illustrating tasks such as sector separation, CH selection, and the establishment of multi-hop routes based on two-level trees. These processes are orchestrated by the BS to alleviate the overall energy consumption burden on the network.

A. Stage 1: Setup

The setup stage is performed in three steps to select leader nodes for clusters and polygons as follows:

1) Step 1: Sector partition with balancing number of nodes
At first, nodes transmit the HELLO messages between them and the BS. Then, the BS divides the sensing field of the entire
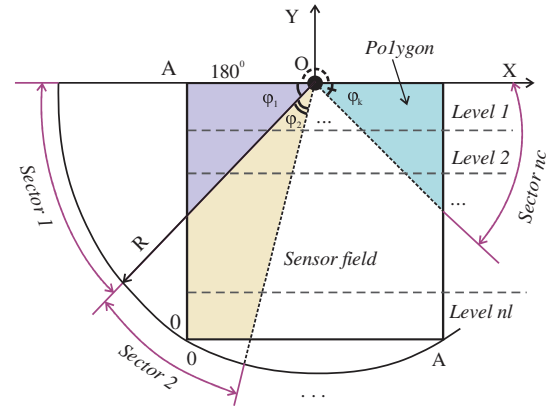

Fig. 3. Partitioning of the network into sectors and levels.

network into nc logical sectors, which are equal to nc clusters with unrealistic arcs covering the entire monitoring field as displayed in Fig. 3. Here we assume that there is an example of the network topology consisting of 100 sensor nodes in a 100 square meters area and the BS at (49,100) [10], [35]. Considering the BS at the polar coordinate origin of the XOY system as shown in Fig. 3, in which the BS determines the angle φ of all nodes in the network according to the location given as follows:

$$\omega = arctan\left(\frac{Y}{X}\right)\frac{180}{\pi}, \quad (11)$$

where $X = \mid x - x_{BS} \mid, Y = \mid y - y_{BS} \mid$ and

$$\varphi = \begin{cases} \omega & , \text{ if } X > 0 \text{ and } Y > 0 \\ \omega + 360 & , \text{ if } X > 0 \text{ and } Y < 0 \\ \omega + 180 & , \text{ otherwise,} \end{cases} \quad (12)$$

where x, y, x_{BS}, and y_{BS} are the position of the nodes and the BS in two-dimensional coordinate system, respectively. Next, BS further divides the sensing field into nl logical levels, which consist of some unequal polygons, the nodes in one polygon will be connected into a MST. To achieve a balanced distribution of nodes among clusters, Algorithm 1 is employed, which consists of steps 1 to 6. This approach takes into consideration that the total number of live nodes may fluctuate over time. Subsequently, steps 7 to 13 are carried out to further divide the nodes within each cluster into several smaller polygons.

2) Step 2: Selecting cluster head nodes
In the context of energy-efficient routing, data fusion, and tree-based clustering in WSNs, we provide the brief definitions outline the key roles of CHs, sub-CHs, and relay-CHs as follows.
− *CHs*: Selected from sensor nodes, CHs perform coordination and advanced functions within their clusters. They aggregate and process data from the sensor nodes and forward it to the BS or sink node, aiming to reduce overall communication energy consumption.
− *sub-CHs*: Within each cluster, sub-CHs assist the CH in data aggregation and routing tasks. They act as intermediaries

Algorithm 1 Sector and level division

Input: N sensor nodes with x, y position and current energy
Output: N nodes are distributed in nc logical clusters and np polygons with $N/nc/np$ nodes in each polygon

1: **for** $i = 1$ to N **do**
2:     Transfer HELLO message to the BS, which contains its ID, remaining energy and location;
3: **end for**
4: Compute the angle φ of every node in (11) and (12);
5: Sort the list of nodes without reducing the angle φ;
6: Distribute the network zone into nc sectors with equivalent N/nc nodes for each sector according to their φ angle as shown in Fig. 3;
7: **for** $i = 1$ to nc **do**
8:     Set nnc = {number of nodes in cluster ith};
9:     Sort nodes in sector ith non-decreasing by y coordinate;
10:     **for** $j = 1$ to nl **do**
11:         Distribute nnc/nl nodes in level jth into the set of polygons jth as shown in Figs. 4 and 5;
12:     **end for**
13: **end for**
14: **return:** {N/nc nodes in nc sectors and list of nodes in np polygons}.

Fig. 4. Determining the distances between sub-CHs located at different levels to construct a two-level tree routing.

between sensor nodes and the CH, facilitating efficient data gathering and transmission within the cluster.

− *Relay-CHs*: In multi-hop communication or large-scale sensor networks, relay-CHs are chosen from CHs. They relay data between different clusters or bridge communication gaps to enable efficient inter-cluster communication. Relay-CHs receive data from their clusters and forward it towards the base station or other designated destinations. The other nodes in the cluster (not CH, sub-CH, or relay-CH) are called cluster member nodes.

In each round, EE-TLT selects a CH or relay-CH for each sector and a sub-CH for each polygon based on the criterion below:

Average residual energy ($EC_{avg}(j)$, $EP_{avg}(j,l)$): We consider $EC_{avg}(j)$ as the average residual energy of alive nodes in sector jth, and $EP_{avg}(j,l)$ as the average residual energy of living nodes in sector jth and polygon lth. These are the most important features of candidate nodes that need to become CH because they use more energy in transmitting data to BS.

$$EC_{avg}(j) = \frac{1}{nnc} \sum_{i=1}^{nnc} E_{res}(i), \qquad (13)$$

$$EP_{avg}(j,l) = \frac{1}{nnp} \sum_{i=1}^{nnp} E_{res}(i), \qquad (14)$$

where $E_{res}(i)$ is the remaining energy of sensing node ith, nnc and nnp are the respective amounts of nodes in the present cluster and polygon.

Distance to BS (d_{toBS}): d_{toBS} is a crucial criterion that should be considered because the longer the data transmission

distance, the more energy is consumed. Accordingly, in cluster jth, EE-TLT chooses a CH node that is near the BS and has a residual energy level greater than the $EC_{avg}(j)$ as the optimal fitness function to make a decision as follows:

$$ff(i) = \frac{c_1 \times E_{res}(i)}{c_2 \times d_{toBS}(i, BS)}, \qquad (15)$$

where $d_{toBS}(i, BS)$ is the geographical distance from the current node ith to the BS, calculated as below:

$$d_{toBS}(i, BS) = \sqrt{(x_i - x_{BS})^2 + (y_i - y_{BS})^2} \qquad (16)$$

Users can set up the coefficient values of fitness function c_1 for energy and c_2 for Euclidean distance in different features of the WSN model.

Inter-level distance (d_{toCH})**:** The objective of this criterion is to minimize the inter-polygon distance between sub-CHs and the respective relay-CH in the two-level tree, which consumes less energy and balances among CHs. To achieve this objective, the sub-CHs will choose another sub-CH as their parent node, which not only pays attention to inter-polygon communication cost but also considers the cost of communicating with the BS of CHs or relay-CHs. Fig. 4 shows that the sub-CH ith node will choose the sub-CH ath to become its parent node because it is not far distance from the sub-CH ath to the BS compared to that of sub-CH bth to the BS, although the distance between sub-CH ith and bth is shorter than the distance to sub-CH ath. The criterion for the vote of sub-CH in polygons of each cluster can be computed as below:

$$sub\_CH(i) = \frac{c_1 \times E_{res}(i) \times nh}{10c_2 \times d_{toBS}(i, BS)}, \qquad (17)$$

where nh is the number of neighbor nodes of the candidate CH ith within the radio range area.

