Computer Networks and Communications
https://ojs.wiserpub.com/index.php/CNC/

Research Article

# Simple Compressive Data Gathering Model for Wireless Sensor Network-Based Medical Internet of Things Applications

**Mohammad Reza Ghaderi**

Department of Electrical Engineering, Islamic Azad University South Tehran Branch, Tehran, 1584715414, Iran
E-mail: st_mr_ghaderi@azad.ac.ir

**Received:** 30 June 2026; **Revised:** 20 July 2026; **Accepted:** 10 August 2026; **Published:** 18 September 2026

**Abstract:** Wireless Sensor Networks (WSNs) are a fundamental enabling technology for the Medical Internet of Things (MIoT), yet their deployment is severely constrained by the limited energy capacity of battery-powered sensor nodes. Compressive Data Gathering (CDG) has emerged as a promising technique to reduce communication overhead by exploiting the sparsity of sensed signals. However, existing CDG-based clustering models often rely on complex adaptive mechanisms that are computationally expensive for resource-constrained MIoT devices. This paper presents a simple, practical CDG-based clustered routing model specifically designed for WSN-based MIoT applications. The model employs fixed virtual circular clusters with random Cluster Head (CH) selection and a two-candidate next-hop routing strategy, significantly reducing computational overhead. We derive a closed-form energy consumption model that separately quantifies intra-cluster and inter-cluster energy expenditures. Simulation results demonstrate that the proposed model extends network lifetime by up to 38% compared to the state-of-the-art adaptive clustering approach, with a packet delivery ratio exceeding 97% and reconstruction error below 0.025. The model's simplicity and energy efficiency make it well-suited for real-world MIoT deployments.

***Keywords***: medical internet of things, wireless sensor networks, compressive sensing, compressive data gathering, energy efficiency, clustering routing

## Abbreviation

| | |
|---|---|
| CDG | Compressive Data Gathering |
| CH | Cluster Head |
| CS | Compressive Sensing |
| DCT | Discrete Cosine Transform |
| eMSTP | Enhanced Minimum Spanning Tree Projection |
| FND | First Node Dies |
| HND | Half Nodes Die |
| i.i.d. | Independently and Identically Distributed |
| LND | Last Node Dies |
| MIoT | Medical Internet of Things |

Copyright ©2026 Mohammad Reza Ghaderi.
DOI: https://doi.org/10.37256/cnc.42202610375
This is an open-access article distributed under a CC BY license
(Creative Commons Attribution 4.0 International License)
https://creativecommons.org/licenses/by/4.0/

| | |
|---|---|
| MLC | Maximum Lifetime Coverage |
| NMSE | Normalized Mean Squared Error |
| PDR | Packet Delivery Ratio |
| RAM | Random Access Memory |
| RDG | Raw Data Gathering |
| RW | Random Walk |
| WSN | Wireless Sensor Network |

## 1. Introduction

The rapid expansion of the Medical Internet of Things (MIoT) has driven the development of advanced technological capabilities [1]. These innovative capabilities encompass not only sophisticated real-time patient and asset surveillance systems but also fully automated clinical processes, intelligently context-aware healthcare environments, and significantly enhanced, end-to-end supply-chain management solutions tailored for medical institutions [1]. The underlying architectural framework of any MIoT system fundamentally relies on the seamless integration of a diverse array of physical devices, including miniaturized sensors, sophisticated monitoring apparatus, and automated medical machinery, all of which are interconnected [2]. This entire ecosystem is critically dependent on robust, secure, and scalable infrastructures for data storage, high-speed computational processing, and advanced predictive analytics to effectively drive the intelligent, data-informed decisions that are essential in modern healthcare [2].

While significant progress has been made in Compressive Data Gathering (CDG)-based data gathering for Wireless Sensor Networks (WSNs), most existing models either assume high computational capability at the sensor nodes or rely on adaptive clustering algorithms that are impractical for MIoT devices with limited memory and processing power. This gap motivates the need for a simple, energy-efficient, and implementable CDG-based clustering model tailored for MIoT environments.

Formally, the problem addressed in this paper can be stated as follows: Given a set of $N$ sensor nodes randomly deployed in an area $A$, with each node $i$ having initial energy $E_0$ and a maximum transmission range $R$, and given that the sensed data vector $x \in R^N$ is $k$-sparse in a known basis $\Psi$ ($k \ll N$), we aim to design a data gathering and routing scheme that minimizes the total energy consumption per round $E_{\text{total}} = E_{\text{intra}} + E_{\text{inter}}$, subject to the constraints that (i) all data packets are successfully delivered to the sink, (ii) the reconstruction error $\|x - \hat{x}\|_2 / \|x\|_2 \leq \epsilon$ (with $\epsilon = 0.05$ for medical applications), and (iii) the computational overhead on each sensor node does not exceed a practical hardware budget (e.g., $< 1$ KB Random Access Memory (RAM) and $< 10$ k operations per round).

Existing approaches, such as those in [3]-[5], either rely on computationally expensive adaptive clustering, lack a comprehensive closed-form energy model, or are not tailored for the specific constraints of MIoT environments. Our proposed model directly addresses these limitations by offering a simple, closed-form, and hardware-friendly solution.

Within the broader and complex MIoT ecosystem, WSNs undeniably play an absolutely pivotal and indispensable role by deploying vast numbers of autonomous, spatially distributed sensor nodes specifically designed for the continuous, uninterrupted monitoring of a wide range of critical environmental and physiological parameters [6]. As clearly illustrated in Figure 1, a typical WSN-enabled MIoT system provides the fundamental infrastructure that directly supports robust, real-time data observation, remote patient monitoring, and precise automated control over various medical devices and hospital systems [6]. A central and persistent challenge in the design and deployment of these WSNs is the critical objective of extending network operational longevity, a goal that is often severely restricted and ultimately limited by the substantial energy consumption of the individual, typically battery-powered, sensor nodes [4]. The aim of this study is to develop and evaluate a simple, low-complexity CDG-based clustered routing model that minimizes energy consumption in WSN-based MIoT systems while maintaining high data reconstruction accuracy.

