THE UNIVERSITY OF DANANG
UNIVERSITY OF SCIENCE AND EDUCATION

MAI HÀ THI

OPTIMIZATION OF ENERGY EFFICIENCT ROUTING IN MULTI DOMAINS WIRELESS UNDERGROUND SENSOR NETWORKS

DOCTORAL DISSERTATION SUMMARY

DaNang, Year 2025

INTRODUCTION

Overview

Wireless sensor networks (WSNs) are being widely deployed across numerous application domains, particularly in smart agriculture. The integration of WSNs with the Internet of Things (IoT) and data analytics enables resource optimization and enhances crop productivity. Nevertheless, sensor nodes in WSNs are intrinsically constrained in computation, memory, and energy, and large numbers of sensors must be deployed to achieve coverage over expansive areas.

In wireless underground sensor networks (WUSNs), these limitations become even more pronounced due to the harsh subsurface environment. Elevated humidity, high pressure, and corrosive agents shorten sensor lifetimes, while communication signals face substantial challenges because of attenuation caused by soil and buried materials. Supplying energy to, and maintaining, sensors under such conditions also pose significant technical and cost challenges.

The Internet is a multi-domain system that uses the Border Gateway Protocol (BGP) to exchange routing information, demanding high scalability and security. In multi-domain sensor networks, relay nodes can extend connectivity; however, concrete routing solutions remain lacking to jointly optimize coverage and network lifetime. Multi-constrained routing optimization remains difficult, and BGP itself exhibits limitations in promptly reacting to structural/topological changes in the network.

Urgentity of the dissertation

Research on WUSNs has achieved substantial advances in devising routing protocols to enhance coverage and reliability; however, several critical issues remain unresolved. Energy optimization is among the foremost concerns, as it is essential for ensuring the long-term sustainability of these networks. Notwithstanding recent progress, notable limitations persist, including:

- In approaching WUSNs, most studies model path loss, employ LoRaWAN for data transmission, and evaluate energy consumption.
- Prior work has largely applied network-routing protocol developments to datasets derived from two-dimensional environments.
- Most prevailing network assumptions concern homogeneous deployments, whereas heterogeneous networks remain comparatively underexplored.
- A principal shortcoming is the lack of a unifying theory for inter-regional data delivery that explicitly accounts for cross-region network lifetime.
- Relay nodes or sink nodes are commonly introduced to communicate with subsurface nodes located far from the base station. Ultimately, the installation costs of relay nodes and base stations can be substantial; hence, optimizing the use and placement of relay nodes within the network is a problem that warrants attention.

Routing research in WUSNs must place particular emphasis on the practical challenges that arise from the unique characteristics of their propagation channels. Traditional routing models are often not suited because of the distinctive subsurface energy-dissipation properties. The energy optimization problem in WSN especially in WUSNs amounts to selecting a path from a source to a destination such that the cumulative energy cost is minimized while preserving end-to-end connectivity across the network.

This dissertation focuses on developing new routing solutions that directly address energy-consumption optimization in WUSNs routing simulations, leveraging a more accurate energy or path loss model. The objective is to demonstrate the efficacy of these routing solutions through simulations over one or multiple regions, thereby evidencing their potential to substantially extend the overall operational lifetime of the network.

Research objectives of the dissertation

This dissertation investigates and proposes energy-optimization solutions for WUSNs, with the goal of developing routing algorithms and models that improve energy efficiency under single and multi areas observational conditions. According to the general goals above, the thesis investigates the idea of creating an energy optimization model through network routing:

- Modeling routing strategies in WUSNs: Formulate mathematical models for routing methods aimed at minimizing energy expenditure in wireless underground sensor networks.
- Proposed energy-optimization algorithms: Design and develop routing algorithms that optimize energy utilization in WUSNs, with operability in both single- and multi-domain settings.
- Experimental implementation of the proposed algorithms: Implement and test the algorithms in realistic testbeds or simulation environments to evaluate energy savings and the maintenance of network lifetime.
- Evaluation and comparative analysis: Conduct concrete measurements to assess the algorithms' energy efficiency and reliability in WUSNs, with the aim of advancing the current state of the art.

Contribution

The dissertation investigates the network lifetime optimization problem for WUSNs under a layered architecture grounded in a path-loss model. The study considers the organization and management of sensor-node data transmissions via region wise clustering and introduces a new metric for multi-region observation. Designing a WUSN routing model remains a principal challenge that this dissertation seeks to address. In parallel, routing is leveraged to prolong the lifetime of underground sensor nodes across diverse applications. To achieve this overarching aim, the dissertation sets forth the following specific research objectives:

- Theoretical contributions

  - **Propose a clustering-based network routing model to optimize energy in WUSNs** .

    - First, the dissertation develops a new mathematical model by extending clustering algorithms to WUSNs through the joint treatment of surface (deployed) and subsurface nodes. To optimize cluster-head (CH) selection, an objective function is established that incorporates the energy-consumption model, with optional CH placement derived in closed form.

    - Second, the dissertation introduces a clustered networking protocol comprising an optimal CH-selection algorithm and an efficient routing method. The protocol is designed for practical case studies, explicitly accounting for path loss. Experiments are conducted to validate the reliability of the proposed model and protocol.

  - **Propose a multigraph-based network routing model for monitoring multiple regions in WUSNs.**

    - Construct a mathematical model and associated solutions for homogeneous wireless sensor networks augmented with relay nodes. Several theorems and properties of the proposed model are established.

    - Develop a new routing algorithm that determines each sensor's shortest path using a hash-table–based approach.

    - Provide experimental results demonstrating the effectiveness of the new algorithm relative to related methods.

- Applied contributions: The proposed routing models and algorithms address key challenges in multi-domain WUSNs. As the backbone of WUSN application systems, the routing algorithms support scalability from single-region to multi-domain settings, enabling energy optimization that extends network lifetime in complex and heterogeneous deployments.

Structure of Doctoral Dissertation

The thesis entitled **"Methods for Modeling Network Topology in Wireless Underground Sensor Networks"** comprises an introduction, four principal chapters, a conclusion, and a bibliography, encompassing the following essential components:

- **Introduction**: The research context is presented, followed by a survey of routing protocols in WSNs and their limitations. The research problems are identified together with the objectives. The methodological approach and research methods are outlined, with details on the study's content, scope, and boundaries. Finally, the principal contributions and the overall dissertation structure are discussed.
- **Chapter 1**: An introduction to WUSNs covering their architecture, constituent components, and subsurface wave-propagation models, with emphasis on how operating frequency, burial depth, soil moisture, and soil composition affect signal attenuation. The chapter also reviews optimization problems in such networks and routing issues in multidimensional settings. Related concepts, a formal statement of the routing problem, and its applications are presented.
- **Chapter 2**: This chapter focuses on proposing energy-optimization methods to extend WUSNs lifetime in both single-region and multi-region environments, with particular emphasis on inter-region routing for multi-regions scenarios. The work proceeds from mathematical modeling to the construction of the proposed algorithms, with validation through analytical assessments. The chapter concludes by detailing the new routing algorithms developed in the dissertation.
- **Chapter 3**: This chapter presents experiments that substantiate the effectiveness of the models and algorithms proposed in Chapter 2 by means of empirical evidence. The algorithms are implemented in MATLAB, with clearly defined network, channel, and energy parameters. Experimental sce-

narios are designed to compare the proposed algorithms (for both single- and multi-region cases) against canonical baselines such as PEGASIS, H-LEACH, and FCM. The scenarios also examine the impact of factors including node density, variations in burial depth, and soil properties.

- **Conclusion and future direction**: The concluding chapter synthesizes the main results of the dissertation, reaffirming the key scientific contributions: a new mathematical model for WUSNs, an efficient routing algorithm for single-region scenarios, and an inter-region routing solution for multi-region scenarios, together with experimental evaluations benchmarked against the criteria established in Chapters 2 and 3. The chapter also identifies remaining limitations encountered during the research and, finally, discusses prospective research directions for WUSNs.

Chapter 1

OVERVIEW OF MULTI-DOMAIN WIRELESS UNDERGROUND SENSOR NETWORKS

1.1 Introduction

This chapter surveys the theoretical foundations of WSNs and their extension to WUSNs. It reviews the principal approaches to routing optimization, with emphasis on heuristic algorithms and approximation algorithms. The notion of wavelength in WUSNs, energy consumption models in WSNs, and multi-domain networking are also introduced. The chapter concludes by describing experimental environments and evaluation criteria for WSN experiments. These foundational concepts provide the basis for routing methods aimed at achieving energy-efficient operation in WUSNs.

1.1.1 Wireless Sensor Network

Wireless sensor networks comprise large populations of sensor nodes that communicate over wireless channels to deliver data to a base station. WSNs have been employed across a broad spectrum of applications, including wildlife monitoring, environmental surveillance, detection of subsurface gas and water leaks, structural health monitoring, fire detection, and healthcare. These networks are capable of acquiring environmental data (e.g., light intensity or temperature readings) and subsequently forwarding this information to the base station.

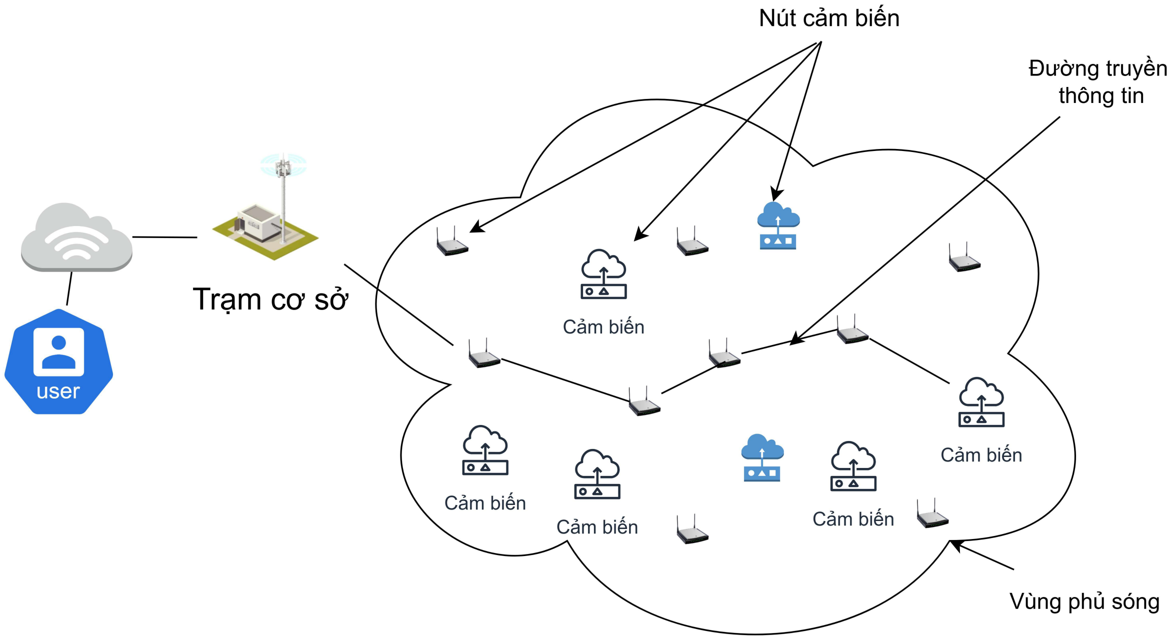

Figure 1.1: Wireless Sensor Network architecture.

1.2 Extensions of Wireless Sensor Networks

1.2.1 Wireless Underground Sensor Networks

Wireless Underground Sensor Networks (WUSN) consist of sensor nodes buried underground, transmitting data to a sink (either underground or above ground) via electromagnetic (EM) waves or magnetic induction (MI) communication [1, 2]. Data is transmitted in multiple hops until it reaches the gateway, which serves as the destination node [3].