Relay cluster head election
In EE-TLT, only several relay-CHs transport aggregated data to the BS to save energy in other nodes, and of course, the distance between them and the BS should be short to prolong the network lifespan. Therefore, if the distance from the CH to the BS does not exceed the average distance D_{avg} between

Algorithm 2 Cluster head election

Input: N *sensor nodes with x, y positions and the current energy within nc clusters and np polygons*

Output: *List of CHs, sub-CHs and relay-CHs in clusters and polygons*

1: **for** $i = 1$ to nc **do**
2:     Calculate the average energy of nodes in cluster ith as in (13)
3:     Choose CH node, which has the highest value of the fitness function as (15)
4:     Append CH to {the list of CHs}
5: **end for**
6: **for** $i = 1$ to nc **do**
7:     **for** $i = 1$ to np **do**
8:       Calculate the average energy of nodes in each polygon as in (14)
9:       Choose a sub-CH node, which has the highest value of the cost function as (17)
10:       Append sub-CH to {the list of sub-CHs}
11:     **end for**
12: **end for**
13: Calculate the average distance D_{avg} among overall CHs and BS as (18)
14: **for** each CH_i in {list of the CHs} **do**
15:     **if** $((d(CHi, BS) < D_{avg})$or$(y_{CHi}$ in *Level 1* or *Level 2*)) **then**
16:       **if** ({number of relay-CH} $< nc/2$) **then**
17:         Append CH_i to {list of the relay-CHs}
18:       **end if**
19:     **end if**
20: **end for**
21: Calculate round size (t_{round}) for this round based on (31) and (35)
22: **return:** {the list of CHs, sub-CHs and relay-CHs};

CHs within the list of CHs to the BS or if the position of CH in *Level 1* or *Level 2* zones (Figs. 4 and 5), then that CH is selected to become relay-CH in this round. The amount of the selected relay-CHs is smaller than half of that of CHs and the D_{avg} is computed as below:

$$D_{avg}(r) = \frac{1}{nc} \sum_{i=1}^{nc} d(CH_i, BS) \qquad (18)$$

In Algorithm 2, steps 1 to 5 identify the CH nodes associated with the clusters, while steps 6 to 12 determine the selection of sub-CH nodes for polygons. The subsequent steps, specifically steps 13 to 20, pertain to the generation of the list of CHs.

3) Step 3: Two-level tree construction
Fig. 5 demonstrates that the network topology comprises $N = 100$ micro-sensor nodes within a 100 square meter region. The network zone is divided into five logical sectors, so each sector contains 20 nodes. Furthermore, each cluster is subdivided into five non-overlapping levels, creating multiple virtual polygons within the network. After that, every node in a polygon is linked to a MST with a sub-CH node as the root.

Fig. 5. Partitioning of the network into sectors and levels.

The sub-CHs are again connected into a two-level tree with the BS as the root. To finish this work, we assume that E_{dn} and E_{dCH} contain a set of links connected by nodes in the polygon and sub-CH, and CH or relay-CH and the BS in the network, respectively. The creation of multiple MSTs and two-level trees is achieved using the Kruskal [44]. Algorithms 3 and 4 below illustrate the process of constructing MSTs and two-level trees, respectively. In Algorithm 3, an MST is constructed using Kruskal's algorithm for the nodes within a polygon, with the sub-CH serving as the root. To establish multi-hop communication routes with minimal communication costs and enhance energy efficiency, Algorithm 4 executes Algorithm 3 np times, building MSTs for np polygons. This process constructs a two-level tree that connects the CHs and the BS, facilitating energy-efficient data transmission and reducing overall energy consumption.

B. Stage 2: Data Gathering and Transmission

In the proposed approach, sensor nodes are first grouped into sectors and CHs are selected. A two-level tree is constructed to identify the data transmission routes with the lowest energy consumption costs. Sensor nodes periodically monitor the environment and send the collected data towards the BS through the two-level tree with intra-polygon and inter-polygon communication. The leaf nodes at the highest level in each MST start data transmitting to their parent or sub-CH node according to the tree, and parent nodes sense the environment and receive data packets from their children. Then aggregate the data packets into a single packet and forward them to the sub-CH nodes at the upper-level. The sub-CHs transmit to their parent nodes on a two-level tree and the relay-CH nodes transfer all the data to the BS. After a fixed interval of time (t_{round}), the next round starts again by repartitioning

Algorithm 3 MST formation

Input: - nnp the number of sensor nodes in the polygon, and sub-CH;
   - E_{dn} the list of edge creating by nodes in polygon;
Output: MST with sub-CH as a root.

1: count = 0;
2: edgi = 0;
3: MST = MST $\cup$ {sub-CH};
4: Set sub-CH as root node;
5: **for** i = 1 to {number of the edges} **do**
6:    $E_{dn}[i].selected$ = FALSE;
7: **end for**
8: Sort {list of the edges in E_{dn}} as not ascending of value
9: **while** (true) **do**
10:    Select edgi in E_{dn}, whose $E_{dn}[edig].selected$ equals FALSE;
11:    du = get root of ($E_{dn}[edig].du$);
12:    dv = get root of ($E_{dn}[edig].dv$);
13:    **if** (du and dv are two points on different trees) **then**
14:       Union($E_{dn}[edig].du,E_{dn}[edig].dv$)
15:       $E_{dn}[edig].selected$ = TRUE;
16:       count = count + 1;
17:       **if** (count = nnp-1) **then**
18:          break;
19:       **end if**
20:    **end if**
21:    Set edgi = edgi + 1;
22: **end while**
23: **for** edgi =1 to {number of edges in E_{dn}} **do**
24:    **if** ($E_{dn}[edgi].selected$ = TRUE) **then**
25:       Set MST = MST $\cup$ {$E_{dn}[edgi].du$, $E_{dn}[edgi].dv$};
26:    **end if**
27: **end for**
28: **return** {MST};

clusters, polygons, as well as reselecting CHs, relay-CHs, sub-CHs, and rebuilding two-level trees in the network for a fresh round.

C. Evaluation of Energy Consumed of Micro-sensors in EE-TLT

1) The setup stage: We assume that there are N micro-sensor nodes implemented uniformly in the region of A square meter with nc clusters (sectors). Therefore, it will have N/nc nodes in each sector consisting of one relay-CH or CH and $(N/nc) - 1$ non-CH node which contains both member nodes and sub-CHs. Let E_{setup} be the energy used by nodes in a sector where nodes transfer control information to the BS at the round ith as follows:

$$E_{setup}(r) = \frac{N}{nc}\left(qE_{elec} + qE_{tworay}d^4_{toBS}\right) + \frac{N}{nc}qE_{elec} \tag{19}$$

The previous part of (19) shows the energy utilized by nodes that broadcast HELLO messages to the BS, which contains residual energy, position, and identity. The last part

Algorithm 4 Two-level tree formation

Input: - N the number of alive nodes in the network, and the BS;
   - E_{dCH} the list of edge created by CH, sub-CHs, relay-CH, and the BS ;
Output: Two-level tree with BS as a root.

1: **for** p = 1 to {number of polygons in the network} **do**
2:    Set nnp = {number of nodes in polygon pth} ;
3:    Call Algorithm 3 for constructing MST for polygon pth with sub-CH as the root;
4: **end for**
5: Set two-levels_TREE = two-levels_TREE $\cup$ {BS};
6: Call Algorithm 3 to construct a two-level tree for the list of sub-CHs with the BS as root
7: Create some time slot (TDMA) for all nodes in clusters
8: Distribute the time slot and two-level_TREE information to the network
9: **return:** {two-level_TREE};

of the equation displays the energy consumed for obtaining advertisement messages containing two-level trees, relay-CHs or sub-CHs, and time slots for communication planned from the BS.

2) The data gathering and transmission stage: The energy consumed by non-CH and relay-CH nodes in the stable data gathering and transmission stage can be described as follows:

(1). E_{mem}-the energy consumed by $(N/nc - nl)$ member nodes within a sector in a single packet is shown in (20), in which we assume that each member directly connects at most cn child nodes. So, after receiving cn packets from child nodes, the member node will aggregate cn packets with the measured data by itself to obtain a packet of the same size, and then send it to its parent node, or sub-CH.

$$E_{mem} = cn\left(\frac{N}{nc} - nl\right)qE_{elec} + \left(\frac{N}{nc} - nl\right)qE_{DA}$$
$$+ \left(\frac{N}{nc} - 1\right)(qE_{elec} + qE_{friis}d^2_{toCH}), \tag{20}$$

where nl is the amount of levels, d_{toCH} is the distance among the leaf nodes and their parent or sub-CH on MSTs and it can be expressed with (5) in the deployed uniform network.