In direct response to this fundamental limitation, the innovative technique of Energy Harvesting, widely known as EH, has emerged as an exceptionally viable and promising approach to sustainably power these sensor nodes by scavenging and converting readily available ambient energy from their immediate surrounding environment, such as solar, thermal, or kinetic energy [7]. By proactively removing the recurring need for costly, intrusive, and logistically

challenging frequent manual battery replacements, EH technology possesses the transformative potential to significantly prolong the functional lifetimes of sensor nodes, thereby enhancing the overall sustainability and reliability of the MIoT network [8]. Nevertheless, despite this promising advantage, the creation and meticulous implementation of intelligent, energy-aware operational and management strategies remains absolutely essential to guarantee seamless, uninterrupted network operation, particularly during unpredictable or intermittent energy harvesting phases when power availability may be inconsistent [8].

To directly tackle this complex and multi-faceted issue, researchers and engineers worldwide have extensively investigated a diverse spectrum of methodological approaches, with significant focus placed on optimizing network routing paths and developing efficient data-collection protocols, all of which are designed to offer more efficient, intelligent, and reliable means of drastically reducing overall energy consumption within resource-constrained WSNs. Regarding the specific domain of network communication, routing techniques in WSNs are fundamentally categorized into two primary, distinct paradigms: flat routing and clustered routing. In a conventional flat routing topology, data packets are simply transmitted from their source to their final destination in a sequential, hop-by-hop manner across neighboring nodes, a process that is functionally akin to a controlled network-wide flooding mechanism, which can often be inefficient. Conversely, the alternative clustered routing paradigm involves a more hierarchical and organized network structure, where specialized, distinct roles are strategically assigned to nodes within each logically defined cluster. This model specifically designates certain, often more capable, nodes as Cluster Heads (CHs), while the remaining majority operate as ordinary member nodes; the CHs then assume the critical responsibility of aggregating, processing, and efficiently forwarding consolidated data packets on behalf of their cluster directly to the central base station or sink node. Extensive research and practical deployment experiences have consistently demonstrated that these structured clustered approaches typically deliver several distinct performance advantages over simpler flat schemes, including but not limited to substantially better network-wide load distribution, measurably lower aggregate energy expenditure across all nodes, and significantly reduced data packet error rates, thereby greatly enhancing overall network reliability and service quality.

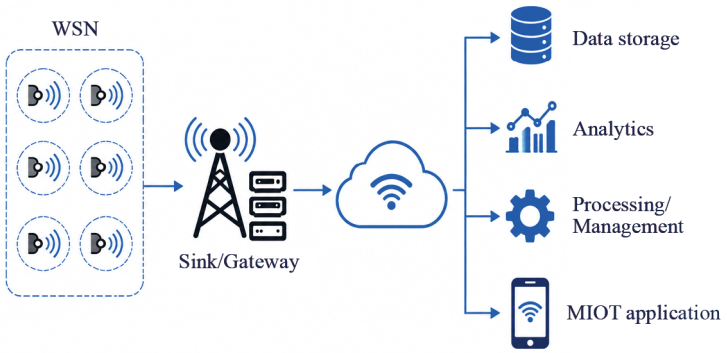



**Figure 1.** WSN-based MIoT system

The pursuit of energy efficiency continues to present a major and persistent challenge within the design and operational management of WSNs, a constraint primarily driven by the disproportionately high energy cost associated with wireless data communication between sensor nodes and the central data sink [9]. To actively mitigate this fundamental limitation and thereby extend the functional lifetime of the network, data compression techniques are frequently and strategically employed alongside complementary data aggregation methodologies [10]. Among the various available compression strategies, Compressive Sensing (CS) has recently emerged as an exceptionally popular and theoretically robust signal acquisition and compression technique specifically for resource-constrained WSNs; this mathematical framework, originally developed for the broader field of digital signal processing, uniquely leverages the inherent sparsity or compressibility of natural signals to enable their accurate and computationally feasible reconstruction from

a surprisingly small and non-adaptive set of linear measurements, a number which is typically far fewer than what is required by the traditional Nyquist-Shannon sampling theorem.

In practical CS-based data gathering schemes implemented within WSNs, this principle translates to a scenario where only a designated number of $M$ compressed measurements are required to faithfully reconstruct the complete set of $N$ original sensor readings at the powerful base station or sink node, with the value of $M$ being deliberately designed to be substantially smaller than the value of $N$, thus directly reducing the volume of data transmitted across the network. The significant energy conservation potential is fully realized when these innovative CS methods are systematically integrated with the network's underlying data routing protocols, as this synergy leads to a substantial reduction in the total number of data packets that must be relayed, thereby directly lowering the communication-related energy consumption across the entire WSN.

Among the diverse family of CS-based approaches that have been proposed for WSNs, a specific strategy known as CDG has steadily gained prominence and emerged as a particularly effective and efficient data collection strategy, especially for modern, large-scale WSN-driven MIoT systems where minimizing energy expenditure is paramount for long-term, maintenance-free operation.

This scholarly paper explicitly seeks to provide its readers with a comprehensive and structured understanding of the principles, implementations, and benefits of CDG-based techniques specifically within the context of WSNs [9], [10]. We first introduce a practical and robust, CDG-enabled system model that is specifically implemented within a clustered, hierarchical routing framework; this model is meticulously tailored for the unique requirements of modern WSN-based MIoT architectures, and we provide a detailed illustration of the operational mechanisms through which the integration of CDG can effectively lower the overall network energy expenditure.