The three basic communication channels [4] – illustrated in Figure 1.2 –

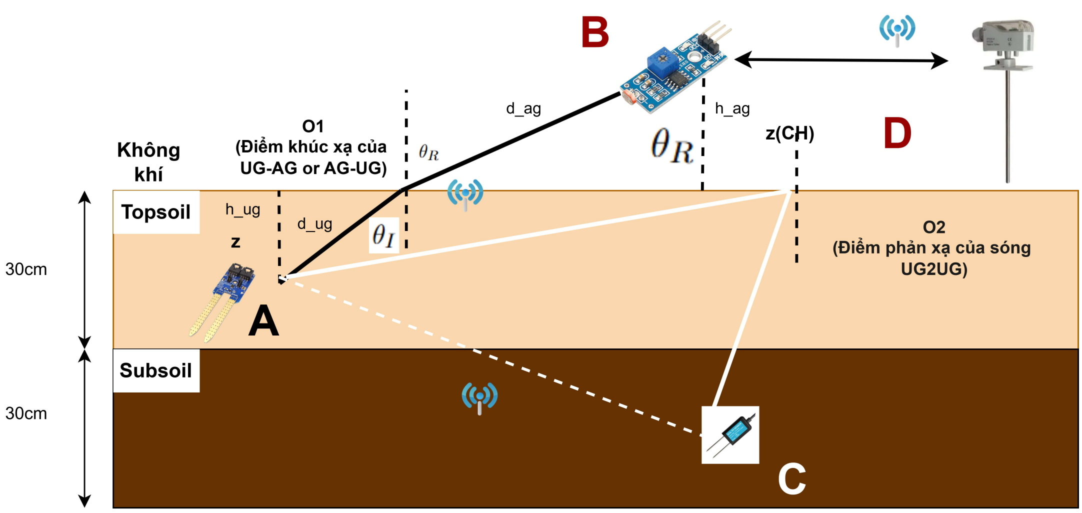

Figure 1.2: Communication channels in wireless underground sensor networks

are:

- UG2UG (underground → underground): Two underground nodes transmit signals through soil layers (from A to C).

- UG2AG (underground → above ground): An underground node sends data to a surface node/station; the signal must pass through the soil and then propagate through the air (A → B), requiring high power and experiencing significant attenuation.

- AG2UG (above ground → underground): A surface node/station transmits to an underground node (B → A), which is the reverse of UG2AG.

These methods reflect the complexity of the WUSN communication channel and serve as the foundation for selecting communication solutions, antennas, and energy management strategies within the network.

1.2.2 Multi-domain Wireless Underground Sensor Networks

WUSNs operate entirely underground, eliminating the need for cables, thus overcoming the limitations of older systems that required cables to be brought from underground sensors to the surface. In a multi-domain architecture, the network is divided into multiple domains (sub-networks). Data destined for the base station (BS) must be routed across domains: each domain has its own router, with the intra-domain protocol determining paths within the domain, while the inter-domain protocol exchanges information between routers to link the domains.

Current research primarily focuses on measuring pathloss attenuation and optimizing hardware, with less consideration given to network layer parameters and the relationship between energy consumption [5].

Practical Applications:

- Smart cities – pressure sensors placed in gas/water pipelines detect leaks; coastal sensors measure vibrations and soil temperature for natural disaster warning.

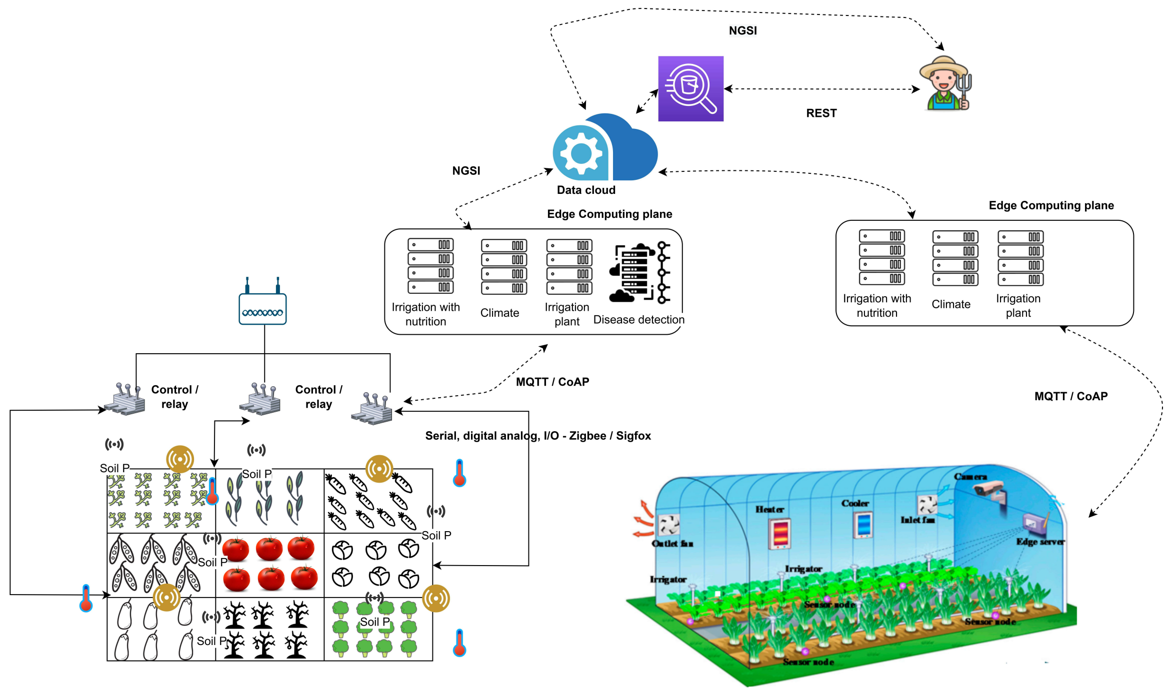

Figure 1.3: Multi-domain Applications for Agricultural Models

- Precision agriculture – each crop is equipped with a specialized sensor monitoring soil nutrition, pH, temperature, humidity (Figure 1.3) [6].

1.3 Routing Protocols

Routing helps routers select paths for packets by referencing destination addresses in the routing table. There are two methods:

- Static routing – administrators manually input fixed routes.
- Dynamic routing – routers autonomously exchange information, including:
  - Distance vector: periodically sends routing tables to neighbors; provides less detailed information, high update bandwidth.
  - Link-state: broadcasts LSA so each router has a complete network view, runs the "shortest path" algorithm; converges quickly, supports CIDR/VLSM but requires substantial CPU and memory resources.

This dissertation will focus on studying dynamic link-state routing, where routers exchange LSA messages for each convergence round to optimize paths.

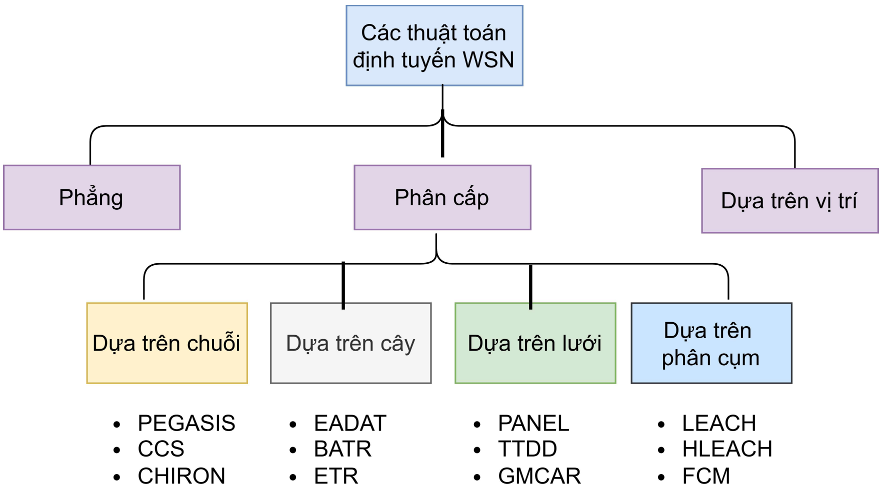

Figure 1.4: Overview of common routing protocols

1.3.1 Common Routing Protocols

In WSNs, node density can reach up to 20 nodes/m^3, meaning the network can contain hundreds or even thousands of nodes; therefore, a sustainable routing architecture is necessary [7]. Figure 1.4 summarizes three main algorithm families aimed at reducing path lengths, minimizing energy consumption, and prolonging network lifetime [8]:

- Flat structure: all nodes are of equal rank, multihop transmission; nodes close to the base station (BS) act as relay nodes.
- Hierarchical structure: nodes are clustered, with cluster-heads aggregating and forwarding data. Chain-based (PEGASIS, CCS, CHIRON), tree-based (EADAT, BATR, ETR), and PANEL.
- Geographical structure: path selection based on geographical coordinates.

In the above protocols, this dissertation uses FCM, LEACH, and PEGASIS algorithms to compare with the proposed algorithm in terms of performance.

1.3.2 Overview of Routing Solutions in Wireless Underground Sensor Networks

Most current WUSN research prioritizes energy loss modeling, with less focus on routing optimization. Some groups experiment with LoRaWAN for data collection but still focus more on transmission loss modeling rather than network performance [9–11]. Nguyễn Thị Tâm proposed selecting the number of relay nodes to extend network lifetime [12], while Panda expanded the path loss model for various environments, but lacked routing mechanisms [13]. Yao used Otsu clustering for energy-saving, but only based on surface images, not applicable to underground sensors [14]. Other 3D studies are still limited by depth or consider the network as 2D [15]; the FCM upgrade in [16] did not account for large losses underground.

Therefore, the optimal network structure is a key factor in ensuring reliability and energy savings for WUSNs. This dissertation will propose a smart three-dimensional architecture, taking into account burial depth limits and sensor placement, prioritizing short multi-hop transmissions instead of a few long hops to reduce losses and prolong network lifetime.

1.3.3 Overview of Routing Solutions in Multi-domain Wireless Underground Sensor Networks

The Internet is a multi-domain network: Autonomous Systems (AS) exchange routing information via BGP, deciding paths based on AS Path, Local Pref attributes [8, 17]. Inside each domain, OSPF (link-state, Dijkstra algorithm) finds the shortest path but relies on topology and must ensure scalability and security. The new network architecture requires more flexible routing mechanisms than older models [18].

In multi-domain WSNs, relay nodes are added to extend coverage and connectivity. Many studies optimize the number of nodes or coverage range [19–22], but none address specific 3D routing. Studies on flow control/improvements in BGP [23, 24] improve quality but still struggle to meet the dynamic constraints of multi-domain networks.

1.4 Path Loss Model

In WSNs, attenuation and noise determine signal quality and network lifetime. This dissertation calculates energy consumption for links between surface nodes using a free-space and multipath model.

$$E_T(l,d) = \begin{cases} lE_{\text{elec}} + l\epsilon_{\text{fs}}d^2 & \text{for } d < d_0 \\ lE_{\text{elec}} + l\epsilon_{\text{mp}}d^4 & \text{for } d \geq d_0 \end{cases} \tag{1.1}$$

For links between underground nodes and surface nodes, the dissertation uses Akyildiz et al.'s modified Friis model [25]. Path loss is computed as L_{UG} (in dB) using the simplified Friis model below. The values α (1/m) and β (radian/m) depend on soil conditions. This represents the attenuation due to material absorption and phase changes.

$$L_{UG-AG}(dB) = 6.4 + 20\log(d_{ug}) + 20\log(\beta) + 8.62\alpha d_{ag} \tag{1.2}$$

$L_{\text{UG-AG}}$ only measures the loss ratio; to estimate energy in WSNs, we use the following steps in the objective functions:

- **Step 1 – dB → Power**

$$P = P_0\,10^{\frac{L_{\text{UG-AG}}}{10}}, \qquad P_0 = 1\,\text{mW} = 10^{-3}\,\text{W}[26].$$

- **Step 2 – Power → Energy**

$$E = P\,t, \qquad t = 15\,\text{minutes} = 900\,\text{seconds}[27].$$

This process follows IEC standards for converting dB-W-J, helping integrate Friis attenuation adjustments into the overall network consumption.