(2). E_{sub-CH}-the energy consumed for accepting, aggregating, and forwarding the data of a sub-CH node in a two-level tree during a single packet is presented as below:

$$E_{sub-CH} = cn \times qE_{elec} + qE_{DA}$$
$$+ (qE_{elec} + qE_{friis}d^2_{toCH}), \tag{21}$$

where, cn is the amount of child nodes directly connected to sub-CH, d_{toCH} is the distance between the sub-CH nodes and their relay-CH, and can be expressed as (22) [10]:

$$E[d^2_{toCH}] = \iint (x^2 + y^2)\rho(x,y)dxdy$$
$$= \iint r^2\rho(r,\theta)drd\theta \tag{22}$$

According to (3) and (7), (22) can be estimated as follows:

$$E[d_{toCH}^2] = \rho \int\limits_{\theta=0}^{2\pi} d\theta \int\limits_{r=0}^{A/\sqrt{nc\pi}} r^3 dr = \rho \frac{A^4}{2nc^2\pi} = \frac{A^2}{2nc\pi} \quad (23)$$

(3). E_{CH}-the energy consumed by a CH, which consists of accepting data packets from np sub-CHs nodes on a two-level tree, aggregating data, and transmitting to other relay-CH or the BS, is presented below:

$$E_{CH} = cn \times qE_{elec} + qE_{DA} + (qE_{elec} + qE_{tworay}d_{toCH}^2) \quad (24)$$

If the CH is selected as the relay-CH for transmitting data to the BS, (24) can be replaced by (25) as shown below:

$$\begin{aligned} E_{relay-CH} = cn \times qE_{elec} + qE_{DA} \\ + (qE_{elec} + qE_{tworay}d_{toBS}^4), \end{aligned} \quad (25)$$

where d_{toBS} represents the distance between the relay-CH node and the BS.

(4). E_{sector}- the overall energy consumed by sensor nodes within a sector, which includes one CH or relay-CH, $(nl-1)$ sub-CH, and $(N/nc - nl)$ member nodes, can be indicated as follows:

$$E_{sector} = E_{relay-CH} + (nl-1)E_{sub-CH} + E_{mem} \quad (26)$$

(5). E_{round}- the overall energy dissipated for one round in sector:

$$E_{\text{round}} = E_{setup} + mE_{sector}, \quad (27)$$

where m is the number of packets sent during the data transmission stage in the round of a node. So, the energy totality consumption of the network running EE-TLT protocol during a round can be presented as:

$$\begin{aligned} E_{\text{total}} &= ncE_{\text{round}} \\ &= nc \times E_{setup} \\ &\quad + nc(m(E_{relay-CH} + (nl-1)E_{sub-CH} + E_{mem})) \end{aligned} \quad (28)$$

3) Network throughput and round length: In the current paper, the throughput of network Q of a WSN is defined as the total of packets successfully delivered from all living microsensor nodes to the destination (BS) per time unit [45]. If m_i data packets are successfully forwarded a node to the BS during the data collecting and transmission stage in round ith, then the throughput Q may be computed as:

$$Q = \sum_{i=1}^{N_{\text{round}}} N \times m_i, \quad (29)$$

where N_{round} denotes the overall amount of rounds and according to (2) and (27), we can calculate N_{round} as below:

$$N_{\text{round}} = \frac{E_{init}}{ncE_{\text{round}}} \quad (30)$$

If t_{round} indicates the time duration in a round, then:

$$t_{\text{round}} = t_{setup} + m \times T_{\text{packet}}, \quad (31)$$

Fig. 6. The energy efficiency of EE-TLT for varying numbers of clusters and levels.

where, t_{setup}, T_{packet} stand for the time length in the setup stage and single packet transmission, respectively. Thus, the throughput Q will enlarge if we diminish t_{setup} and enlarge $m * T_{\text{packet}}$ components. However, both t_{setup} and T_{packet} are fixed in every round; therefore, we may only enlarge m in the data collecting and transmission stage (the size of rounds). Furthermore, if the size of rounds is greatly increased, then the energy of relay-CH nodes will be quickly used up due to receiving and transmitting data packets at a distance farther from the BS than that of other nodes. Consequently, in EE-TLT protocol, we calculate suitably the dynamic t_{round} for different rounds. This make enhances efficient energy when applying EE-TLT protocol. Thanks to (8), (10), (25), and Table I, we can estimate the round length (t_{round}) in the data transmission stage in transmitting m data packets as follows:

$$\begin{aligned} E_{relay-CH}(i) = E_{res}(i) - m(ch \times qE_{elec} + qE_{DA} \\ + q(E_{elec} + E_{tworay}d_{toBS}^4)) \geq E_{theshold}, \end{aligned} \quad (32)$$

where $E_{theshold}$ is the threshold value of energy, which should be above zero to assure that the relay-CHs do not die when this round is completed. $E_{res}(i)$ and ch correspondingly express the residual energy of relay-CH ith and the number of connected nodes on the two-level tree, and:

$$m \leq \frac{E_{relay-CH}(i) - E_{theshold}}{q((ch+1)E_{elec} + E_{tworay}d_{toBS}^4 + E_{DA})} \quad (33)$$

The energy efficiency (EE) of a wireless sensor network (WSN) is defined as the amount of data packets received by the BS (measured as throughput Q) per unit of energy consumed (KB/J) during the network operation:

$$EE = \frac{q \times Q}{1000 \times E_{\text{total}}} \quad (34)$$

As a result, the objective function is given below:

$$f = argmax\,(EE) = argmax\,(N_{\text{round}} \times m) \quad (35)$$

Fig. 6 shows our simulation results in NS2 comparison of the energy efficiency (KB/J) with changing numbers of clusters and levels. It is clearly seen that EE-TLT achieves the highest EE when the number of sectors and levels are between 3 and 5 values in the network model with $N = 100$ nodes, $A = 100$ m, and distance from nodes to the BS = 175 m.

D. Complexity Analysis

The computational complexity of the LEACH protocol is $O(N)$, while LEACH-VA, which employs the Voronoi algorithm for clustering, has a complexity of $O(Nlog_2N)$, where N represents the number of nodes in the network. Similarly, the STDC and PEGCP protocols adopt a greedy algorithm to construct trees and chains, resulting in a computational complexity of $O(Nlog_2N)$. In EE-TLT, the sorting task in Algorithm 1 employs the QuickSort algorithm, leading to a complexity of $O(Nlog_2N)$. Algorithm 2 has a complexity of $O(N)$ due to the construction of CHs. Constructing an MST using the Kruskal algorithm in Algorithm 3 has a worst-case complexity of $O(((N/(nc \times nl))^2)$. Algorithm 4, which constructs two-level trees by invoking Algorithm 3 ($np + 1$) times, has a worst-case complexity of $O(N^2)$. Consequently, the worst-case complexity of EE-TLT is $O(N^2)$, exceeding the computational complexity of the three existing protocols, which is $O(Nlog_2N)$. Nonetheless, since the tasks of clustering, CH selection, and two-level tree construction are performed by the BS, EE-TLT can still be effectively employed in real-world applications.

V. SIMULATION AND PERFORMANCE EVALUATION

A. Performance Metrics

The performance of EE-TLT may be evaluated and compared with three different protocols according to the metrics below [14], [16], [18].

- *Network lifespan (t_{network})-* The network lifespan is determined as the duration of stable network operation until the occurrence of 1% node death, 50% node death, or 100% node death:

$$t_{\text{network}} = \sum_{i=1}^{N_{\text{round}}} t_{\text{round}}(i) \qquad (36)$$

- *Energy dissipation (E_{network})-* The total energy consumed by all participating sensing nodes during the entire network operation is computed using the following equation:

$$E_{\text{network}} = nc \sum_{i=1}^{N_{\text{round}}} E_{\text{total}}(i) \qquad (37)$$

- *The total number of data packets accepted by the BS (Q)-* The overall amount of data packets accepted by the BS from entire living nodes during the network operation which may be calculated by (29) above.
- *EE-* This metric indicates the proportion between the number of data packets successfully sent to the BS and the total energy consumed by the overall nodes (KB/J)

that can be computed by (34) above. An efficient energy routing protocol should not only reduce the total of energy consumed but also balance the energy distributed in all sensing nodes in WSNs.

B. Simulation Parameters

To verify the effectiveness of the EE-TLT, we simulated EE-TLT, LEACH-VA, PEGCP, and STDC by using the network simulator tool NS2 (v.2.34) [46], [47] with the scenarios and the parameters that are set up as in Table I [10], [35].