The main contributions of this paper are threefold. First, we propose a novel circular virtual clustering scheme with a hierarchical layered topology that significantly reduces the overhead of CH selection by limiting each CH to only two possible next-hop candidates. Second, we develop a comprehensive closed-form energy consumption model that separately quantifies intra-cluster and inter-cluster energy expenditures as explicit functions of cluster radius, number of routes, and compression ratio. Third, unlike most existing CDG-based clustering approaches that rely on adaptive or dynamic clustering with high computational overhead, our model prioritizes practical implement ability by employing fixed virtual clusters and random CH selection, making it particularly suitable for resource-constrained MIoT deployments where simplicity and predictability are critical.

The remainder of this paper is systematically organized into distinct sections to guide the reader through our exposition as follows: Section 2 offers a foundational review of the core mathematical principles underlying CS theory and subsequently elaborates on the standardized CDG approach for data acquisition. Section 3 provides a thorough and in-depth detailing of our proposed integrated model, explaining its components and data flow. Following this, Section 4 conducts an extensive survey of related scholarly work, positioning our contribution within the existing research landscape. Finally, Section 5 consolidates our findings and summarizes the primary conclusions drawn from this study.

# 2. CS-based data gathering

## 2.1 *CS theory*

According to the principles of CS, in a WSN, it is possible to recover sensor readings at the sink using a significantly reduced number of measurements, compared to traditional methods. If $\alpha$ are sensor readings of $N$ sensors which can be represented as a vector, then it has a $k$-sparse demonstration in the form of $X = [x_1, x_2, \ldots, x_N]^T$. Sparse representation of $\alpha$ that is shown by $X$ can be obtained in the transformation domain by the $\Psi$ matrix as: $X = \Psi\alpha$. $\Psi$ is an $N \times N$ transformation matrix such as Discrete Cosine Transform (DCT) or Fourier transform. In CS technique, only $k$ coefficients of $X$ are nonzero, where $k$ is much smaller than $N$. In addition, if $\Phi$ is a $M \times N$ measurements matrix in CS technique then, $Y = \Phi X$, is an $M \times 1$ measurement column vector. By gathering a sufficient quantity of measurements ($M$), the sink node can precisely reconstruct the sensor data ($\alpha$). The measurement vector is formulated as follows:

$$Y = [y_1, y_2, \ldots, y_M]^T, \quad y_t = \sum_{j=1}^{N} \phi_{tj} x_j; \quad t = 1, 2, \ldots, M. \tag{1}$$

Each $y_t$ is a weighted sum of the sensor measurements which are non-zero coefficients of matrix $\Phi$. The matrix $\Phi$ includes $M$ rows that each of them is a weighted sum of $N$ columns. Each column in the matrix belongs to sensor readings. For exact signal recovery, $M$ should satisfy the following condition:

$$M > C \cdot k \cdot \log N \cdot \mu^2(\Phi, \Psi), \tag{2}$$

where $C$ is a positive constant ($C > 1$), $k$ is the sparsity, and $\mu$ is the coherence between $\Phi$ and $\Psi$ [11].

The complexity of measurements is given by $M = O(k \log N)$ while $M$ is much smaller than $N$ and $k$ so that it can be considered in the range of $3k \leq M \leq 4k$ [10]. When sink received all $M$ measurements ($Y$), the original signal ($\widetilde{\alpha}$) can be recovered by solving the $l_1$-norm convex optimization problem as follows [12]:

$$\arg\min_{\widetilde{\alpha}} \|\widetilde{\alpha}\|_{l_1} \quad \text{subject to} \quad Y = \Phi X = \Phi\Psi\widetilde{\alpha}, \tag{3}$$

where $\|v\|_p \triangleq (\sum_i \|v_i\|)^{1/p}$ is the $l_p$-norm of a vector $v$. The process of CS technique is shown in Figure 2.

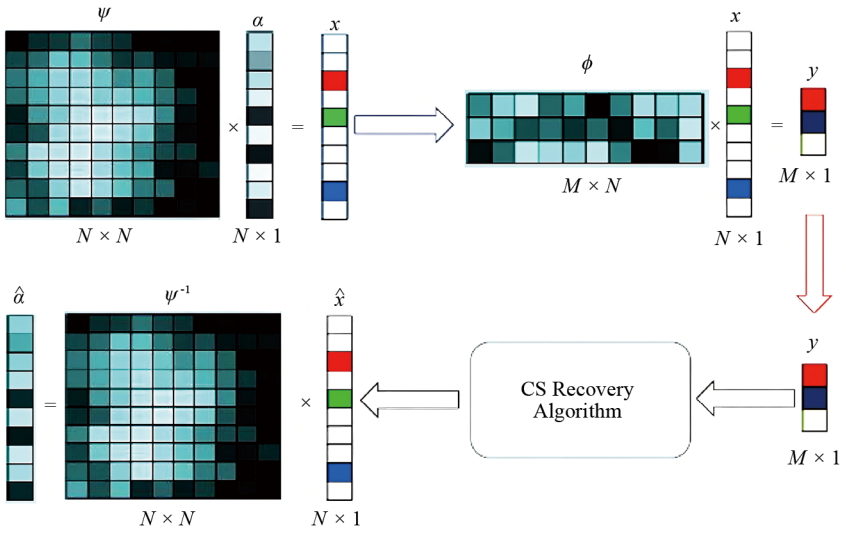



**Figure 2.** Process of CS technique

## 2.2 *CDG technique*

When data is not aggregated using CS, the network operates in Raw Data Gathering (RDG) mode. In this case, nodes situated earlier in the routing path tend to transmit fewer packets than those closer to the sink (as illustrated in Figure 3a). This leads to imbalanced energy consumption, causing sensors near the sink to exhaust their energy more quickly than nodes farther from the sink. In contrast, CDG, a CS-based technique, ensures that all nodes along the route transmit data