1.5 Simulation Environment and Evaluation Criteria

This dissertation uses MATLAB as a simulation tool to implement the proposed algorithms based on analytic solutions. The routing algorithms will be evaluated against four main criteria, summarized as follows:

1. Network lifetime: Determines when the network ceases to function due to loss of connectivity or important coverage areas; this is the primary goal to maximize.

2. Energy consumption: Measures total energy used and load balancing among nodes; energy savings help reduce battery size and increase network lifetime.

3. Delay to destination: Prioritizes selecting short paths to minimize the packet transmission time to the sink, suitable for nodes with minimal memory.

4. Packet delivery success rate: Reflects the reliability of data transmission; especially important as nodes have limited buffering capacity.

In this dissertation, two key metrics—network lifetime and energy consumption—will be used for comparison to find the optimal balance between lifetime, efficiency, and network reliability.

1.6 Chapter Conclusion

This chapter provides an overview of WUSNs, including the basic concepts and differences between intra-domain and inter-domain routing, as well as clustering routing algorithms. In particular, the dissertation analyzes the network lifetime optimization problem, a critical factor in WSNs. This chapter also introduces typical routing algorithms that will serve as benchmarks in the experimental section, while establishing the simulation framework and specific performance evaluation criteria for the algorithms proposed in subsequent chapters.

- The results of this research are published in the paper [1] of the dissertation.

Chapter 2

PROPOSED ROUTING SOLUTIONS FOR MULTI-DOMAIN WIRELESS UNDERGROUND SENSOR NETWORKS

Wireless Underground Sensor Networks (WUSNs) consist of sensor nodes deployed underground. The problem of optimizing the lifetime of WUSNs has recently attracted attention from researchers, as sensor nodes deployed underground deplete energy rapidly. This chapter delves into finding new routing solutions based on analytical modeling and proposes algorithms aimed at optimizing energy consumption and extending network lifetime. The following two tables provide an overview of the solutions implemented in this dissertation.

New Routing Solution for Single-Domain Wireless Underground Sensor Networks Based on Fuzzy Clustering
Monitored area: 1 region and sensor set
Key Contributions - Developed a model to compute energy consumption across the network, considering underground transmission. - Developed a routing algorithm based on the proposed model. - Built a sensor location database with underground characteristics and implemented experiments in MATLAB to validate the accuracy of the proposed algorithms.

New Routing Solution for Multi-Domain Wireless Underground Sensor Networks Based on Sensor Metrics

Monitored area: k observation regions, k sensor sets, and p relay nodes to connect the regions

Key Contributions
- Proposed a new metric that considers not only distance and remaining energy but also the impact on neighboring nodes and sensor types.
- Developed a new network model for inter-domain and intra-domain routing based on relay nodes.
- Created algorithms to build routing matrices with the new metric and update the entire network after routing.
- Built datasets for sensor locations with various distributions, representing the depth or height of sensor nodes.
- Conducted experiments with different scenarios to demonstrate the accuracy of the network model.

2.1 New Routing Solution for Single-Domain Wireless Underground Sensor Networks Based on Fuzzy Clustering

For a WUSN in three-dimensional space, with M sensor nodes distributed evenly and scattered above and below the ground to collect information, the sensors are fixed in position and uniformly distributed in terms of configuration and function. The sensors are placed within each other's coverage range. The following assumptions are made, as illustrated in Figure 2.1:

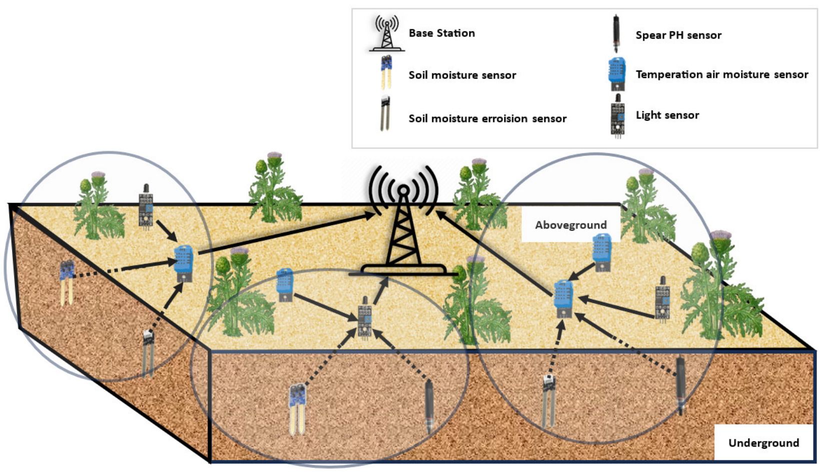

Figure 2.1: Overview of the proposed network model assumptions

- The energy of the sensor nodes is initialized equally, even though in the model, the sensors are used for different purposes.

- M sensor nodes are assigned to a specific cluster C. The number of nodes above and below the ground is divided based on the ratio of datasets. The number of members in clusters is unequal.
- All CH members are placed on the ground, and a cluster is formed that includes nodes both on and below the ground. Then, CHs will collect information from both underground and above-ground nodes and transmit it to the BS. Transmission schedules will be established at a fixed time.
- In a cluster, there will be k underground nodes responsible for collecting information from underground to the CHs above the ground.
- Single-hop communication is used as the transmission step in this model. This means that cluster members connect directly to the CH sensors, and CHs forward data to the BS within communication range.
- The BS will be fixed at the central position of the monitored area.
- After each round, all nodes will calculate their remaining energy for the BS to review and update the nodes in the network for the next rounds based on the remaining energy. From there, sensor node assignment to clusters and cluster-head selection will be recalculated.

2.1.1 Mathematical Model

This dissertation proposes a routing model for WUSN based on fuzzy clustering. Unlike crisp clustering, where each node belongs to a single group, fuzzy clustering allows a node to participate in multiple groups with different degrees of membership, reflecting uncertainty and overlapping data. This approach is particularly suitable for multi-source, multi-characteristic data in underground sensor networks. The FCM algorithm of Bezdek [28] has shown superior quality in clustering regions where clusters overlap, providing a foundation for the new solution design in this dissertation.

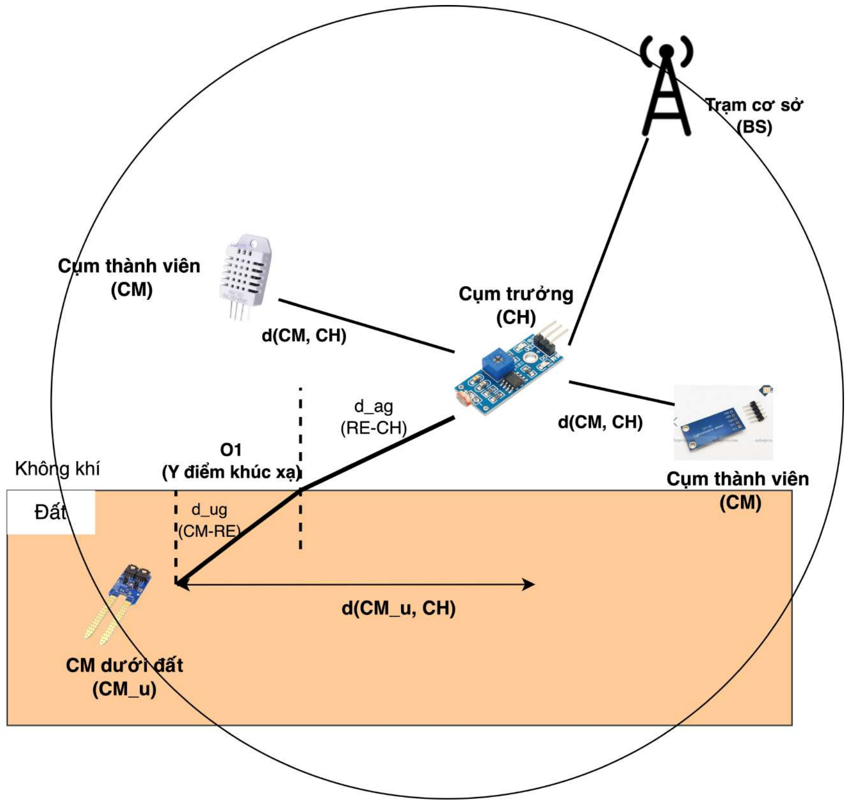

Figure 2.2: Modeling cluster members

Previous studies have only considered surface energy, neglecting underground loss. This dissertation adds this loss component, calculated based on the modified Friis model for sensor links underground to cluster heads [29], incorporated into the routing objective function. Combining fuzzy clustering with the loss model, the solution optimizes the WUSN network in real-world conditions.

$$E_{CM_u-CH} = 6 + 20\log\|X_i - V_j\| + 20\log\beta + 8.69\alpha\|X_i - V_j\| \tag{2.1}$$

Definition [30] : The total energy consumption model for the network, as assumed in the initial network hypothesis.

$$E_{total} = C(E_{CH-BS}) + (M-C)(E_{CM-CH}) + k(E_{CM_u-CH}) \tag{2.2}$$

Noise information and energy consumption are two parts of the linear function, so we can rewrite formula (2.2) as:

$$\begin{aligned} E_{total} &= C(lE_{DA} + l\epsilon_{mp} \sum_{j=1}^{C} \|V_j - X_{BS}\|^4 \ \ ) \\ &+(M-C)(lE_{elec} + l\epsilon_{fs} \sum_{i=1}^{M}\sum_{j=1}^{C} \|X_i - V_j\|^2 \ \ ) \\ &+k(6 + 20\log\|X_i - V_j\| + 20\log\beta + 8.69\alpha\|X_i - V_j\|) \end{aligned} \tag{2.3}$$

This dissertation adds the loss component for underground nodes (from coefficient k onwards) into the total network energy calculation. The new model separately considers losses when the signal passes through soil and air, helping to more accurately estimate the loss energy of underground nodes and improving the overall network efficiency.

The dissertation focuses on minimizing energy consumption across the entire network, from which the objective function of the model is constructed as follows:

$$\begin{aligned} J &= C(\sum_{i=1}^{M}\sum_{j=1}^{C} \|V_j - X_{BS}\|^4 \ \ ) \\ &+(M-C)(\sum_{i=1}^{M}\sum_{j=1}^{C} \|X_i - V_j\|^2 \ \ ) \\ &+k\left(6 - \frac{20}{\ln 10} + 20\log\beta + (8.69\alpha + \frac{20}{\ln 10})\sum_{i=1}^{M}\sum_{j=1}^{C}\|X_{i_{ug}} - V_j\| \ \ \right) \longrightarrow Min \end{aligned} \tag{2.4}$$

The fuzzification of the total energy function (2.4) gives a new objective function.

$$\begin{aligned} J_{(u_{ij},V)} &= C(\sum_{i=1}^{M}\sum_{j=1}^{C} u_{ij}^m \|V_j - X_{BS}\|^4 \ \ ) \\ &+(M-C)(\sum_{i=1}^{M}\sum_{j=1}^{C} u_{ij}^m \|X_i - V_j\|^2 \ \ ) \\ &+k\left(6 - \frac{20}{\ln 10} + 20\log\beta + (8.69\alpha + \frac{20}{\ln 10})\sum_{i=1}^{M}\sum_{j=1}^{C} u_{ij}^m \|X_{i_{ug}} - V_j\| \ \ \right) \longrightarrow Min \end{aligned} \tag{2.5}$$

The routing problem's objective function is represented by (2.5), where the membership degree u_{ij} and the cluster center V_j are two important factors that need to be assigned for each cluster in this scenario. Specifically, u_{ij} represents the membership degree of sensor i to cluster j in the fuzzy clustering model, indicating how much sensor i belongs to cluster j. Meanwhile, V_j is the center of cluster j, where data from the sensors in the cluster will be gathered and processed. The main goal of the problem is to distribute the sensors between clusters, then choose the cluster head (CH). Data will be transmitted from CM to CH and from CH to BS directly. After each iteration, the BS will evaluate the remaining energy of the sensors in the network and, if necessary, reassign clusters or reselect the CH. To ensure the feasibility of the model, several communication constraint conditions are included, such as the following constraints:

$$\sum_{i=1}^{M}\sum_{j=1}^{C} u_{ij}^{m} = 1 \quad u_{ij} \in [0,1]; \quad \forall i = \overline{1, M} \tag{2.6}$$

$$\|X_i - V_j\| \quad \leq 2Tr \tag{2.7}$$

$$\|V_j - X_{BS}\| \quad \leq 9Tr \tag{2.8}$$

This nonlinear optimization problem places connection constraints: sensor to CH not more than 2Tr; CH to BS not more than 9Tr (BS has a radius of approximately 8 Tr). Each sensor belongs to only one cluster, ensuring that the entire network is always connected. When analyzing the objective function, the dissertation derives two explicit solutions for u_{ij}, the membership degree of each node to the clusters, and V_j, the centers of the clusters chosen as the cluster heads, as demonstrated in the full text.