C. Simulation Scenarios

In this experiment, we have simulated the LEACH-VA [14] of block-based scheme, PEGCP [16] of chain-based scheme, STDC [18] of tree-based routing scheme, and the proposed EE-TLT protocols in many different scenarios to advance the reliability of the protocols in practice. Specifically, we used the "setdest" command in NS2 to randomly generate many different scenarios with the same simulation parameters. To determine how many scenarios (n_{sc}) are needed to run the simulation, we perform some steps as follows:

Step 1: Generate 100 random scenarios with 100 sensor nodes each, deployed in a stationary state over a simulation area of 100×100 m^2.

Step 2: Simulate the LEACH-VA, PEGCP, STDC, and EE-TLT protocols in the first scenario ($i = 1$). Then, create a table to document the percentage of node deaths, total energy consumption, and data packet reception rate at the base station.

Step 3: Select performance metrics, specifically energy efficiency and network lifespan, to evaluate the protocols.

Step 4: Run the next scenario ($i = i + 1$), and record the proportion of dead nodes, total energy consumed, and the number of data packets received by the BS.

Step 5: Calculate the mean (m_x), standard deviation (δ), and standard deviation ratio (ξ) using (38), (39), and (40).

Step 6: Compare the obtained results with previous scenarios. If the ratio of standard deviation is less than $\xi\%$, stop the simulation and proceed to **Step 7**. If not, return to **Step 4**.

Step 7: Graph the simulation results based on the mean and standard deviation with the scenario number ($n_{sc} = i$). Evaluate the performance of the protocols (i.e., LEACH-VA, PEGCP, STDC, and EE-TLT) in terms of dead nodes, energy consumption, and the quantity of data packets accepted by the BS.

We assume the mean m_x, standard deviation σ, and standard deviation ratio ξ can be expressed as below:

$$m_x = \frac{1}{n_{sc}} \sum_{i=1}^{n_{sc}} x_i, \qquad (38)$$

$$\sigma_i = \sqrt{\frac{1}{n_{sc}} \sum_{i=1}^{n_{sc}} (x_i - m_x)^2}, \qquad (39)$$

$$\xi = \sigma_i/max(x_i) \text{ with } i = 1, \cdots, n_{sc}, \qquad (40)$$

where x_i is the simulation results of LEACH-VA, PEGCP, STDC, and EE-TLT protocols at ith scenario. Specially, we

TABLE I
THE SIMULATION SCENARIOS INVOLVED VARIOUS PARAMETER VALUES.

Symbol	Parameters	Value
A	Simulation area	100×100 m^2
N	Number of micro-sensor nodes	100 nodes
E_{friis}	Energy amplification for free space	10 pJ/bit/m^2
E_{tworay}	Energy amplification for two ray ground	0.013 pJ/bit/m^4
E_{elec}	Electric energy	50 nJ/bit
E_{DA}	Energy consumption for data aggregation	5 nJ/bit
q	Packet size	1024 bytes
x_{BS}	The X-axis coordinate of BS	49 m
y_{BS}	The Y-axis coordinate of BS	175 m, 225 m, 265 m, and 300 m
c_1	Coefficient factors of energy	100/J
c_2	Coefficient factors of distance	5 m
	For the homogeneous network setup	
E_0	The initial energy of all nodes	2 J
	For the heterogeneous network setup	
E_0	The initial energy of a *normal* node	1 J
M_1	The parentage of *intermediate* in N nodes	30%
M_2	The parentage of *advanced* in N nodes	20%
α	Initial energy factor of *intermediate* nodes	0.5
β	Initial energy factor of *advanced* nodes	2

also simulate many different scenarios for the homogeneous and heterogeneous network model, as given in Table I.

D. Experimentation Results and Analysis

1) Homogeneous network: In Fig. 7, the ratio of living nodes and the overall network lifespan are plotted. The red lines in the figures which is correspond to the standard deviation ratio of the mean (m_x) for the EE-TLT protocol, with $\xi = 3.1\%$ at $n_{sc} = 19$ scenarios. This indicates that additional simulation scenarios will not yield results beyond the standard deviation curve. We choose $\xi = 6.8\%$ ($n_{sc} = 17$) for LEACH-VA, 5.6% ($n_{sc} = 19$) for PEGCP, and 9.3% ($n_{sc} = 24$) for STDC when the network had 95% of dead nodes. The ratio of living nodes in the EE-TLT protocol is higher by 25%, 15%, and 10% when compared to the LEACH-VA, PEGCP, and STDC protocols, respectively. The results in Fig. 7 indicate that the EE-TLT protocol with two-level tree-based clustering not only reduces energy consumption but also balances energy among nodes in the entire network.

Fig. 8 illustrates the total energy dissipation by the living nodes of the four protocols versus the network lifespan. It is apparent that EE-TLT uses less energy than the other protocols because it selects CHs based on the residual energy and distance between the candidate CH nodes and the BS. Furthermore, in LEACH-VA, PEGCP, and STDC, most of the CHs immediately forward the observed data to the BS, whereas in EE-TLT, only relay-CHs forward the collected data to the BS over a short distance, and other nodes send fused data packets in a two-level tree. As a result, EE-TLT achieves better energy efficiency and improves the network's lifespan in comparison to LEACH-VA, PEGCP, and STDC.

Fig. 9 presents the ratio of dead nodes versus the network lifespan in s. The results demonstrate that EE-TLT outperforms
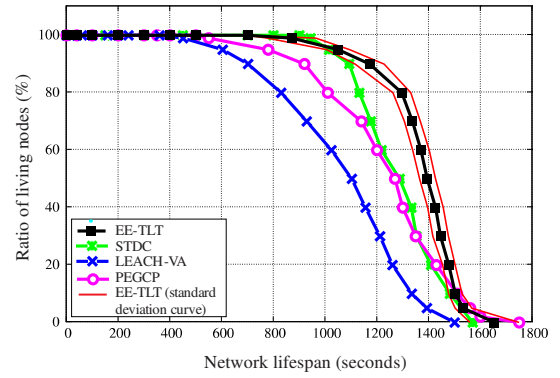

Fig. 7. The ratio of alive nodes of various routing protocols of the homogeneous network.
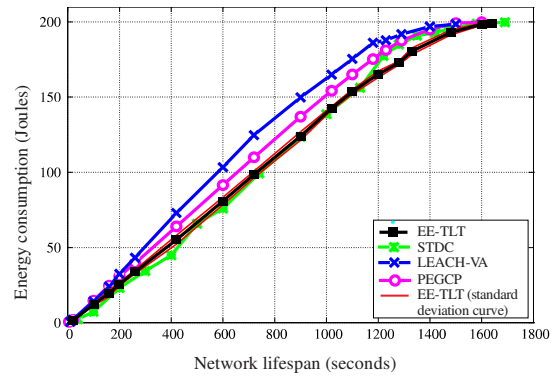

Fig. 8. Total energy utilization of the homogeneous network.

LEACH-VA, PEGCP, and STDC in terms of the network lifespan with FND and HND, considering the simulation scenarios and parameters as mentioned above.