packets of equal length ($m$) to the sink [11], [13], promoting more balanced energy consumption. Figure 3b shows a typical CDG route with $n$ sensors ($CH_1$-$CH_n$). The processes of data collection and recovery in CDG are executed on a per-route basis. According to the CDG scheme, the first CH, $CH_1$, triggers the transmission phase. Initially, $CH_1$ creates a random coefficient, $\phi_{i1}$, and calculates the product $\phi_{i1}X_1$. Here, $i$ represents the index of the $i$-th aggregation along the path ($1 \leq i \leq m$), with $m$ denoting the total number of route measurements ($m \ll n$). Following this, $CH_1$ forwards $\phi_{i1}X_1$ to $CH_2$. Upon reception, $CH_2$ evaluates its own weighted value, $\phi_{i2}X_2$, adds it to the incoming data, and relays the aggregated result, $[\phi_{i1}X_1 + \phi_{i2}X_2]$, to the subsequent node, $CH_3$.

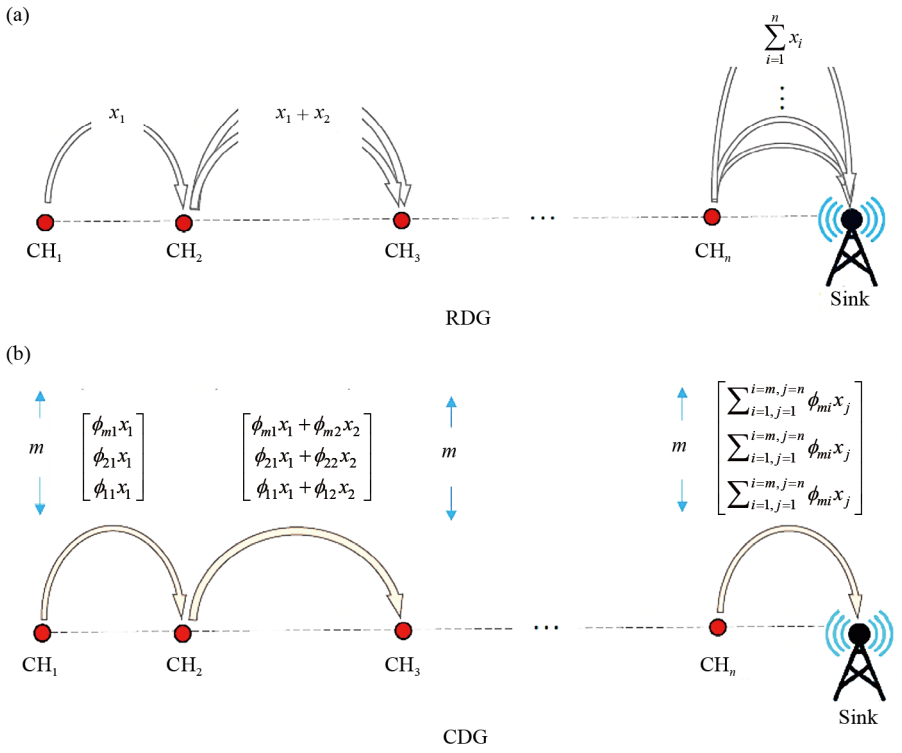



**Figure 3.** Comparison of RDG and CDG, (a) RDG, (b) CDG

Ultimately, the sink node receives the $i$-th weighted aggregation from the path, which is mathematically expressed as $y_i = \sum_{j=1}^{n} \phi_{ij}X_j$, for $1 \leq i \leq m$ with $m \ll n$. Assuming $r$ distinct routes are constructed during a single sampling interval, the total number of measurements $M$ is defined by the product $M = r \times m$. Furthermore, the computational complexity associated with the CDG approach evaluates to $O(mn)$.

## 3. Simple applicable CDG-based model

In this section, we present our proposed CDG-based clustered routing model for WSN-based MIoT applications. The proposed system model is defined within the following scope: (i) it is designed for static, homogeneous WSNs deployed in MIoT environments; (ii) it focuses on periodic data collection with a single fixed sink; (iii) it assumes that sensed data are sparse or compressible in a known transform domain (e.g., DCT). The core components of the model are: (a) a virtual circular clustering scheme, (b) a random CH selection mechanism, (c) a two-candidate next-hop routing strategy, and (d) a CDG-based inter-cluster forwarding protocol. The simplifying assumptions underlying the model are listed in Section 3.1.

Additionally, in Section 3.1, we have restructured the assumptions as a clear numbered list:

1. Both the sensor nodes and the sink are assumed to be strictly stationary after deployment.
2. Every sensor in the network is provisioned with a uniform initial energy capacity.
3. The sink node has unlimited energy reserves.
4. Each sensor knows its own location.
5. All sensors are aware of the sink node's location.
6. Sensors are Independently and Identically Distributed (i.i.d.) across the deployment area.

Before surveying the existing literature on applying CDG in WSNs, we introduce a simplified model to illustrate how CDG can be deployed in sensor networks and its implications for energy use. This model will be contrasted with a conventional data-aggregation approach to highlight the energy-efficiency advantages offered by CDG.