$$\begin{aligned} V_j = & \frac{\frac{B}{A}}{\sqrt[3]{\frac{B(X_{BS}-X_i)+C}{2A} + \sqrt{\frac{1}{4}\left(\frac{B(X_{BS}-X_i)+C}{A}\right)^2 + \frac{B^3}{27A^3}}}} \\ & - \sqrt[3]{\frac{B\left(X_{BS}-X_i\right)+C}{2A} + \sqrt{\frac{1}{4}\left(\frac{B\left(X_{BS}-X_i\right)+C}{A}\right)^2 + \frac{B^3}{27A^3}}} + X_{BS} \end{aligned} \tag{2.9}$$

$$u_{ij} = \frac{1}{\sum_{i=1}^{C}\left(\frac{\epsilon_{mp}\|V_j - X_{BS}\|^4 + (M-C)\epsilon_{fs}\|X_i - V_j\|^2 + k\left(\frac{20}{\ln 10} + 8.69\alpha\right)\|X_{i_{ug}} - V_j\|}{\epsilon_{mp}\|V_j - X_{BS}\|^4 + (M-C)\epsilon_{fs}\|X_i - V_j\|^2 + k\left(\frac{20}{\ln 10} + 8.69\alpha\right)\|X_{i_{ug}} - V_j\|}\right)^{\frac{1}{m-1}}} \tag{2.10}$$

2.1.2 FCM-WUSN Algorithm

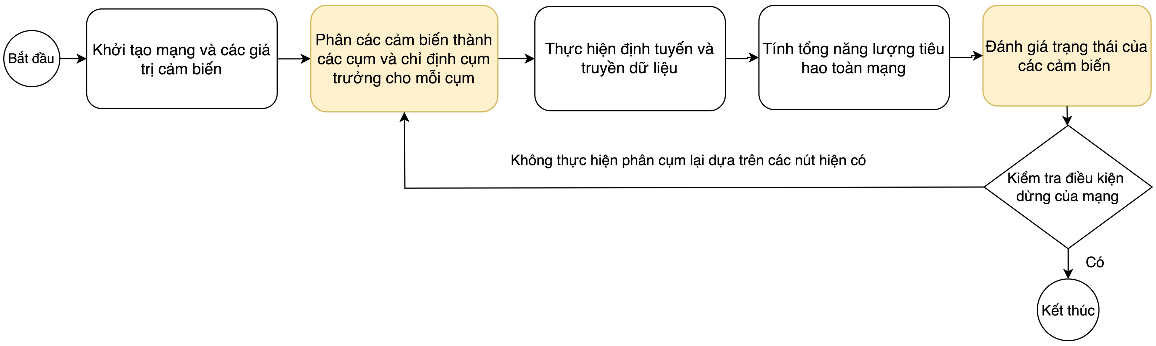

Figure 2.3: Overview of the routing process in the network

The algorithm initializes the positions, energy, signal levels, and sampling frequency for each sensor, as well as defining the coverage area. Next, the nodes are clustered, and cluster heads are selected. After each network update, the algorithm re-clusters based on the solutions v_j and u_{ij}.

Algorithm 1 FCM-WUSN Fuzzy Clustering Algorithm

Input: Number of sensors M; Number of clusters C; fuzzy parameter m; set of sensors $X = [X_1, X_2, ..., X_M]$

Output: Membership matrix μ, and cluster center matrix V

$t \leftarrow 1$
$j \leftarrow 1$
Initialize membership matrix u_{ij} satisfying constraint (2.6)
while $\|u^t - u^{t-1}\| > \epsilon$ **do**
    $t \leftarrow t + 1$
    Compute cluster center vector V using formula (2.9)
    Update membership matrix u using formula (2.10)
end while

Assign clusters to each node:
Get index j such that $U_{ij} = \max(U_{i1}, U_{i2}, ..., U_{iC})$.
The cluster that node i belongs to is cluster j.

While loop for additional conditions to assign cluster heads:
while $j < C$ **do**
    if $(d(X_i, V_j) = min(\forall X_i \in C_j)\ \ and\ X_j\ on\ surface)$ **then**
        if $d(X_i, BS) < 9Tr$ **then**
            $X_i(CH) \leftarrow true$ – Assign as cluster head at position X_i
        end if
    end if
    if $d(X_i, X_j(CH)) < 2Tr$ **then**
        $X_i(non - CH) \leftarrow true$
    end if
end while

The FCM-WUSN algorithm improves FCM by calculating fuzzy membership for each node based on both distance and energy loss. The cluster centers are established, and the cluster head is selected from the surface nodes within the BS coverage. This method ensures homogeneous clustering, optimizes resources, and reduces energy consumption, making the WUSN network more stable.

Algorithm 2 Routing Algorithm in the Network

Input: Number of sensors M; sensor set $X = [X_1, X_2, ..., X_M]$; threshold for total energy $thresTotalEC$; threshold for stopping rounds $thresStopRounds$; BS position BS

Output: flagStopNetwork

```
flagStopNetwork ← false
totalEC ← 0
deadNodes ← 0
repeatRounds ← 0
while flagStopNetwork = false do
    for i ← 1 to M do
        for j ← 1 to C do
            if d(X_i, X_j(CH) > 2Tr then
                Find connection with a nearby node within 2Tr, and mark the connection.
            else
                Mark the nodes X_i and X_j(CH)
            end if

        end for
    end for
    Compute totalEC based on formula (2.3)
    if (totalEC <= thresTotalEC or deadNodes >= thresDeadNode) and repeatRounds = thresStopRounds then
        flagStopNetwork ← true
    else
        if totalEC <= thresTotalEC or deadNodes >= thresDeadNode then
            repeatRound ← repeatRound + 1
        end if
        Update energy and state of X_i.
    end if
end while
```

The proposed clustering-based routing protocol in this study enables efficient routing in WUSNs by routing the packets collected by the CMs to the CH through cluster member nodes or directly. The CHs forward directly to the BS, maintaining the route for efficient data routing. The routing algorithm describes how the database is obtained from CMs to BS.

2.1.3 Algorithm Complexity

Algorithm 1: Fuzzy Clustering FCM-WUSN The complexity of the algorithm to form clusters based on the proposed FCM-WUSN algorithm, with N data points and C clusters. Computing the centers $V : O(N \times C)$ for each iteration, updating membership values $u : O(N \times C^2)$. Since the FCM-WUSN algorithm repeats these steps until convergence (with T), the overall complexity of the FCM-WUSN algorithm is $O(T \times N \times C^2)$. If the number of clusters C and iterations T are small, FCM-WUSN can work efficiently for small to medium-sized clustering problems. However, for large datasets and a large number of clusters, the computational complexity may become an issue. Therefore, after routing, at the end of each round, the algorithm updates the states of the nodes in the network, potentially reducing FCM-WUSN's complexity by filtering out low-energy nodes, thus making the algorithm more efficient in large sensor networks.

Algorithm 2: Routing in the Network Algorithm 2 contains the following main loops:

- The `While` loop (Lines 5-15): This loop continues until `flagStopNk` is set to `true`. The loop can run up to `thresStopRounds` times.
- The first `For` loop (Lines 7-10): This loop iterates through all sensors from 1 to `M`, meaning this loop runs `M` times.
- The second `For` loop (Lines 8-10): This loop iterates through the child sensors of each sensor from 1 to `C`. Thus, this loop runs `C` times for each sensor in the outer loop.

Each iteration of the second `For` loop, the operations inside (checking conditions, finding connections, marking connections, computing total energy `totalEC`, and checking conditions) takes `O(1)` time. Thus, the complexity for each iteration in the `While` loop is $O(M \times C)$. If the `While` loop runs for a maximum of `thresStopRounds` times, the total complexity of the algorithm will be:

$$O(\text{thresStopRounds} \times M \times C) \tag{2.11}$$

2.2 New Routing Solution for Multi-Domain Wireless Underground Sensor Networks Based on Sensor Metrics

The fuzzy clustering-based routing algorithm saves energy in single-domain networks, but it can lead to congestion and rapid battery depletion at the cluster head (CH). In WUSNs, underground nodes outside the CH's coverage lose

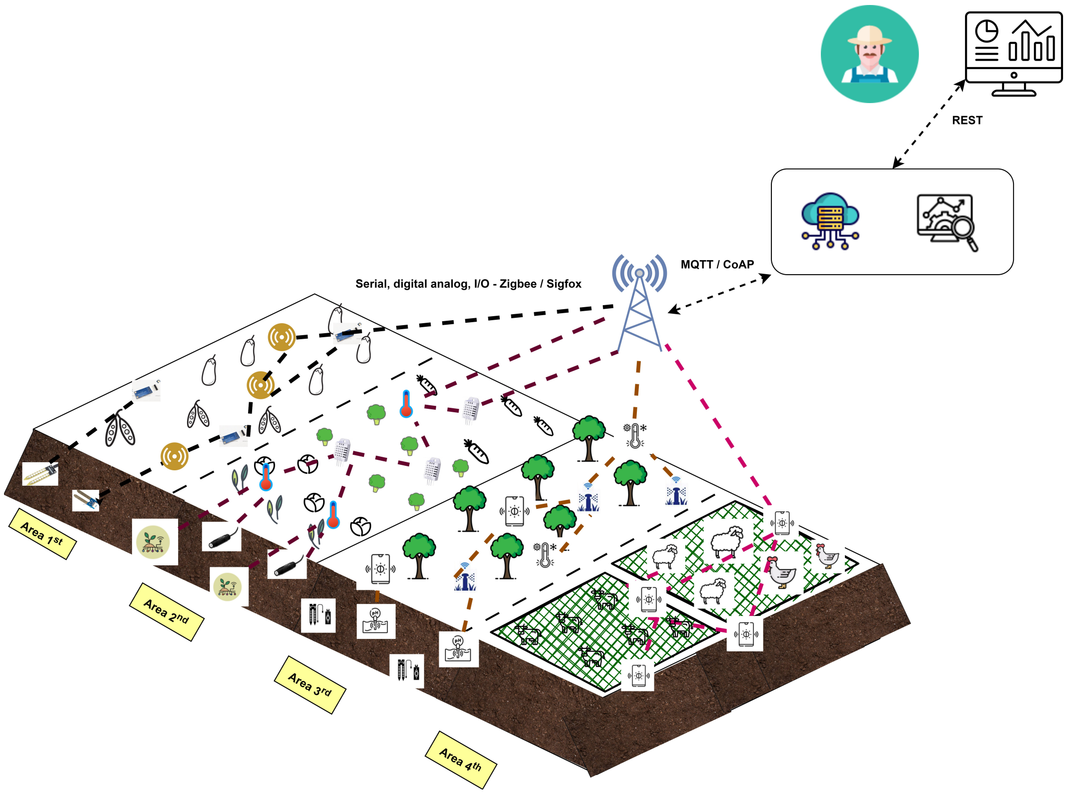

Figure 2.4: General concept of the multi-domain transmission model

communication, causing energy loss to increase due to the soil environment. Therefore, there is a need to transition to multi-domain routing. This dissertation proposes a new metric to form a flexible multi-domain structure, enabling faster scanning of all observation domains, leveraging underground transmission characteristics, and extending the lifetime of large-scale WUSN networks.