In Fig. 10, we display the percentage of data packets
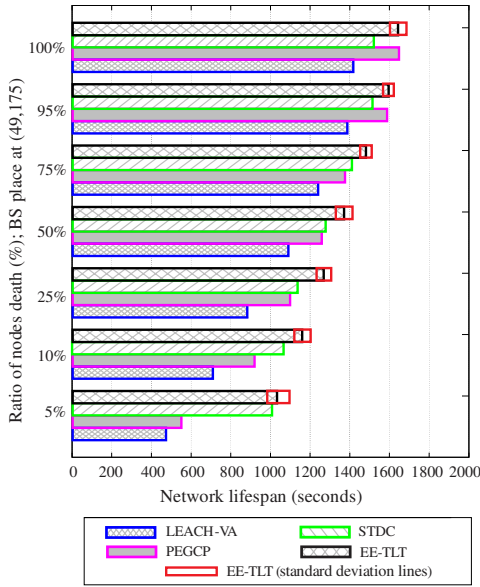
Fig. 9. The ratio of node deaths in different routing protocols of a homogeneous network.
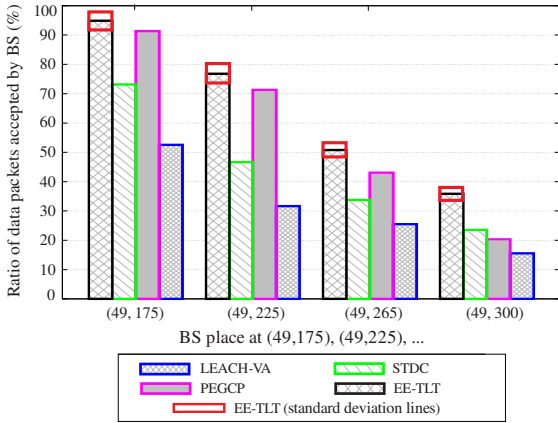

Fig. 10. The ratio of data packets obtained at the BS of the homogeneous network when the BS position changes.

accepted by the BS as its location changes. It was observed that when the BS is moved from its initial position (49, 175) to the endpoint (49, 300) beyond the simulation area, the total number of data packets received by the BS considerably reduces. Nevertheless, our proposed protocol outperforms LEACH-VA, STDC, and PEGCP by approximately 35%, 20%, and 15%, respectively. These results suggest that EE-TLT balances and reduces the energy consumption of sensor nodes within the WSN, which leads to an increase in the quantity of data packets received by the BS [27]. Additionally, the red lines in the figure represent the standard deviation ratio obtained by performing simulations from **Step 2** to **Step 7** as outlined in Section V.C.

Table II presents the evaluation results and comparison of the proportion of dead nodes and energy efficiency for WSNs in the homogeneous configuration as the t_{round} changes from 10 to 700 s. After running 19 scenarios applied with EE-TLT protocol ($t_{round} = 10$, $EE = 950$, and $\xi = 1.0\%$), 22 scenarios with LEACH-VA protocol ($t_{round} = 10$, $EE = 504$, and
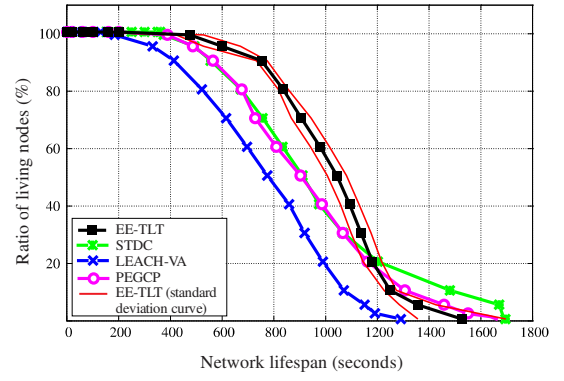

Fig. 11. The ratio of alive nodes of various routing protocols of a heterogeneous network.

$\xi = 4.1\%$), 18 scenarios with PEGCP protocol ($t_{round} = 10$, $EE = 670$, and $\xi = 10.4\%$), and 18 scenarios with STDC protocol ($t_{round} = 10$, $EE = 730$, and $\xi = 4.5\%$) that is displayed in bold text in the table. We selected the mean and the percentage of standard deviation based on the simulation results. It is observed that increasing t_{round} leads to an increase in energy efficiency, but the first nodes will die early. For instance, the EE of STDC protocol increases from 730 to 838 (KB/J) as t_{round} grows from 10 to 100 s. However, the first node dies sooner, from 804 to 112 s, respectively. Notably, EE-TLT significantly reduces from $EE = 1264$, $\xi = 1.8\%$ to $EE = 1092$, $\xi = 9.6\%$ as t_{round} increases from 500 to 700 s, respectively, due to the early death of CH. Thus, the question arises: how to select the appropriate t_{round} to achieve optimal performance? Table II shows that EE-TLT has the best performance with $EE = 1110$ (KB/J), the first node dies at 821 when $t_{round} = 100$ s, and the last node dies at 1702 s, which means that our proposal achieves a very good energy consumption balance. Meanwhile, when $t_{round} = 300$ s, SDTC protocol has the FND at 142 s and the LND at 3621 s, indicating that SDTC unbalances the power consumption load among nodes within the entire network.

2) Heterogeneous network: In the heterogeneous network, we ensure fairness by setting the *intermediate* and *advanced* nodes to 30% and 20% of the total nodes, respectively, as described in previous studies [36], [42]. The *intermediate* and *advanced* nodes have an energy initialization level that is 1.5 and 3 times greater than normal nodes, with $E_0 = 1$ Joule. Table I shows the parameters used to configure the simulated protocols. To evaluate the performance of EE-TLT, we follow the same simulation procedure as described in Section V.C, and compute the mean and the percentage of standard deviation based on the simulation results.

We present the simulation results in Fig. 11 showing the change in the ratio of dead nodes versus increasing the network lifespan in seconds. Our results indicate that the EE-TLT protocol, with carefully selected CHs, has a more balanced energy consumption and higher energy efficiency than LEACH-VA, PEGCP, and STDC. However, STDC has a longer LND due to unbalanced energy consumption among nodes within the entire network.

Fig. 12 displays the total energy dissipation by all sensing

TABLE II
THE ENERGY EFFICIENCY AND NETWORK LIFETIME WITH CHANGING t_{round} IN CONFIGURED HOMOGENEOUS WSN.

Protocols	Rounds (second)	EE (KB/J) (ξ%)	The percentage of dead nodes in the network (% dead nodes)						
			1% (ξ%)	10% (ξ%)	25% (ξ%)	50% (ξ%)	75% (ξ%)	95% (ξ%)	100% (ξ%)
LEACH-VA	**10**	**504(4.1)**	620(9.4)	777(7.8)	923(8.6)	1096(8.9)	1222(8.4)	1366(6.2)	1362(8.1)
	50	614(12)	92(80.7)	382(28)	704(17.2)	1008(12.2)	1276(11.8)	1534(10.7)	1541(10.0)
	100	781(7.5)	63(41.1)	315(28)	625(13.9)	1059(11.2)	1558(14.9)	1924(12.6)	1935(12.7)
	300	799(11.9)	47(33.3)	590(18.6)	1899(29.1)	2497(20.5)	2823(18.7)	3053(13.1)	3162(12)
	500	787(19.3)	55(30.3)	896(42.1)	2328(31.7)	3312(19.9)	3750(16.1)	4005(13.8)	4033(13.8)
	700	707(32.1)	54(28.6)	1296(21.6)	2898(9.9)	3565(8.8)	4091(4)	4393(5.5)	4453(4.9)
PEGCP	**10**	**670(10.4)**	465(29.0)	817(6.2)	995(4.3)	1204(3.9)	1343(5.0)	1546(7.5)	1903(19.5)
	50	1071(1.2)	519(19.5)	919(7.8)	1124(5.4)	1353(3.5)	1535(4.4)	1690(4.1)	1718(4.1)
	100	1095(1.5)	410(32.4)	898(7)	1139(7.9)	1373(5.7)	1552(4.7)	1781(5.4)	1807(5)
	300	1183(1.2)	232(16.3)	827(5.4)	1127(4.1)	1433(2.7)	1674(5)	1923(5.6)	2056(4.8)
	500	1241(2.4)	235(16.1)	816(14.1)	1160(5.4)	1460(3.3)	1758(4.6)	2075(8.8)	2240(7.8)
	700	903(10.2)	207(27.9)	864(15.9)	1186(17.1)	1470(16.3)	2329(15.9)	2628(13.6)	2628(13.6)
STDC	**10**	**730(4.5)**	**804(12.5)**	1082(8.9)	1314(8.3)	1494(7.4)	1603(6)	1831(8.5)	1941(7.9)
	50	**840(9.4)**	142(66)	636(26.8)	938(20.5)	1190(17.5)	1455(15.4)	1636(15.9)	1678(14.8)
	100	838(9.8)	**112(54.2)**	413(27.9)	809(15.1)	1253(10.5)	1682(9.3)	2064(7.6)	2071(7.3)
	300	576(6.8)	**142(69.3)**	626(31.6)	1512(17.1)	2746(14.2)	3328(8.2)	3544(9.6)	**3621(7)**
	500	570(8.3)	81(68.1)	953(26.1)	2433(23.3)	3351(23.1)	3800(13.7)	4066(11.4)	4151(12.1)
	700	557(9.5)	79(51.8)	1161(30.5)	2861(13.6)	3643(5.5)	4048(5.5)	4240(6.3)	4324(5.9)
EE-TLT	**10**	**950(1)**	751(12.5)	1065(7.6)	1227(1.9)	1273(2.2)	1334(1.7)	1429(4.2)	1544(9.3)
	50	1080(1.1)	811(26.1)	1163(7.1)	1309(1.8)	1387(1.2)	1446(1.1)	1528(2.3)	2037(27.8)
	100	**1110(1.2)**	821(27.3)	972(38.9)	1300(2.7)	1415(1.5)	1503(2.1)	1643(13.4)	**1702(10)**
	300	1202(1.5)	732(29.5)	924(38.4)	1303(3.7))	1500(1.5)	1636(1.9)	1867(2)	2159(9.4)
	500	**1264(1.8)**	607(41.1)	1065(23.2)	1326(23.2)	1564(23.1)	1800(23.5)	2057(23.2)	2458(25.6)
	700	**1092(9.6)**	414(32.7)	875(17.8)	1239(12.6)	1501(2.8)	1668(6.4)	2144(11.7)	2248(14.4)