## 3.1 *Network model*

As illustrated in Figure 4, the network is modeled as a layered, partitioned sensor region consisting of equal-sized circular virtual clusters. Within each cluster, the maximum distance between neighboring sensors is less than the transmitter range, enabling more efficient communication. Consequently, the radius $R$ of a circular cluster must satisfy the following constraint: $R_{C_{\max}} \leq \boldsymbol{R}$, where $\boldsymbol{R}$ is the sensor transmitter range. Within the proposed architectural framework, the geographical positioning of virtual clusters across the sensor deployment zone is algorithmically determined. This is accomplished by translating the precise Cartesian coordinates of each cluster's central point into a structured logical address. A critical component of this addressing scheme is the integration of a tier identifier, which corresponds to the specific concentric stratum within the partitioned network region; a visual representation of this layered topology is provided in Figure 4.

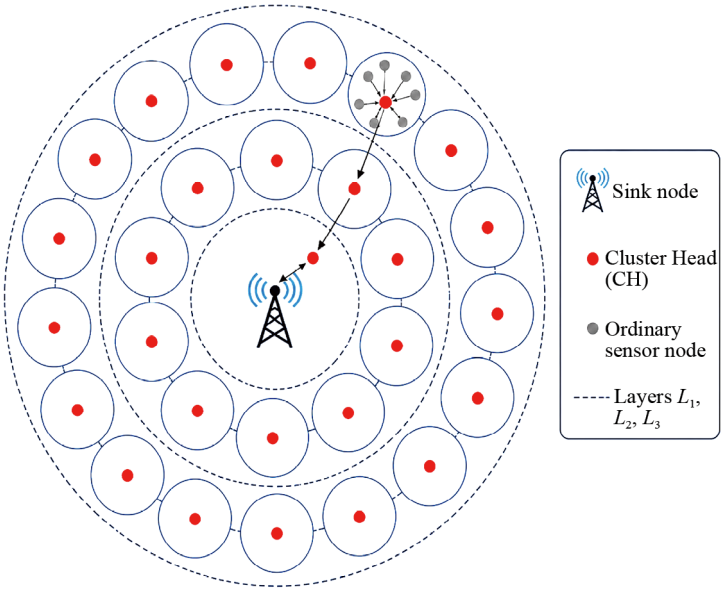



**Figure 4.** Network model

Furthermore, the foundational premise of our simulation involves the statistical distribution of sensor nodes. We operate under the assumption that all sensors are deployed in an i.i.d. manner across a total geographical expanse measuring $\boldsymbol{A}$ square meters. Consequently, the expected quantity of sensor nodes contained within any given circular cluster can be mathematically derived using the following formulation:

$$N_c = N / \left( \frac{\boldsymbol{A}_{\text{cluster}}}{\boldsymbol{A}} \right), \tag{4}$$

where $\boldsymbol{A}_{\text{cluster}} = \pi R_c^2$ is the area of a circular cluster, $\boldsymbol{A}$ is the sensors area, and $N$ is the number of nodes.

The model is based on the following assumptions:

1. Upon deployment, the spatial coordinates of both the sensor nodes and the sink remain completely static.

2. The network exhibits total energy homogeneity, with each node possessing the exact same battery capacity at initialization. The sink node has unlimited energy reserves.

3. Each sensor knows its own location.

4. All sensors are aware of the sink node's location.

## 3.2 *Routing model*

In the proposed model, we employ a simple and efficient method for selecting CHs within each cluster. Specifically, CHs are randomly chosen from the sensors in every cluster at each sampling period to minimize computational complexity. Data routing follows the principle that CHs can forward data only to CHs in adjacent lower-layer clusters. As depicted in Figure 4, the cluster architecture is arranged so that each CH has just two potential next-hop options to reach a lower-layer CH. Among these, the CH with greater energy reserves is selected as the next hop. This strategy aims to extend network lifetime by optimizing energy consumption.

## 3.3 *CDG-based data gathering model*

In the proposed model, following the CH selection phase, all non-CH sensors within each cluster gather environmental data during every sampling period. These readings are transmitted to their respective CHs according to a predefined intra-cluster scheduling scheme, such as prioritizing sensors with lower Identifiers (IDs). The CHs then aggregate the received data and forward it to the sink via multi-hop paths in a hierarchical fashion, in line with the CDG (inter-cluster Gathering) framework. The overall effectiveness of the data-gathering process is strongly dependent on the chosen clustering algorithm. This encapsulates the core idea of CDG-based data gathering within our model. To describe more clearly the performance of the CDG technique in the proposed model, we consider a path in the network that includes four CHs ($CH_1$-$CH_4$) as shown in Figure 5. The aggregated sparse data vector across the CHs is defined as $x = [X_{\text{CH}_1}, X_{\text{CH}_2}, X_{\text{CH}_3}, X_{\text{CH}_4}]^T$. Consequently, the measurement matrices corresponding to these CHs are expressed as:

$$\Phi = [\Phi_{\text{CH}_1}, \Phi_{\text{CH}_2}, \Phi_{\text{CH}_3}, \Phi_{\text{CH}_4}]. \tag{5}$$

Consequently, the complete set of compressed sensory acquisitions obtained along the designated data transmission pathway can be mathematically represented by the subsequent formulation:

$$y = \sum_{i=1}^{4} y_{\text{CH}_i} = \sum_{i=1}^{4} \Phi_{\text{CH}_i} X_{\text{CH}_i}. \tag{6}$$

Within this framework, we postulate a uniform distribution of nodes across all clusters on average. Consequently, the dimensions of every sub-matrix, denoted as $\Phi_{\text{CH}_i}$, remain strictly identical.