For a wireless sensor network in three-dimensional space with k observation areas, each containing M sensor nodes responsible for collecting data both above and below the ground, the sensors are uniformly distributed both above and below the ground to efficiently gather information. The sensors are placed close to each other and within each other's coverage radius. Several assumptions are made, as illustrated in Figure 2.4:

1. A WUSN consists of k regions, each containing M sensor nodes as shown in Figure 2.5. A set of sensor nodes in all regions is considered: $S^k = \{S_1^k, \ldots, S_M^k\}$.

2. There are l relay nodes functioning as intermediary nodes to extend coverage. $R = \{R_1, \ldots, R_l\}$ for the entire network. All relay nodes are placed

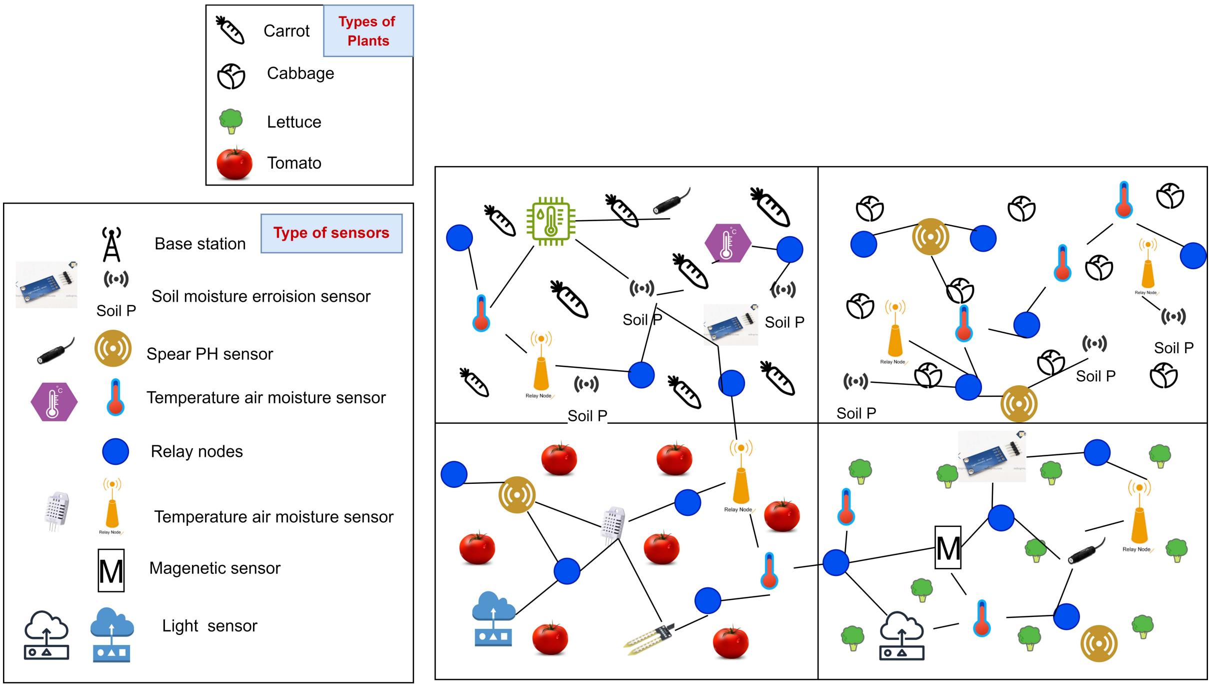

Figure 2.5: Hypothetical simulation for multiple observation domains

on the ground. The relay nodes have large energy and can communicate inter-region.

3. The sensor nodes in a region have identical configurations regarding initial energy and communication radius. However, for each different region, the sensor node configurations vary.

4. Multi-hop communication is used for transmission in this model.

2.2.1 Constructing a New Sensor Metric

For the multi-domain WUSN, this dissertation proposes a new index to evaluate the shortest path in the wireless sensor network, called EDTNR, consisting of five criteria that make up this index. This new index is designed to comprehensively evaluate the efficiency of communication paths by integrating several important factors and considering the signal transmission capability of sensor nodes buried underground in a large space with various types of sensors. Typically, each sensor node can be evaluated based on two main factors: distance and remaining energy. In this model, each sensor node in the network is considered a vector with the following values, where k is the index of the observation region:

$$^{[k]}v_j^{(i)} = \left\{ ^{[k]}v_j^{(1)}, ^{[k]}v_j^{(2)}, ^{[k]}v_j^{(3)}, ^{[k]}v_j^{(4)}, ^{[k]}v_j^{(5)} \right\} \tag{2.12}$$

1. **Remaining energy (E)** $^{[k]}v_j^{(1)}$ represents the remaining energy of the sensor v_j in region k, evaluating the power of sensor nodes along the path. This ensures that the selected path is energy-efficient, extending the network's operational lifetime. The unit is J.

2. **Sensor location (D)** $^{[k]}v_j^{(2)}(x, y, z)$ represents the location of the sensor v_j. From this information, the distance between two nodes can be measured to form a path.

3. **Sensor type (T)** $^{[k]}v_j^{(3)}$ represents the type of sensor. Typically, this will be the same for k regions. A sensor can transmit data to another sensor of the same type.

4. **Communication influence of neighbors (N)** $^{[k]}v_j^{(4)} : N(^{[k]}v_j^{(i)})$ represents the neighbors of sensor v_j. It reflects the influence of the node in the network based on the number of intersections with neighboring nodes.

5. **Sensor radius (R)** $^{[k]}v_j^{(5)}$ is the communication radius value of sensor v_j. The communication radius for different regions may vary. It can be considered as an estimate of the strength and reliability of communication links between regions.

Evaluating the Correlation Between Two Sensors in the New Metric

EDTNR can be represented as a vector in a multidimensional space corresponding to various components: $EDTNR = (E, D, T, N, R)$. The explanation for each criterion of the correlation between two sensors is as follows:

1. A large difference in remaining energy between two nodes leads to low correlation between the two sensor nodes. A small energy difference between two nodes results in higher correlation.

2. A short distance between two sensors helps strengthen the signal and transmit data better, thereby reducing latency and saving energy. Conversely, if the distance is too far and beyond the connection range, the signal weakens and data cannot be transmitted, leading to low correlation between the two nodes.

3. The relationship between two nodes is based on the type of region. If they belong to the same region, the relationship value is 1. If they belong to different regions, they cannot communicate with each other. This will be handled similarly between region types and the range of sensor nodes.

4. The more intersections of sensor nodes in neighboring regions, the lower the correlation. If the correlation between two sensors is high, they can be selected as the next destination for data transmission. However, if there are too many nodes in the neighborhoods of the two sensors, it creates loops, and the information will circulate within these nodes. These loops cannot break, preventing finding new paths to forward data to the base station.

2.2.2 Evaluating the New Metric

This dissertation proposes the Flexmetric (F) metric based on Euclidean distance. F satisfies the four criteria of a metric and directly influences energy consumption, transmission range, and coverage radius. At the same time, F must demonstrate the correlation between neighboring nodes, ensuring accurate reflection of the connectivity and interaction of sensors in the region.

$$e^{-\left\|N_{[k]}v_j^{(4)} \wedge N_{[k]}v_l\right\|} + e^{-\left\|N_{[k]}v_j^{(4)} \wedge N_{[k]}v_t^{(4)}\right\|} > e^{-\left\|N_{[k]}v_t^{(4)} \wedge N_{[k]}v_l^{(4)}\right\|} \tag{2.13}$$

With the condition $x > y,\ y > z,\ x > z$, these assumptions are applied to these constraints.

$$\begin{cases} |y| < |x| + |y| - 2|y| \Leftrightarrow 2|y| < |x| \\ |z| < |y| + |z| - 2|z| \Leftrightarrow 2|z| < |y| \\ |z| < |z| + |x| - 2|z| \Leftrightarrow 2|z| < |x| \end{cases}$$

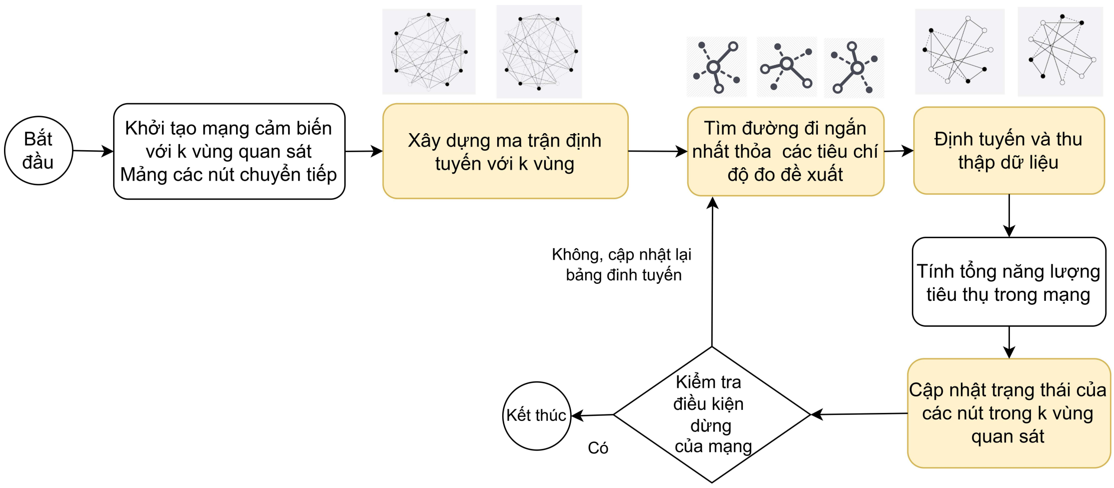

Figure 2.6: Flowchart of the WUSN routing model for multiple observation domains

2.2.3 Multi-Graphs WUSN Algorithm

The flowchart in Figure 2.6 describes the operation of a sensor network system, where sensor nodes and observation domains are continuously connected and monitored. The system constantly checks the node status; if a node is not performing well, the route is reconfigured, and energy consumption is computed and updated to maintain optimal performance. In the intra-domain layer, the sensor sends data to the relay node for processing or compression; in the inter-domain layer, the relay nodes transmit data through multiple hops to the base station. This two-layer structure shortens transmission distances, saves energy, and enhances network scalability.

Algorithm 3 Forming the Routing Matrix

Input :k observation domains; number of sensors in k regions $M = [M_1, ..., M_k]$; p number of relay nodes in the region; list of relay nodes $RN = [RN_1, RN_2, ..., RN_p]$; List of sensor sets in the k^{th} region $X = [X_1, X_2, ..., X_M]$ where $X1 = [X_1^1, X_2^1, ..., X_M^1]$

Output : Routing matrix

$routingMatrix \leftarrow containers.Map()$
while $t \ < regions$ **do**
  while $i \ < X_t$ **do**
    $j \leftarrow i + 1$
    while $j \ < X_t$ **do**
      if distance $(X_t^i, X_t^j) \leq 2radius(X_t^i)$ **then**
        Set X_t^j in the routingMatrix of X_t^i
      end if
    end while
    if distance $(X_t^i, BS) \leq 3radius(X_t^i)$ **then**
      Mark X_t^i as able to connect to BS
    end if
  end while
end while

The connection matrix is created initially to record the relationships between sensors, relay nodes, and all possible links within each region. Since the nodes are fixed, the matrix needs to be created only once, and links that cannot be established initially will remain unchanged. This allows any subsequent tracing or optimization of transmission paths to simply refer to the matrix, eliminating the need for reconfiguration.

Marking the nodes that can transmit directly to the base station ensures fast and stable data transmission. The connection information stored in the matrix helps create simple routing and convenient look-up. When testing or improving, the matrix reference minimizes the complexity of operations.