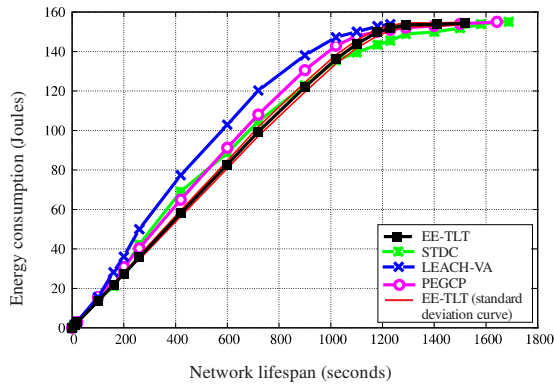Fig. 12. The total energy consumed by different routing protocols in a heterogeneous network.

Fig. 13. The node death ratio of different routing protocols in the heterogeneous network.

nodes in the heterogeneous WSN for four protocols during the network lifespan. It is evident that EE-TLT protocol has lower energy consumption in comparison to LEACH-VA, PEGCP, and STDC when the network lifespan is between 0 to about 1000 s. After that, EE-TLT consumes a little more energy than only STDC.

In Fig. 13, the percentages of node death in the heterogeneous network (5%, 10%, 25%, 50%, 75%, 95%, and 100%) are presented in reverse order to Fig. 11. It can be observed that EE-TLT improves the network lifespan by reducing the ratio of node death from 5% to 75% when compared to existing protocols such as LEACH-VA, PEGCP, and STDC. However, STDC has a longer network lifespan at 95% and 100% of nodes died because STDC unequally distributes
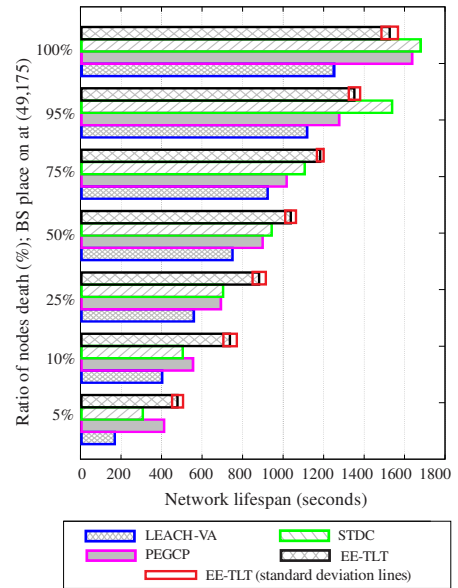
energy consumption. STDC sets up many *advanced* nodes with a high initial energy level, but they are not selected as CH, thus having more battery energy than other nodes.

Fig. 14 demonstrates the ratio of data packets acquired by the BS when the position of the BS is changed. This is also a valuable performance measurement for evaluating the high energy efficiency utilization of routing protocols because a network with more energy efficiency makes the BS receive

TABLE III
THE ENERGY EFFICIENCY AND NETWORK LIFETIME WITH CHANGING t_{round} IN CONFIGURED HOMOGENEOUS WSN.

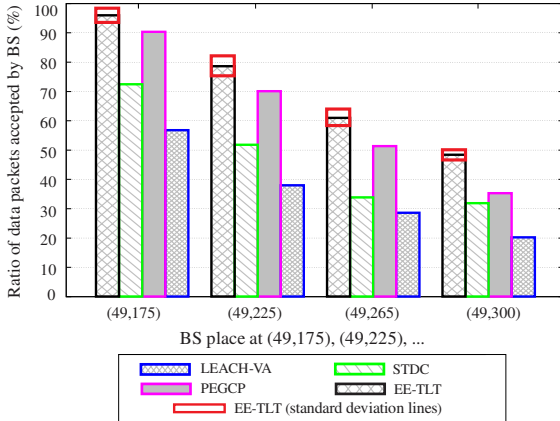| Protocols	Rounds (second)	EE (KB/J) ($\xi\%$)	The percentage of dead nodes in the network (% dead nodes)						
			1% ($\xi\%$)	10% ($\xi\%$)	25% ($\xi\%$)	50% ($\xi\%$)	75% ($\xi\%$)	95% ($\xi\%$)	100% ($\xi\%$)
LEACH-VA	10	494(4.4)	267(26.8)	472(10.7)	610(10.7)	778(7.2)	950(10.7)	1165(5.3)	1170(5.3)
	50	688(7.8)	36(44.4)	238(11.6)	466(12.7)	826(15)	1143(9.5)	1339(5.9)	1342(6.1)
	100	930(9.9)	35(26.7)	233(28.2)	562(20.3)	1064(21)	1420(21)	1630(14.8)	1636(14.8)
	300	758(23.7)	38(33.4)	570(23.3)	1431(19.7)	1853(18.9)	2195(13.3)	2585(13)	2593(13)
	500	1007(19.6)	45(70.6)	935(25.3)	1773(16.2)	2197(13)	2602(14.1)	3265(13)	3318(13.2)
	700	965(19.7)	44(34)	1239(23.3)	2026(15.4)	2470(13.1)	2888(13)	3842(11.8)	3921(8.9)
PEGCP	10	660(3.4)	323(26)	502(7.5)	647(2.7)	820(5.9)	1023(5.6)	1364(9.2)	2083(26.9)
	50	1069(1.7)	349(22.3)	557(8.7)	706(1.5)	906(4.8)	1121(4.6)	1467(11.9)	1568(13)
	100	1093(1.9)	326(32.6)	562(8.5)	710(1.3)	940(6.2)	1169(6.7)	1526(6.7)	1569(6.8)
	300	1142(2.0)	203(48)	544(8.7)	725(6.6)	970(8.5)	1234(6.9)	1630(6.2)	1702(6.6)
	500	1238(22.5)	184(30.1)	536(24.9)	715(22.7)	1029(24.6)	1435(28)	1863(24)	2041(25.4)
	700	936(2.7)	173(37.5)	536(11.6)	727(2.2)	1015(9)	1448(5.8)	2367(10.2)	3456(22.4)
STDC	10	753(12.9)	105(118)	468(16.8)	684(15.7)	846(28.1)	1131(14.5)	1546(10.8)	1548(10.8)
	50	879(6.7)	69(50.5)	314(20.6)	608(10.5)	942(9.3)	1221(10.8)	1421(8.0)	1426(8.1)
	100	1010(13.4)	62(78.3)	340(17.1)	679(16.9)	1205(9.6)	1482(12.2)	1695(24.2)	1707(24.2)
	300	**861(39.7)**	55(76.2)	551(25.2)	1359(12.5)	1813(14.5)	2152(14)	2625(11.1)	2628(11.1)
	500	**612(7.6)**	87(58.2)	1001(22.1)	1915(22.5)	2415(16.6)	2879(12.7)	3685(8.9)	3702(8.7)
	700	**595(7.3)**	61(55.7)	1217(34.4)	2021(14)	2522(10)	2926(11.3)	3873(7.5)	3889(7.6)
EE-TLT	10	963(20.3)	399(24.4)	627(7.0)	752(3.8)	909(5)	1024(3.6)	1211(7)	1577(20.3)
	50	1078(1.8)	470(13.6)	753(2.8)	853(7.1)	1050(4)	1167(5.4)	1356(14.2)	1577(28.5)
	100	1110(2.5)	441(33.2)	744(8.4)	832(4.2)	1044(3.3)	1160(2.8)	1361(7.9)	1675(23.7)
	300	1187(1.8)	438(15.8)	730(5.9)	826(3.4)	1091(4.5)	1276(1.4)	1576(5)	1860(12.9)
	500	**1250(1.8)**	**400(28.6)**	705(11.6)	838(10.7)	1147(6.7)	1430(6.7)	1730(6.2)	2178(11.7)
	700	1126(5)	403(18.3)	647(14.1)	797(2.4)	1135(2.7)	1461(1.1)	1880(18.2)	2095(11)