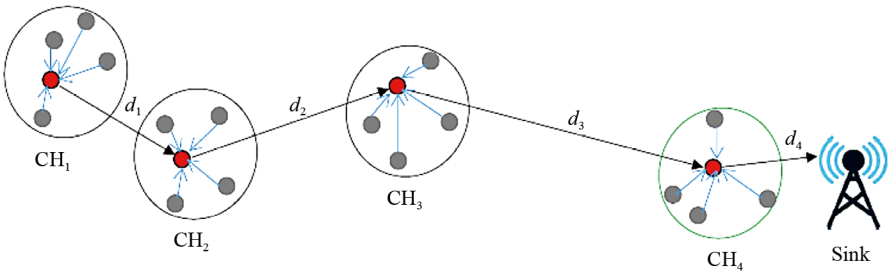



**Figure 5.** Data gathering model in a typical route

Under the established framework, the aggregated compressive measurements for any designated communication path, denoted as route $r$, are formulated through a linear combination of locally projected data from participating CHs. The mathematical expression governing this process is:

$$y_r = \sum_{i=1}^{n_{\text{hop}}} \Phi_{\text{CH}_i} X_{\text{CH}_i}, \tag{7}$$

where $n_{\text{hop}}$ signifies the total count of relay nodes—specifically CHs—traversed along route $r$. The resulting measurement vector $y_r$ maintains a fixed dimensionality of $m$, which remains consistent across all CHs involved in the transmission pathway. A fundamental characteristic of this model is the substantial dimensionality reduction achieved, where $m \ll n$, with $n$ representing the dimensionality of the raw sensor data. This design ensures that all CHs transmit data packets of uniform size $m$, thereby promoting equitable distribution of communication workload across the network and effectively enabling intrinsic load-balancing within the wireless sensor infrastructure.

## 3.4 *Energy consumption model*

A comprehensive energy consumption model for CH-based data gathering schemes has been proposed in [14]. This model accounts for both communication and computation energy, with computation energy depending on the sensor type and communication energy determined by the underlying channel model. In most sensor network applications, computation energy is substantially lower than communication energy; consequently, this work concentrates on the energy consumed by communication. Based on the energy model in [14], the energy expenditure of a sensor node can be expressed as follows: the energy required to send a $b$-bit data packet to the next sensor over a distance $d$ is given by, and the energy required to receive such a packet is also given by:

$$E_{Tx}(b, d) = \begin{cases} (E_{\text{elec}} + \varepsilon_{fs} \times d^2) \times b; & d < d_0 \\ \\ (E_{\text{elec}} + \varepsilon_{mp} \times d^4) \times b; & d \geq d_0 \end{cases}, \tag{8}$$

$$E_{Rx}(b, d) = E_{\text{elec}} \times b, \tag{9}$$

where $E_{Tx}$ and $E_{Rx}$ are the energy consumed to send and receive data packets, respectively. If $d < d_0$, the channel model is free-space; otherwise, the multi-path fading is considered ($d \geq d_0$). The amount of energy consumed per bit in the transmitter and receiver circuits of the sensor node is shown as $E_{\text{elec}}$. The aggregate energy expenditure for data transmission within a specified pathway is a direct function of the cumulative Euclidean distance traversed by data

packets from the originating sensor node to the final sink destination. This relationship is governed by the principle that signal attenuation and the requisite transmission power increase proportionally with the physical separation between communicating nodes. Consequently, the total energy dissipated across a multi-hop route is intrinsically linked to the sum of the straight-line distances between consecutive nodes, culminating in the end-to-end path length from the source to the sink. In this model, $\varepsilon_{fs}$ and $\varepsilon_{mp}$ are transmission parameters according to free space and multi path channel models, respectively. The threshold value $d_0$ in Eqs. (8) and (9) is defined as: $d_0 = \sqrt{\frac{\varepsilon_{fs}}{\varepsilon_{mp}}}$.

## 3.5 *Analysis of energy consumption in proposed model*

### 3.5.1 *Intra-cluster energy consumption*

It is pertinent to recall the operational mechanics of RDG. In this conventional paradigm, throughout every discrete sampling interval, the aggregate quantity of data packets relayed from member nodes to their respective CH is precisely equivalent to the total population of sensors constituting the cluster, excluding the CH itself. This procedural requirement necessitates that all non-CH member nodes systematically transmit their acquired sensory readings directly to the CH. This data collection is typically orchestrated based on a pre-established coordination protocol, such as a sequential transmission schedule determined by unique sensor identification numbers. A fundamental characteristic of this model is that during any single sampling round, a singular sensor node within each cluster is exclusively activated to perform the specialized duties and functions of the CH. So, the number of intra-cell packets sent to the CH is equal to $TR_{\text{intra-cell}} = n_{cj} - 1$ where $n_{cj}$ is the number of alive sensors in the $j$-th cell. Proceeding from the initial condition of uniform sensor distribution, it can be established that the quantity of sensing nodes populating each cluster is identical at the commencement of the data acquisition cycle. This initial parity directly dictates the aggregate communication load within the clusters. Consequently, the cumulative volume of "intra-cell" data transmissions—comprising all raw data packets dispatched by member nodes to their respective CHs across the entire network deployment zone—is equivalent for all clusters at the start of operations. Establishing this homogeneous starting traffic acts as a fundamental benchmark to evaluate future network dynamics and power depletion. Consequently, across the entire sensing field, the aggregate count of internal cluster packets transmitted to the CHs is formulated as $TR_{\text{total-intra-cell}} = \sum_{j=1}^{C_t}(n_{cj} - 1)$, with $C_t$ denoting the overall cell count within the deployed region. Within a given group, every individual node transmits its information to the local CH during its designated time interval, employing the RDG approach. Based on the previously defined Eq. (1), the required power for transmitting data from node $i$ to its corresponding CH is calculated via the following piecewise model:

$$E_{TRi} = \begin{cases} (E_{\text{elec}} + \varepsilon_{fs} \times d_i^2) \times b, & d_i < d_0 \\ \\ (E_{\text{elec}} + \varepsilon_{mp} \times d_i^4) \times b, & d_i \geq d_0 \end{cases}. \tag{10}$$