Algorithm 4 Find the Shortest Path for a Node

Input : Index of the node under consideration, routing matrix; array of relay nodes connecting to BS
Output : Shortest path from index-node

```
touchBS ← false
Initialize shortestPath(1, 1) ← index − node
Retrieve the set of connected nodes X_t from the routing matrix connectedNodes
index ← 1
while touchBS = false do
    Initialize numConIntersect ← maxNumInter
    Initialize next position numberIndex ← maxIndex
    Loop to find the next node with favorable connection to forward energy
    for node i in connectedNodes do
        if node i is not in shortestPath then
            intersectSet1 = getConnectTableNodes(i)
            intersectSet2 = getConnectableNodes(lastNodeInShortestPath)
            intersect = intersectSet1 ∩ intersectSet2
            if intersect ≤ numConIntersect then
                numConIntersect ← intersect
                numberIndex ← X_t^i
            else
                if numberIndex ≠ maxIndex & C = numConIntersect then
                    Nti ← size(X_t^i)
                    Nte ← size(X_t^numberIndex)
                    if Nte > Nti then
                        numberIndex ← X_t^i
                    end if
                end if
            end if
        end if
    end for
    index ← index + 1
    Check if the current path reaches the BS
    shortestPath(1, index) ← numberIndex
    if X_t^numberIndex can connect to BS then
        touchBS ← true
    else
        if numberIndex ≠ maxIndex then
            X_t ← routing − matrix_numberIndex
        else
            touchBS ← true
        end if
    end if
end while
```

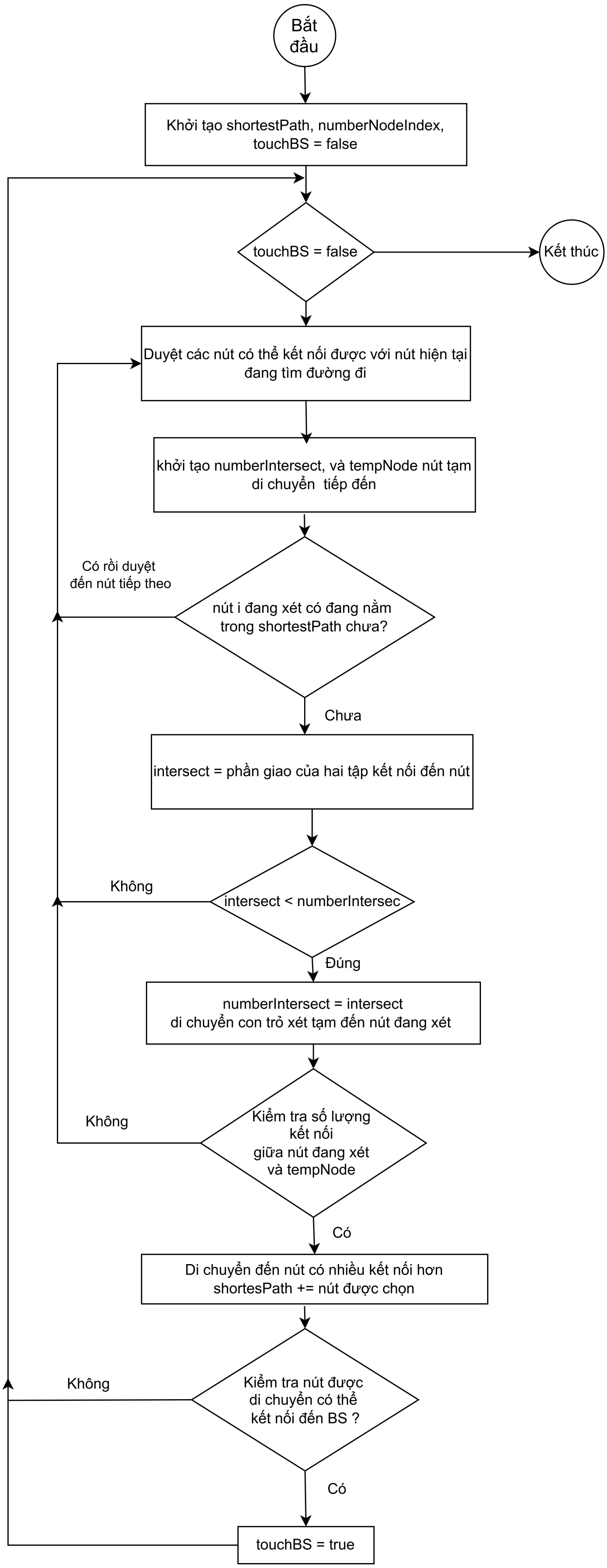

Figure 2.7: Flowchart of the shortest path selection process for a node

The process of finding the shortest path between nodes in a network is a key part of routing algorithms, particularly in communication networks or wireless sensor networks. Once the algorithm determines the shortest path from one node to others, the next step is to perform routing across the entire network, ensuring that data can be transmitted efficiently from source to destination.

Algorithm 5 General Routing Algorithm in the Network

Input :k observation regions; number of sensors in k regions $M = [M_1, ..., M_k]$ p number of relay nodes in the region; list of relay nodes $RN = [RN_1, RN_2, ..., RN_p]$; List of sensor sets in the k^{th} region $X = [X_1, X_2, ..., X_M]$ where $X1 = [X_1^1, X_2^1, ..., X_M^1]$; initial energy for the entire network; $\alpha = 0.03, \beta = 0.1$: parameters for assumptions in the soil

Output : flagStopNetwork

$routingMatrix \leftarrow containers.Map()$
while $flagStopNetwork = false$ **do**
  while $t < regions$ **do**
    while $i < X_t$ **do**
      $shortestPath_t^i \leftarrow Algorithm\ 2$
      if node under consideration is connected to BS **then**
        while $j \leq shortestPath_t^i.size() - 1$ **do**
          $next \leftarrow j + 1$
          Compute energy for transmission - based on energy calculation defined by the proposed metric
          if $next = shortestPath_t^i.size()$ **then**
            Mark node X_t^j as connected to node X_t^i in the routing matrix
          end if
        end while
      end if
    end while
    Compute total energy consumption $totalEC$ and number of connected nodes $connectedNodes$
    if $totalEC \leq (\alpha initEnergy X_t.size()) \mid (X_t.size() - connectedNodes) \leq (\beta X_t.size())$ **then**
      $flagStopNetwork \leftarrow true$
    else
      Update states based on the algorithm
    end if
  end while
end while

This research focuses on optimizing routing in multi-graph WUSNs by selecting each node in the routing matrix and determining the shortest path from these nodes to the BS. The selection of the next node for transmission is influenced by factors such as sensor type, distance, energy level, and the impact of neighboring nodes within the set. The flowchart for the proposed routing algorithm is detailed in Algorithm 4. Relay nodes are considered capable of transmitting across multiple domains, serving as intermediaries for nodes or domains located far from the BS. The routing algorithm computes and adjusts routes to optimize network resources, including updating the network's energy levels.

2.2.4 Algorithm Complexity

Routing Matrix Construction Algorithm: The algorithm will iterate over k regions, and for each sensor node, it will check the other sensors, giving a complexity of $\frac{X_t(X_t-1)}{2}$. Thus, with two nested loops, the overall complexity will be $O(regions \times X_t^2)$.

Shortest Path Finding Algorithm: The complexity of finding the shortest path from one node to the BS is $O(n \times t)$ for t executions, where the outer loop can iterate up to n times. t is the number of nodes in $connectedNodes$, and n is the number of steps required to reach the BS.

General Routing Algorithm: The algorithm runs until the $flagStopNetwork$ variable is set to true. Within the main loop, this will execute for a number of iterations equal to the number of observation regions $regions$, giving a complexity of $O(regions)$ at line 3. At line 4, it calculates the number of nodes in each region X_t. Then, the algorithm checks the nodes and finds the shortest path to the BS, with complexity depending on Algorithm 4. In line 8, the algorithm computes energy for transmission, giving a complexity of $O(k')$, where $k' = shortestPath_t^i.size()$ is the number of steps in the path. Thus, the overall complexity of this algorithm is described by:

$$O(\text{regions} \cdot X_t \cdot k \cdot n \cdot t + \text{regions} \cdot M \cdot z)$$

Routing Matrix Update Algorithm: The complexity of updating the routing matrix is $O(regions \times M \times z)$, where M is the number of nodes in each region, and z is the number of relay nodes connected to the nodes.

2.3 Chapter Conclusion

Chapter 2 introduced two WUSN routing models, developed using Taylor expansions and validated by the Cauchy inequality to prove stability, convergence, and energy efficiency. The multi-region model collects and merges data through short hops to relay nodes or cluster heads, while the inter-region model transmits multi-hop data between regions to the base station, with all path costs stored in the routing matrix. The fuzzy clustering algorithm optimizes routing for each observation domain, while a new metric helps quickly build routing tables for multi-domain structures. The results of this research are published in papers [2, 3] of the dissertation and one paper currently under review [5].

Chapter 3

SIMULATION RESULTS

3.1 Environment Setup

This chapter describes the testing and evaluation of WUSN routing models on a self-constructed node coordinate dataset to verify the effectiveness and reliability of the algorithm. The author analyzes datasets and simulation environments from previous studies, highlighting the strengths and weaknesses, accuracy, reproducibility of underground conditions, and practical deployment limitations.

3.1.1 Introduction to the Simulation Environment

The multi-domain underground sensor network (WUSN) consists of several clusters of buried nodes. Each domain routes data internally and then transmits it to a gateway or base station, requiring optimization of both intra-domain and inter-domain paths to reduce energy consumption and maintain stability of the underground nodes. The experimental section will: compare the clustering of the new algorithm with FCM; measure total energy consumption, network lifetime (number of rounds), number of dead nodes, and underground dead nodes; contrast FCM-WUSN with the proposed algorithm on a sample area; and test the new algorithm across various multi-region scenarios with Gamma, Poisson, Gaussian distributions, multiple gateways, and BS positions to evaluate inter-domain capability.

3.1.2 Dataset Construction

To optimize routing performance in WUSNs, building a high-quality dataset is crucial. This dataset not only serves as a solid foundation for development but also acts as a key tool for evaluating effective routing algorithms. This dissertation focuses on constructing a specialized dataset, including the precise coordinates of sensor nodes deployed underground. Collecting and organizing coordinate data for underground sensor nodes presents unique challenges, especially the complexity of accurately determining the depth and height of the sensors. To address this, the dataset in this dissertation is designed to realistically simulate diverse deployment scenarios, providing a reliable testing environment for examining and comparing different routing algorithms.

Design Principles and Dataset Construction The goal is to create a dataset with sensor node positions for routing in WUSNs. This is based on the scientific foundation using the Linear Congruential Generator (LCG) and Middle Square Method (MSM). The dataset is then integrated with selection or exclusion criteria for use in routing. The design principles are outlined as follows:

- **Step 1**: Determine the coordinate set (x, y, z); x, y represent horizontal position, and z indicates whether the sensor is above or below the ground.
- **Step 2**: Generate random data
  - **LCG**: $X_{n+1} = (aX_n + c) \bmod m$ to generate the seed sequence.
  - **MSM**: Use X_n from LCG, square it, and extract the middle digits as the next value to break cycles and increase randomness.
- **Step 3**: Filter by routing criteria (coverage radius, density, distribution); exclude points outside the observation region, negative coordinates, or duplicates.

Scenario	Description
S1	**Underground Factor Analysis** Evaluate total energy consumption and the number of underground connected nodes.
S2	**Routing Performance Evaluation** Assess the effectiveness of data routing in extending network lifetime in multiple regions.
S3	**Routing Performance using Relay Nodes** Evaluate the impact of relay node repetition and base station (BS) placement conditions.
S4	**Variance Testing with Different Distributions** Demonstrate the proposed method's capability across different distributions.

3.1.3 Evaluation Scenarios

3.2 Routing Experiment Results in a Single Region

The experimental results from evaluating the WUSN energy optimization techniques are presented here. The experiments considered how the proposed solutions improve energy usage, network lifetime, and data transmission efficiency.

3.2.1 Main Experimental Results

This section presents the experimental results validating the proposed algorithms. Scenarios are provided to compare FCM and FCM-WUSN in clustering and data aggregation. The results, as shown in Figure 3.1, illustrate the total energy consumption and the number of rounds after running FCM and FCM-WUSN. We chose the average results from ten runs of datasets (100 nodes and 10% underground nodes) across six different datasets.