Fig. 14. The ratio of data bytes received at the BS in the heterogeneous network when the BS position changes.

more data bytes. From this figure, we may be seen there is a remarkable decline in the ratio of data bytes acquired by the BS in four protocols when the place of the BS is moved from the nearest point (49, 175) to the farthermost (49, 300) in the simulation zone. Nevertheless, the ratio of the packet acquired by EE-TLT is still higher than those of LEACH-VA, PEGCP, and STDC by about 25%, 15%, and 10%, respectively. This is because our proposed protocol calculates the suitable round size for every round in the stable data transportation stage, making the EE-TLT protocol outperform in terms of efficient energy and network lifespan in both heterogeneous and homogeneous network models for WSNs.

Table III presents the efficient energy and the percentage of dead nodes in FND, HND, and LND for LEACH-VA, PEGCP, STDC, and EE-TLT by changing t_{round} from the smallest 10 to 700 s in the heterogeneous WSN model. As may be seen in Table III, if we lengthen the size of round (time length within the data transportation stage t_{round}), the energy-efficient will also increase, but the FND will be earlier because relay-CHs consume more energy in transmitting to BS. Specifically, when $t_{\text{round}} = 300$, $EE = 861$, $t_{\text{round}} = 500$, $EE = 612$, and $t_{\text{round}} = 700$, $EE = 595$, using the single-hop inter-cluster communication method, the energy efficiency of the STDC protocol decreases deeply because CHs run out of energy early, and as a result, cluster member nodes waste energy. Whereas, EE-TLT protocol has better energy efficiency ($t_{\text{round}} = 500$ s, $EE = 1250$ KB/J, standard deviation ratio $\xi = 1.8\%$, though, the FND = 400 s, $\xi = 28.6\%$) than LEACH-VA, PEGCP and STDC. Besides, the proposed EE-TLT improves the network lifespan in terms of FND at ($t_{\text{round}} = 50$ s) by about 30.8%, 20.7%, and 15.1% at ($t_{\text{round}} = 10$ s) in comparison with LEACH-VA, PEGCP, and STDC protocols, respectively. This is because our proposed model chooses nodes with high residual energy and adjacent the BS to make relay-CHs. Also, EE-TLT constructs a two-level tree in each cluster to reduce overall the distance communication between nodes in the heterogeneous network model.

VI. CONCLUSION

Our paper presents a new energy-efficient routing protocol called EE-TLT, which aims to reduce energy consumption

in transmitting data for both homogeneous and heterogeneous network models in WSNs. Our work is focused on three primary contributions. Firstly, the proposed EE-TLT protocol balances energy consumption and network lifespan by distributing nodes into clusters and selecting CHs based on residual energy and distance to the BS, while avoiding "long links" communication through a two-level tree built using the Kruskal algorithm. Secondly, the protocol enhances the delivery of data packets by analyzing different transmission time durations and selecting optimal ones to improve throughput. Thirdly, simulation results show that EE-TLT outperforms several state-of-the-art protocols. Specifically, EE-TLT outperforms LEACH-VA, PEGCP, and STDC in terms of throughput, energy efficiency and balanced energy consumption, by about 25%, 15%, and 10%, respectively, in both heterogeneous and homogeneous networks. Although the overall time complexity of EE-TLT in the worst-case scenario is $O(N^2)$, which surpasses the computational complexity of the three existing protocols ($O(Nlog_2N)$), it is important to note that the tasks of clustering, CH selection, and two-level tree construction are performed by the BS. As a result, EE-TLT remains fully applicable in real-world applications. The code and the simulation results of EE-TLT may be found at https://tinyurl.com/ee-tlt-wsn. Our plan for future work includes enhancing the energy-efficient routing protocol by reducing the size of data packets transmitted to the base station in the data transmission stage, through the utilization of data fusion algorithms.

<div align="center">REFERENCES</div>

[1] U. Dampage, L. Bandaranayake, R. Wanasinghe, K. Kottahachchi, and B. Jayasanka, "Forest fire detection system using wireless sensor networks and machine learning," *Scientific Reports*, vol. 12, no. 46, 2022.

[2] T. Alhmiedat, "Low-power environmental monitoring system for ZigBee wireless sensor network," *KSII Trans. Internet Inf. Syst.*, vol. 11, no. 10, pp. 4781–4803, 2017.

[3] D. Kandris, C. Nakas, D. Vomvas, and G. Koulouras, "Applications of wireless sensor networks: An up to date survey," *MDPI Appl. Syst. Innov.*, vol. 14, no. 3, pp. 1–24, 2020.

[4] M. S. Al-kahtani, F. Khan, and W. Taekeun, "Application of Internet of things and sensors in healthcare," *Sensors*, vol. 22, no. 15, pp. 1–20, 2022.

[5] L. Wan *et al.*, "Distributed parameter estimation for mobile wireless sensor network based on cloud computing in battlefield surveillance system," *IEEE Access*, vol. 3, pp. 1729–1739, 2015.

[6] K. Dalal and P. Dahiya, "State-of-the-art in VANETs: The core of intelligent transportation system," *IUP J. Elect. Electron. Eng.*, vol. 10, no. 1, pp. 27–39, 2017.

[7] X. Huibin and Z. Mengjia, "An energy-efficient clustering routing for wireless sensor networks based on energy consumption optimization," *Int. J. Digit. Multimedia Broadcast.*, vol. 2022, pp. 1–11, 2022.

[8] O. J. Pandey and R. M. Hegde, "Low-latency and energy-balanced data transmission over cognitive small world WSN," *IEEE Trans. Veh. Technol.*, vol. 67, no. 8, pp. 7719–7733, 2018.

[9] L.-L. Hung, F.-Y. Leu, K.-L. Tsai, and C.-Y. Ko, "Energy-efficient cooperative routing scheme for heterogeneous wireless sensor networks," *IEEE Access*, vol. 8, pp. 56 321–56 332, 2020.

[10] W. B. Heinzelman, A. P. Chandrakasan, and H. Balakrishnan, "An application-specific protocol architecture for wireless microsensor networks," *IEEE Trans. Wireless Commun.*, vol. 1, no. 4, pp. 660–670, 2002.

[11] H. D. Viet and V. N. Dinh, "DF-SWin: Sliding windows for multi-sensor data Fusion in Wireless Sensor Networks," in *Proc. IEEE KSE*, Oct. 2017.

[12] F. Ye, J. Chen, and Y. Tian, "A robust DS combination method based on evidence correction and conflict redistribution," *J. Sensors*, vol. 2018, pp. 1–12, 2018.

[13] T. M. Behera *et al.*, "Energy-efficient routing protocols for wireless sensor networks: Architectures, strategies, and performance," *Electron.*, vol. 11, no. 15, 2022.

[14] H. Liang, S. Yang, LiLi, and J. Gao, "Research on routing optimization of WSNs based on improved LEACH protocol," *EURASIP J. Wireless Commun. Netw.*, vol. 1, no. 2019, pp. 1–12, 2019.

[15] S. Lindsey and C. S. Raghavendra, "PEGASIS: Power-efficient gathering in sensor information system," in *Proc. IEEE AeroConf*, Mar. 2002.

[16] F. Bouakkaz and M. Derdour, "Maximizing WSN life using power efficient grid-chain routing protocol (PEGCP)," *Wireless Personal Commun.*, vol. 117, 03 2021.