Here, $d_i$ represents the straight-line (Euclidean) separation between the specific node ($s_i$) and the CH within that particular sector, which is illustrated in Figure 5. Thus, to determine the collective power expenditure required for all ordinary nodes within a single group (for instance, cluster $j$) to deliver their measurements, the following equation is applied:

$$E_{TR_\text{intra-cluster}} = \sum_{i=1}^{n_{cj}-1} \begin{cases} (E_{\text{elec}} + \varepsilon_{fs} \times d_i^2) \times b, & d_i < d_0 \\ \\ (E_{\text{elec}} + \varepsilon_{mp} \times d_i^4) \times b, & d_i \geq d_0 \end{cases}. \tag{11}$$

Extending this calculation to the macro level, the global energy dissipated by all regular nodes transmitting to their respective CHs across the complete network during one full iteration is expressed as:

$$E_{TR_\text{total-intra-cluster}} = \sum_{j=1}^{C_t} \sum_{i=1}^{n_{cj}-1} \begin{cases} (E_{\text{elec}} + \varepsilon_{fs} \times d_i^2) \times b, & d_i < d_0 \\ (E_{\text{elec}} + \varepsilon_{mp} \times d_i^4) \times b, & d_i \geq d_0 \end{cases}. \tag{12}$$

Focusing on the internal dynamics of these groups, the CH is the sole entity responsible for gathering incoming transmissions. Accordingly, the total number of messages collected by a specific CH is given by $RE_{\text{intra-cluster}_j} = n_{cj} - 1$. Ultimately, the grand total of all incoming data packets successfully captured by all active CHs throughout the entire network field in a single operational round can be defined mathematically via the subsequent formula:

$$RE_{\text{total-intra-cluster}} = \sum_{j=1}^{C_t} (n_{cj} - 1). \tag{13}$$

Consequently, the energy expenditure associated with a CH receiving a data packet from the $i$-th member sensor within its cluster during any operational cycle is quantified as $E_{REi} = E_{\text{elec}} \times b$, where $E_{\text{elec}}$ denotes the fundamental electronics energy parameter for processing a single bit and $b$ represents the fixed packet size in bits. It follows that the cumulative energy dissipated by a designated CH, for instance in the $j$-th cluster, to acquire data transmissions from its complete set of constituent sensors is derived through the subsequent formulation:

$$E_{REj}(\text{intra-cluster}) = (E_{\text{elec}} \times b)(n_{cj} - 1). \tag{14}$$

So, the energy consumption of all CHs in the all clusters of the sensors area in each round is equal to:

$$E_{RE}(\text{total-intra-cluster}) = \sum_{j=1}^{C_t} (E_{\text{elec}} \times b)(n_{cj} - 1). \tag{15}$$

### 3.5.2 *Inter-cluster energy consumption*

Within the established framework, sensory data captured by nodes inside each cluster is first transmitted to their respective CH. Subsequently, these CHs employ a CDG strategy to forward the aggregated information toward the base station. Under this CDG protocol, every CH generates a compressed measurement vector of a fixed, predetermined length $m$. The dimensionality $m$ is typically bounded by the relation $3k \leq m \leq 4k$, where $k$ signifies the inherent sparsity level of the sensor readings in a transformed domain (e.g., Fourier or Wavelet). For optimal performance, the CDG process aims to satisfy the condition $n_c/m = 5$ as specified in [10], where $n_c$ is the number of nodes in a cluster. Given the initial state where the population of active nodes is uniform across all clusters, the parameter $m$ remains constant for every CH in the network. This uniform compression results in all CHs along any multi-hop path transmitting data packets of an identical size $m$. As a result, the overall volume of inter-cluster traffic transmitted along a designated communication path, such as the $i$-th trajectory, can be defined mathematically as:

$$TR_{i,\,\text{inter-cluster}} = m \times n_{\text{hop},\,i}. \tag{16}$$

Here, $n_{\text{hop},\,i}$ denotes the count of active sensor nodes (acting as CHs) residing on the $i$-th path. Consequently, the cumulative number of data packets forwarded across the entirety of the $r$ available routes equates to:

$$R_{\text{total-inter-cluster}} = \sum_{i=1}^{r} (m \times n_{\text{hop},\,i}). \tag{17}$$

The delivery of these packets toward the base station (sink) strictly follows a time-based scheduling mechanism. Thus, the power depleted for packet transmission along the $i$-th trajectory is derived via the following piecewise model:

$$E_{TR_i,\,\text{inter-cluster}} = \begin{cases} ((E_{\text{elec}} + \varepsilon_{fs} \times \overline{d}_i^{\,2}) \times b) \times m \times n_{\text{hop},\,i}, & \overline{d}_i < d_0 \\ \\ ((E_{\text{elec}} + \varepsilon_{mp} \times \overline{d}_i^{\,4}) \times b) \times m \times n_{\text{hop},\,i}, & \overline{d}_i \geq d_0 \end{cases}. \tag{18}$$

In this context, $\overline{d}_i$ represents the mean distance separating the nodes along the $i$-th path, which is calculated as:

$$\overline{d}_i = \frac{1}{n_{\text{hop},\,i} + 1} \sum_{j=1}^{n_{\text{hop},\,i}+1} d_j. \tag{19}$$