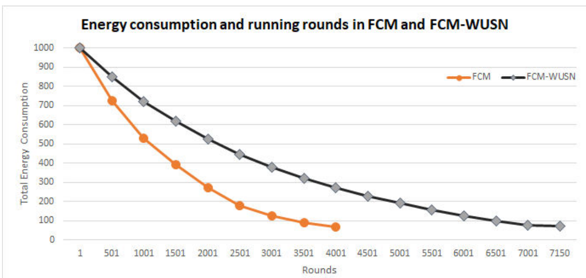

Figure 3.1: Total energy consumption of FCM and FCM-WUSN

The table above shows the average results from ten runs of each dataset, demonstrating that FCM-WUSN consistently outperforms FCM across all six datasets. Network lifetime increased significantly (2.1× with 100 nodes, 20% underground; achieving ≈ 4969 rounds with 200 nodes, 30% underground) while

Table 3.1: Network Lifetime of FCM and FCM-WUSN in Six Datasets

	100 nodes & 10% UN	100 nodes & 20% UN	100 nodes & 30% UN	200 nodes & 10% UN	200 nodes & 20% UN	200 nodes & 30% UN
Number of Rounds						
FCM	4555.5	4198.5	4669.3	5823.8	5972.1	5460
FCM-WUSN	**8085.1**	**8582.4**	**8459.2**	**11230.8**	**10817.4**	**10429.3**
Energy Consumption (J)						
FCM	**950.017**	**950.018**	**950.01**	**1900.022**	**1900.019**	1900.01
FCM-WUSN	953.214	958.41	958.41	1905.6	1912.022	**1886.1**

energy consumption only increased by $\approx 0.02\%$.

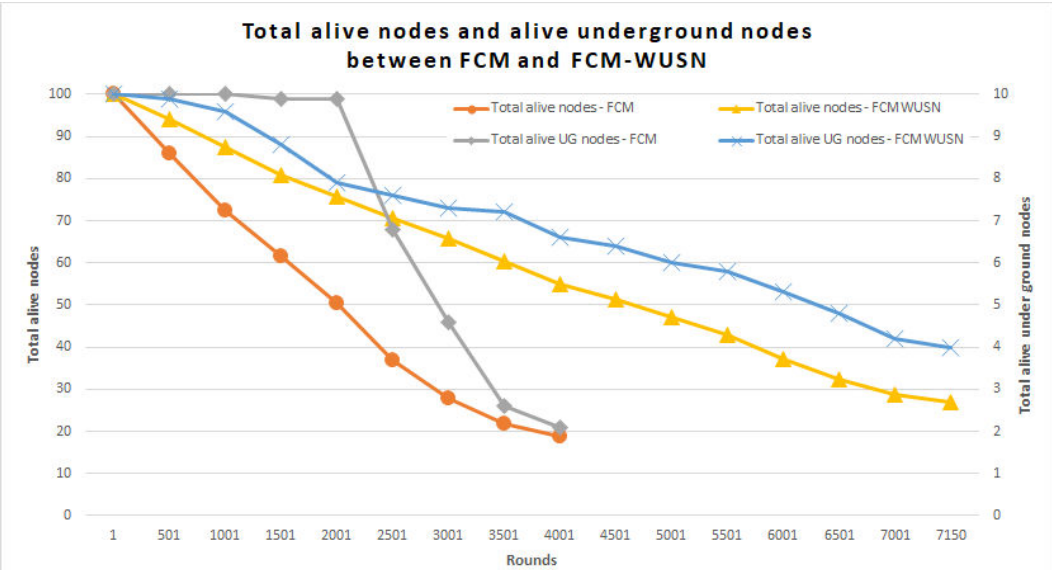

Figure 3.2: Total number of alive nodes, and underground alive nodes between FCM and FCM-WUSN

Figure 3.2 shows the total number of alive nodes and underground alive nodes between FCM and the proposed algorithm. The results conclude that the number of underground nodes in FCM remains unchanged for the first 2000 rounds, whereas in the proposed algorithm, the number of underground nodes gradually decreases from 10 to 3 by round 3174. On the other hand, in the FCM-WUSN algorithm, the number of underground nodes decreases slowly, with a reduction of one node starting from the first round until around the 1337th round. Finally, the number of underground dead nodes in the proposed algorithm remains lower than that in FCM. The clustering formation based on Algorithm 1 leads to harmony between above-ground and underground nodes. Underground nodes do not need to transmit data to distant CHs, leading to slower energy depletion.

Table 3.2: Total number of alive nodes and underground alive nodes of FCM and FCM-WUSN in six datasets

	100 nodes & 10% UN	100 nodes & 20% UN	100 nodes & 30% UN	200 nodes & 10% UN	200 nodes & 20% UN	200 nodes & 30% UN
Total number of alive nodes						
FCM	16.3	16	15.4	29.3	27.7	29.1
FCM-WUSN	**19.4**	**19.3**	**19.1**	**39.2**	**39.2**	**43**
Number of underground alive nodes						
FCM	1.9	5.8	6.5	3.5	11.1	14.8
FCM-WUSN	**2.8**	**7**	**9.8**	**6.3**	**16.6**	**23.2**

Table 3.2 shows the averages from ten runs for each dataset and technique concerning the alive nodes of the network. The findings suggest that the proposed technique outperforms FCM in both total alive nodes and underground alive nodes. Compared to FCM, the proposed method is 1.19 times more successful in evaluating a dataset with 100 nodes, with 10% underground. For a dataset of 200 nodes with 30% underground, the proposed method has about 14 more nodes than FCM.

The distribution of six datasets can affect the network's running cycles and alive nodes. When processing sensor data underground with a wide distribution, the FCM-WUSN method is 1.2 times and 1.4 times more efficient than FCM with 100 and 200 nodes, respectively. With the same number of sensor nodes but varying underground node ratios, FCM-WUSN and FCM show a difference of 4 nodes in total alive nodes, from 9.9 nodes with the data and 200-10% underground nodes to 13.9 nodes with the data and 30% underground nodes.

3.2.2 Conclusions from the Experiments

This study recommends implementing a real-world case study, such as smart agriculture, by setting the BS parameters, the number of sensors in the network, and sensor placement based on depth for underground deployment. In summary, three main findings are presented after conducting the experiments:

- Performance in terms of network lifetime and the number of alive nodes, both above and below the ground, significantly improved when using FCM-WUSN. The proposed method has a network lifetime 2.1 times greater than the standard clustering method, FCM. The FCM-WUSN algorithm has 1.4 times more alive nodes (total and underground) than FCM. This does not change much when running datasets with more sensors in the network and sensor placement based on sensor depth when deployed underground.

- After the experiments, deploying a large number of underground sensors improved the total number of alive sensors and underground nodes compared to traditional algorithms like FCM. Therefore, the underground observation area in FCM-WUSN remains larger and lasts longer than in FCM.
- The algorithm has some disadvantages. Energy usage is still inefficient. Energy consumption is about 1.0002 times higher than FCM. Signal loss due to the transmission between land and air contributes to this.

3.3 Results of Routing Experiments in Multiple Regions

3.3.1 Experimental Results for Different Observation Areas

This section discusses the results of various scenarios. The evaluation is based on the number of rounds, network durability, energy cost during data transmission from nodes to base stations, and the number of active nodes. It considers the number of observation areas, the number of sensors in each area, the base station location, and the different distributions of sensors across regions when analyzing the algorithm's characteristics.

Criteria for Network Lifetime

A network is considered to have failed when ≤ 10 % nodes are active or $\leq 5\%$ energy remains, so load balancing is critical. In Scenario 1, the Multi-Graphs-WUSN algorithm extends the network lifetime 1.4 times longer than FCM-WUSN and 3 times longer than FCM by considering underground connectivity. In Scenario 2, Multi-Graphs-WUSN consistently outperforms PEGASIS in both 2 and 4 regions; the highest difference is nearly 10,000 rounds (4 regions, 100 nodes, 10% underground). When increasing to 200 nodes and a higher underground node ratio, PEGASIS narrows the gap (from 1.259 times to 1.008

times) but still does not surpass Multi-Graphs-WUSN.

In Scenario 3, Multi-Graphs-WUSN still shows superior performance when the BS is placed in the Northern and Eastern regions, with performance percentages of 41.72% and 47.78% in terms of lifetime. However, when placed in the Western and Southern regions, PEGASIS gives better lifetime performance with percentages of 58.41% and 43.31%. Multi-Graphs-WUSN handles better when expanding from 2 regions to 4 regions, especially when placing the BS in the North (31.61%) and East (62.69%). PEGASIS shows stability when expanded, but there is no strong growth in network lifetime compared to the proposed algorithm.

Compared to the PEGASIS + HLEACH model (each algorithm handles one region, equivalent to four regions in the entire network), Multi-Graphs-WUSN is almost always ahead. Network lifetime increased significantly in Gamma (16,522 vs. 14,193 rounds) and small-scale Gaussian (13,713 vs. 9,568). For small Poisson, PEGASIS slightly outperforms by 5,401 rounds, but when the observation area is doubled, Multi-Graphs reverses the situation, becoming 2.5 times longer (9,976 vs. 3,978). Gamma still maintains an advantage; Gauss expands, and PEGASIS + HLEACH only exceeds by 842 rounds. In uniform distribution, Multi-Graphs-WUSN leads in every scale.

Criteria for Total Number of Alive Nodes and Underground Alive Nodes

In Scenario 1, the results show that the proposed algorithm improves coverage for connected nodes while considering the underground factor. All 10 nodes connected to the base station in the early rounds of a scenario with 100 nodes operating at 10%. This increases the data accuracy and the number of operational rounds. The number of remaining nodes decreases by 2 compared to the case where 30% of the nodes are underground. Expanding the scenario to include 10-30% underground nodes improves the performance of the proposed method for the remaining underground nodes. Distributing the data evenly to the file reduces unnecessary concentration on CH nodes, as seen in the FCM-WUSN scenario. As a result, more underground nodes remain.

With uniform distribution, moving the BS determines the number of alive nodes. When placed in the North or East, Multi-Graphs WUSN retains more nodes than PEGASIS by 2.4 times and 1.8 times (2 regions) or 1.7 times and 2.6 times (4 regions); placing it in the West reduces the advantage (1 node), while the South lags PEGASIS by 9 nodes. In Gamma, Gaussian, and Poisson distributions, results fluctuate: Multi-Graphs outperforms PEGASIS + HLEACH in Gamma and Gaussian (Gaussian +4 nodes); small Poisson leans toward PEGASIS + HLEACH (+2 nodes), but when the number of regions increases, Multi-Graphs jumps over 5 times, with Gamma still showing a result of approximately 2.1 times better.

Energy Consumption Criteria

In Scenario 1, Multi-Graphs WUSN consumes the most: about 400 J more than FCM-WUSN when 100 nodes (10% underground) and still higher when increased to 200 nodes, only reducing by 24 J when the underground node ratio is higher. The results confirm that energy consumption decreases when more nodes are connected, and it is proportional to network lifetime. In Gamma, Gaussian, and Poisson distributions across large observation areas, Multi-Graphs saves more energy: energy consumption decreases by 30.3% (Gamma, 4 regions) and 61.9% (Poisson), while Gamma with 2 regions only reduces by 1.36%. However, with Gaussian in 4 regions, Multi-Graphs consumes more energy than PEGASIS+HLEACH by 24.8%, showing that the algorithm may become less efficient when the network expands and sensors are distributed too densely.

3.4 Evaluation

We also used Two-Way Analysis of Variance (ANOVA) (including datasets and methods) to determine if the number of underground alive nodes in the FCM-WUSN network correlates with network durability. The arithmetic results of the algorithms were tested using two-way ANOVA with repetition. Our model evaluates to demonstrate the main point: *"The underground alive nodes in the FCM-WUSN network are slightly higher than in FCM, but network lifetime is* *much better than FCM."*

A two-way ANOVA was conducted to check the variation in the experiment by considering the number of underground alive nodes and the number of rounds as network lifetime. The numerical results of the methods were analyzed using two-way ANOVA with repetition. The effects of the methods and datasets on the obtained results were considered with a significance level of $\alpha = 0.05$. The null hypotheses for this analysis are described when we recognize the following columns as methods:

- H_0: The number of underground alive nodes depends on any method, such as FCM or FCM-WUSN.
- H_1: The number of underground alive nodes does not depend on any method, such as FCM or FCM-WUSN.