[17] M. Mittal, U. Singh, and R. Salgotra, "Tree based threshold-sensitive energy efcient routing approach for wireless sensor networks," *Wireless Personal Commun.*, vol. 108, no. 1, pp. 473–492, 2019.

[18] K. T. Kim, J. C. Park, M. Kim, U. M. Kim, and H. Y. Youn, "A spanning tree based data collection for real-time streaming sensor data," in *Proc. IEEE DASC*, Aug. 2014.

[19] V. Doryanizadeh, A. Keshavarzi, T. Derikvand, and M. Bohlouli, "Energy efficient cluster head selection in Internet of things using minimum spanning tree (EEMST)," *Appl. Artif. Intell.*, 2021.

[20] N. D. Tan and N. D. Viet, "DFTBC: Data fusion and tree-based clustering routing protocol for energy efficient in wireless sensor networks," in *Proc. AISC*, Oct. 2015.

[21] L. Chan, K. G. Chavez, H. Rudolph, and A. Hourani, "Hierarchical routing protocols for wireless sensor network: A compressive survey," *Wireless Networks*, vol. 26, no. 5, pp. 3291–3314, 2020.

[22] S. M. Altowaijri, "Efficient next-hop selection in multi-hop routing for IoT enabled wireless sensor networks," *Future Internet*, vol. 14, no. 2, 2022.

[23] S. Rani and S. H. Ahmed, *Multi-hop routing in wireless sensor networks*, Springer Briefs Elect. Comput. Eng., 2016.

[24] N. Sabor, S. Sasaki, M. Abo-Zahhad, and S. M. Ahmed, "A comprehensive survey on hierarchical-based routing protocols for mobile wireless sensor networks: Review, taxonomy, and future directions," *Wireless Commun. Mobile Comput.*, vol. 2017, pp. 1–23, 2017.

[25] N. Mittal, U. Singh, and B. S. Sohi, "A stable energy efficient clustering protocol for wireless sensor networks," *Wireless Netw.*, vol. 23, pp. 1809–1821, 2017.

[26] Q. Liu, K. Zhang, X. Liu, and N. Linge, "Grid routing: an energy efficient routing protocol for WSNs with single mobile sink," *Int. J. Sensor Netw.*, vol. 25, no. 2, pp. 93–103, 2017.

[27] Y. Liu, Q. Wu, T. Zhao, Y. Tie, F. Bai, and M. Jin, "An Improved Energy-Efficient Routing Protocol for Wireless Sensor Networks," *Sensors*, vol. 19, no. 20, pp. 2–20, 2019.

[28] L. Tang, Z. Lu, and B. Fan, "Energy Efficient and Reliable Routing Algorithm for Wireless Sensors Networks," *Appl. Sci.*, vol. 10, pp. 1–16, 2020.

[29] S. Firdous, N. Bibi, M. Wahid, and S. Alhazmi, "Efficient clustering based routing for energy management in wireless sensor network-assisted Internet of things," *Electron.*, vol. 11, no. 23, 2022.

[30] H. A. Marhoon, M. Mahmuddin, and S. A. Nor, "DCBRP: A deterministic chain based routing protocol for wireless sensor networks," *SpringerPlus*, vol. 5, no. 1, 2016.

[31] L. Zi, W. Chen, X. Liu, and X. Chen, "A novel chain-based routing protocol, BranChain, in wireless sensor networks," *Int. J. Embedded Syst.*, vol. 11, no. 3, pp. 259–268, 2019.

[32] A. Aziz, W. Osamy, A. M. Khedr, and A. Salim, "Chain-routing scheme with compressive sensing-based data acquisition for Internet of Things-based wireless sensor networks," *IET Netw.*, vol. 10, pp. 43–58, 2021.

[33] K. T. Kim, C. H. Lyu, S. S. Moon, and H. Y. Youn, "Tree-based clustering (TBC) for energy efficient wireless sensor networks," in *Proc. IEEE WAINA*, Apr. 2010.

[34] K. Karunanithy and B. Velusamy, "Cluster-tree based energy efficient data gathering protocol for industrial automation using WSNs and IoT," *J. Ind. Inf. Integration*, vol. 19, p. 100156, 2020.

[35] W. Osamy, A. M. Khedr, A. Aziz, and A. A. E. Sawy, "Cluster-tree routing based entropy scheme for data gathering in wireless sensor networks," *IEEE Access*, vol. 6, pp. 77 372–77 387, 2018.

[36] S. Dutt, S. Agrawal, and R. Vig, "Cluster-head restricted energy efficient protocol (CREEP) for routing in heterogeneous wireless sensor networks," *Wireless Personal Commun.*, vol. 100, no. 4, pp. 1477–1497, 2018.

[37] Y. Zhang, L. Liu, M. Wang, J. Wu, and H. Huang, "An improved routing protocol for raw data collection in multihop wireless sensor networks ," *Comput. Commun.*, no. 188, pp. 66–80, 2022.

[38] J. Anzola, J. Pascual, G. Tarazona, and R. G. Crespo, "A clustering WSN routing protocol based on k-d tree algorithm," *Sensors*, vol. 18, no. 9, pp. 1–26, 2018.

[39] W. Osamy, Ahm, Salim, and A. M. Khedr, "An information entropy based-clustering algorithm for heterogeneous wireless sensor networks," *Wireless Netw.*, vol. 26, no. 3, pp. 1869–1886, 2020.

[40] J. Yanfei, C. Guangda, and Z. Liquan, "Energy-efficient routing protocol based on zone for heterogeneous wireless sensor networks," *J. Elect. Comput. Engi.*, vol. 2021, 2021.

[41] M. Kaur, A. Jain, and A. K. Goel, "Energy efficient two level distributed clustering scheme to prolong stability period of wireless sensor network," in *Proc. IEEE ICACCI*, Sep. 2014.

[42] D. N. Quoc, N. Liu, and D. Guo, "A hybrid fault-tolerant routing based on Gaussian network for wireless sensor network," *J. Commun. Netw.*, vol. 24, no. 1, pp. 37–46, 2022.

[43] B. Han, F. Ran, J. Li, L. Yan, H. Shen, and A. Li, "A novel adaptive cluster based routing protocol for energy-harvesting wireless sensor networks," *Sensors*, vol. 22, no. 4, 2022.

[44] J. B. Kruskal, "On the shortest spanning subtree of a graph and the traveling salesman problem," in *Proc. AMS*, vol. 7, no. 1, pp. 48–50, 1956.

[45] X. Li, B. Keegan, F. Mtenzi, T. Weise, and M. Tan, "Energy-efficient load balancing ant based routing algorithm for wireless sensor networks," *IEEE Access*, vol. 7, pp. 113 182–113 196, 2019.

[46] VINT Project, "The network simulator - NS2," http://www.isi.edu/nsnam/ns (Accessed: 20 Jun 2022).

[47] W. R. Heinzelman, A. Chandrakasan, and H. Balakrishnan, "The MIT uAMPS ns Code Extensions Version 1.0," http://www-mtl.mit.edu/research/icsystems/uamps/research/leach/leach_code.shtml (Accessed: 20 Jun 2022), 2002.

Nguyen Duy Tan was born in Hung Yen, Viet Nam, in 1977. He received his M.S. and Ph.D. degree in Ha Noi University of Technology, Vietnam National University, Hanoi in 2009 and 2017, respectively. His research interests consist of data communication, computer networks, network security and IoT. Currently, he is a Lecturer of the Faculty of Information Technology, Hung Yen University of Technology Education. Email: tanndhyvn@gmail.com.

Van-Hau Nguyen Dr. Van-Hau Nguyen received an Engineer's degree in Applied Informatics Mathematics, and Master's degree in Information Technology in 2003 and 2006, respectively, at Hanoi University of Science and Technology. In 2015, he obtained the Ph.D. degree in Computer Science from the Artificial Intelligence lab at Technische Universitaet Dresden, Germany. He is currently the Director of the AI center at the Faculty of Information and Technology, Hung Yen University of Technology and Education, Vietnam. He has published more than 30 papers in international conferences and journals. His research interests include: Automated reasoning, machine learning, artificial intelligence, and IoT.