The null hypothesis of this analysis is described when the rows are tracked as different datasets:

- H_0: The number of underground alive nodes depends on any changes in input values, such as increasing the number of sensors and changing the ratio between above-ground and underground nodes.
- H_1: The number of underground alive nodes is not affected by any changes in input values, such as increasing the number of sensors and changing the ratio between above-ground and underground nodes.

Both the algorithms and datasets have P-values of 0.01071 and 0.0007, respectively. The rows and columns below have P-values below 0.05. These two ideas about the number of underground alive nodes are rejected. The number of underground alive nodes is solely based on the method and statistics. The network lifetime effects of FCM and FCM-WUSN using different ANOVA criteria have P-values for both the algorithm and dataset under 0.05. This means that network durability depends on the method used and may vary with each dataset.

3.5 Summary of Chapter 3

This chapter of the dissertation discusses detailed experiments, methods, datasets, and results. The next chapter will discuss performance measures, trends, and statistical analysis of these experiments, confirming the hypotheses of the dissertation.

CONCLUSION AND FUTURE DIRECTIONS

Key Results of the Dissertation

The research aims to propose and develop energy-efficient routing strategies in wireless underground sensor networks (WUSNs), addressing both single and multi-area configurations:

- **Firstly**, the new routing model in WUSNs includes two main contributions: calculating path loss in the model and routing to estimate the network's durability. To address the proposed model for a specific underground network, the dissertation applied Taylor series level 1.

  - This dissertation proposes the FCM-WUSN algorithm for clustering based on model variance. The proposed method extends network lifetime by 2.1 times compared to the conventional clustering method, FCM. The FCM-WUSN method has 1.4 times more live nodes (calculated as the total nodes and underground nodes) compared to the FCM algorithm. Moreover, when expanding the observation to 200 sensors with 30% underground nodes, the number of underground live nodes in FCM is 1.6 times smaller than in FCM-WUSN.

- **Secondly**, multi-region routing is supported by relay nodes in the graph. A new metric is proposed with five main factors: energy, sensor type, distance between nodes, influence of neighboring nodes, and the communication radius of the nodes. From there, inter-region routing algorithms are developed.

  - Experiments in this dissertation show that network lifetime has been significantly improved; in a single region, it is 1.6 times higher than the previous

FCM-WUSN algorithm, while in multiple regions, it is on average 1.8 times higher than the PEGASIS algorithm.

- After the experiments, with uniform distribution, better results were obtained in all aspects such as network lifetime, total energy consumption in the network, and the number of alive nodes. This result remains unchanged when we double the number of observation regions from 2 to 4. However, if in the future, a system cannot distribute nodes uniformly, a Gaussian distribution also yields similar results. For network lifetime, the Multi-Graphs-WUSN algorithm has more rounds, and the number of alive nodes is also higher (32 nodes compared to 23 nodes in the PEGASIS algorithm).

- The base station (BS) position is also an important factor that we extracted from the experimental results. If the BS is located in the eastern side of the observation area, the network lifetime is extended, and nodes can be found and transmitted according to energy balancing, so the number of alive nodes is higher. This result remains unchanged when increasing from 2 regions to 4 regions.

- The number of relay nodes needed in the network: In the experiment, 50 relay nodes are used to transfer information between regions. We found that the highest number of relay nodes is used in 3 boundary areas: the edges of neighboring regions, and sensors near the BS. The relay nodes are located near underground sensors but in distant regions. For 50 nodes, only 27 sensor nodes are used frequently.

Limitations of the Dissertation

In addition to the results presented, the investigations presented in this work have some limitations, including: The data used in the experiments is self-generated. The routing experiments have not been conducted in specific environments like NS2 or NS3, OMET++.

Currently, the dissertation only addresses the problem of determining the next node or path mechanism based on distance, energy has not integrated the transmission bandwidth signal, or packet quality. The multi-region routing with relay nodes plays a role in connecting different regions to increase inter-region connectivity. However, there is no specific deployment process to determine the optimal number of relay nodes or analyze the costs and benefits of using relay nodes in the network.

Future Developments of the Dissertation

The next direction of research could develop in the following directions:

– Evaluate the proposed algorithm with larger datasets.

– Implement the algorithm in alternative network routing simulation frameworks, including NS2 or OMET++. This means integrating packet characteristics, bandwidth in transmission, and signal transmission between underground nodes with obstacles in the ground.

– Research to integrate meta-heuristic algorithms for optimizing energy consumption in WUSNs could be considered as an improvement and development.

- Research to propose a process for deploying relay nodes for routing in a multi-domain environment.

– Implement and integrate the proposed research into systems designed for monitoring crop health in the agriculture sector for a monitoring domain. For multi-domain, routing models could be tested for multiple different observation systems.

Bibliography

[1] M. C. Vuran and I. F. Akyildiz, "Channel model and analysis for wireless underground sensor networks in soil medium," *Physical Communication*, vol. 3, no. 4, pp. 245–254, 2010.

[2] S. Kisseleff, B. Sackenreuter, I. F. Akyildiz, and W. H. Gerstacker, "On capacity of active relaying in magnetic induction based wireless underground sensor networks," in *Proceedings of the IEEE International Conference on Communications (ICC)*, pp. 6541–6546, IEEE, 2015.

[3] L. Li, M. C. Vuran, and I. F. Akyildiz, "Characteristics of underground channel for wireless underground sensor networks," in *Proceedings of Med-Hoc-Net*, vol. 7, pp. 13–15, 2007.

[4] B. Yuan, H. Chen, and X. Yao, "Optimal relay placement for lifetime maximization in wireless underground sensor networks," *Information Sciences*, vol. 418-419, pp. 463–479, 2017.

[5] A. Alrabea, O. Alzubi, and J. Alzubi, "A task-based model for minimizing energy consumption in wsns," *Energy Systems*, vol. 13, pp. 671–688, 2022.

[6] N. Zhang, M. Wang, and N. Wang, "Precision agriculture—a worldwide overview.," *Computers and Electronics in Agriculture*, vol. 36, pp. 113–132, 2002.

[7] W. Dargie and C. Poellabauer, *Fundamentals of Wireless Sensor Networks: Theory and Practice.* Wireless Communications and Mobile Computing, John Wiley & Sons, 2010.

[8] P. Pandya, "Chapter e73 - transmission control protocol/internet protocol packet analysis," in *Computer and Information Security Handbook (Third Edition)* (J. R. Vacca, ed.), pp. e205–e218, Boston: Morgan Kaufmann, third edition ed., 2013.

[9] S. Phaiboon and P. Phokharatkul, "Wireless underground sensor network path loss models for durian tree," in *2021 Photonics Electromagnetics Research Symposium (PIERS)*, pp. 284–288, 2021.

[10] F. K. Banaseka, F. A. Katsriku, J.-D. Abdulai, K. S. Adu-Manu, and F. Engmann, "Signal propagation models in soil medium for the study of wireless underground sensor networks: A review of current trends," *Wirel. Commun. Mob. Comput.*, vol. 2021, pp. 8836426:1–8836426:12, 2021.

[11] D. Wohwe Sambo, B. O. Yenke, A. Förster, J. Ndong, P. Dayang, and I. Sarr, "A new fuzzy logic approach for reliable communications in wireless underground sensor networks," *Wireless Networks*, vol. 28(7), 3275-3292, 2022.

[12] N. T. Tam, H. T. T. Binh, D. A. Dung, P. N. Lan, L. T. Vinh, B. Yuan, and X. Yao, "A hybrid clustering and evolutionary approach for wireless underground sensor network lifetime maximization," *Information Sciences*, vol. 504, pp. 372–393, 2019.

[13] M. D. Hitesh Panda and B. Sahu, "A generalized mixed path loss (mpl) model for wireless underground sensor networks," *Wireless Personal Communications*, vol. 8, 03 2023.

[14] Y.-D. Yao, X. Li, Y.-P. Cui, J.-J. Wang, J.-J. Wang, and C. Wang, "Energy-efficient routing protocol based on multi-threshold segmentation in wireless sensors networks for precision agriculture," in *IEEE SENSORS JOURNAL - IEEE Sensors Council*, vol. 22, APRIL 1, 2022.

[15] D. Wu, D. M. Chatzigeorgiou, K. Youcef-Toumi, S. Mekid, and R. Ben-Mansour, "Channel-aware relay node placement in wireless sensor networks for pipeline inspection," *IEEE Transactions on Wireless Communications*, vol. 13, no. 7, pp. 3510–3523, 2014.

[16] L. H. S. Dang Thanh Hai, Nguyen Thi Tam and L. T. Vinh, "A novel energy-balanced unequal fuzzy clustering algorithm for 3d wireless sensor networks," in *SoICT '16: Proceedings of the 7th Symposium on Information and Communication Technology*, p. 180–186, December 2016.

[17] P. Maigron, "World - autonomous system number statistics - sorted by number," 2019. Truy cập ngày 3 tháng 9 năm 2024.

[18] D. King and A. Farrel, "Rfc 6805: The application of the path computation element architecture to the determination of a sequence of domains in mpls and gmpls," 2012.

[19] C. Ma, W. Liang, and M. Zheng, "Set covering-based approximation algorithm for delay constrained relay node placement in wireless sensor net-

works," *Journal of Network and Computer Applications*, vol. 59, pp. 1–9, 2015.

[20] D. Yang, X. Fang, and G. Xue, "Near-optimal relay station placement for power minimization in wimax networks," in *2011 IEEE Global Telecommunications Conference (GLOBECOM 2011)*, pp. 1–6, IEEE, 2011.

[21] F. Che, E. L. Lloyd, J. O. Hallstrom, and S. S. Ravi, "Topology control with a limited number of relays," in *2012 IEEE Global Communications Conference (GLOBECOM 2012)*, pp. 1–6, IEEE, 2012.

[22] Y. Xu, O. Ding, R. Qu, and K. Li, "Hybrid multi-objective evolutionary algorithms based on decomposition for wireless sensor network coverage optimization," *Applied Soft Computing*, vol. 68, pp. 268–282, 2018.

[23] J.-P. Vasseur and R. Zhang, "Mpls inter-autonomous system (as) traffic engineering (te) requirements." RFC 4216, Nov. 2005.

[24] L. Buzzi, M. Conforto Bardellini, D. Siracusa, G. Maier, F. Paolucci, F. Cugini, L. Valcarenghi, and P. Castoldi, "Hierarchical border gateway protocol (hbgp) for pce-based multi-domain traffic engineering," in *2010 IEEE International Conference on Communications*, pp. 1–6, 2010.

[25] I. F. Akyildiz, Z. Sun, and M. C. Vuran, "Signal propagation techniques for wireless underground communication networks," *Physical Communication*, vol. 2, no. 3, pp. 167–183, 2009.

[26] W. Ward and B. Barrow, "Another view of dimensions and db," *Proceedings of the IEEE*, vol. 60, no. 6, pp. 743–743, 1972.

[27] M. Lustgarten and B. Barrow, "Expressing watts per hertz in decibel units," *Proceedings of the IEEE*, vol. 59, no. 11, pp. 1640–1640, 1971.

[28] J. C. Bezdek, "Pattern recognition with fuzzy objective function algorithms," *Plenum Press*, vol. 1, pp. 1–224, 1984. ISBN: 0-306-41835-X.

[29] D. Abdorahimi and A. Sadeghioon, "Comparison of radio frequency path loss models in soil for wireless underground sensor networks," *Journal of Sensor and Actuator Networks*, vol. 8, p. 35, 06 2019.

[30] W. Lu, X. Xu, G. Huang, B. Li, Y. Wu, N. Zhao, and F. R. Yu, "Energy efficiency optimization in swipt enabled wsns for smart agriculture," *IEEE Transactions on Industrial Informatics*, vol. 17, no. 6, pp. 4335–4344, 2021.