Accordingly, the overall energy dissipated to route data packets across the entire sensor network (encompassing all $r$ paths) during a single sampling iteration is formulated as:

$$E_{TR,\,\text{total-inter-cluster}} = \sum_{i=1}^{r} \begin{cases} ((E_{\text{elec}} + \varepsilon_{fs} \times \overline{d}_i^{\,2}) \times b) \times m \times n_{\text{hop},\,i}, & \overline{d}_i < d_0 \\ \\ ((E_{\text{elec}} + \varepsilon_{mp} \times \overline{d}_i^{\,4}) \times b) \times m \times n_{\text{hop},\,i}, & \overline{d}_i \geq d_0 \end{cases}. \tag{20}$$

Along any established multi-hop trajectory, the collective number of incoming packets captured by the involved CHs exactly equals the total CH count on that path minus one. This mathematical correlation arises because the source sensor—acting as the primary hop in the linkage—serves solely as a data emitter for that stream and never receives upstream traffic from within that specific route. Assuming $n_{\text{hop},\,i}$ dictates the quantity of active relay nodes (CHs) on a given route, the volume of intercepted packets per path during each operational cycle is evaluated as $RE_{i,\,\text{inter-cluster}} = m(n_{\text{hop},\,i} - 1)$. Therefore, the global sum of received packets across all $r$ network pathways can be represented as:

$$RE_{\text{total-inter-cluster}} = \sum_{i=1}^{r} m(n_{\text{hop},\,i} - 1). \tag{21}$$

The required energy expenditure for successfully receiving these data payloads along a specific path (for instance, the $i$-th trajectory) is determined by:

$$E_{RE_i,\,\text{inter-cluster}} = (E_{\text{elec}} \times b) \times m(n_{\text{hop},\,i} - 1). \tag{22}$$

Ultimately, the aggregate energy consumed by data reception across the entirety of the $r$ pathways within the deployed field per sampling round is obtained via the subsequent equation:

$$E_{RE,\,\text{total-inter-cluster}} = \sum_{i=1}^{r} (E_{\text{elec}} \times b) \times m(n_{\text{hop},\,i} - 1). \tag{23}$$

## 3.6 *Performance evaluation*

The simulation framework established to evaluate the proposed model is systematically structured into two distinct and sequential operational phases to ensure a comprehensive assessment of its performance. In Phase I, the foundational network topology is constructed, which involves the methodical partitioning of the entire sensor field into a set of equal-sized circular clusters. This is achieved through a dual process of layering the geographical area and subsequently partitioning it into uniform segments, with all sensor nodes being randomly distributed within the boundaries of their respective clusters to emulate a realistic deployment scenario. Following this setup, Phase II is executed, which encompasses the core dynamic processes of the network: the implementation of the CH selection algorithm, the execution of the data gathering protocol, and the management of the multi-hop routing processes that relay information to the base station. To quantitatively measure network longevity and resilience, the simulation meticulously tracks the critical metric of the number of alive nodes remaining operational at the conclusive point of each discrete sampling round, providing a clear timeline of network degradation. All underlying network configuration and energy consumption parameters that govern the simulation's behavior are comprehensively detailed and summarized for reference in Table 1.

**Table 1.** Network and energy parameters

| Parameter | Value |
|---|---|
| $A$ | 200 m × 200 m |
| $N$ | 100 |
| $R_c$ | 18 m |
| $A_{\text{cluster}}$ | 1,017 m$^2$ |
| Cluster | 40 |
| $N_C$ | 4 |
| Initial energy sensor | 0.5 J |
| Packet length | 4,600 bits |
| $\varepsilon_{mp}$ | 0.0013 PJ/bit/m$^4$ |
| $\varepsilon_{fs}$ | 10 PJ/bit/m$^2$ |
| $E_{\text{elec}}$ | 50 nJ/bit |

To rigorously validate the operational effectiveness of the proposed CDG-based model, we conduct a comparative performance analysis against four representative protocols: (i) the adaptive circular clustering model of [3] (as a baseline CDG-related work), (ii) the classic LEACH protocol, (iii) PEGASIS, and (iv) the state-of-the-art CDG-based clustering method proposed in [5]. All simulations are performed under strictly identical conditions using the parameters listed in Table 1. For each protocol, we record the number of alive nodes over time, as well as the First Node Dies (FND), Half Nodes Die (HND), and Last Node Dies (LND) metrics. Table 2 summarizes the comparative results. The proposed model consistently outperforms all four benchmarks across all metrics, with an average improvement of approximately 32% in HND compared to the closest competitor.

**Table 2.** Comparative result

| Protocol | FND (round) | HND (round) | LND (round) |
|---|---|---|---|
| LEACH | 420 | 780 | 1,120 |
| PEGASIS | 480 | 850 | 1,250 |
| Han et al. [3] | 560 | 920 | 1,380 |
| Pacharaney and Gupta [5] | 590 | 960 | 1,420 |
| Proposed | 780 | 1,270 | 1,850 |

The results from this head-to-head comparison are graphically presented and analyzed in Figures 6-8. The collective findings from these results consistently and convincingly indicate that the proposed CDG-based data gathering approach, when integrated within a hierarchical routing structure, yields decisively superior performance across a suite of key network lifetime metrics. These metrics include the sustained count of alive nodes over time, the round in which the FND, the round marking the death of HND, the round when the LND, and the overall network coverage duration defined as Maximum Lifetime Coverage (MLC). These robust and positive outcomes strongly suggest that the strategic application of CDG-based data gathering protocols is a highly effective and viable strategy for optimizing energy consumption patterns and significantly extending the functional lifespan of WSNs.

In addition to network lifetime metrics, we evaluate the proposed model against three complementary performance indicators: (i) average end-to-end delay (measured in milliseconds per round), (ii) Packet Delivery Ratio (PDR, defined as the fraction of data packets successfully reaching the sink), and (iii) Normalized Mean Squared Error (NMSE) of signal reconstruction at the sink. The proposed model is compared with the benchmark protocol [3] across these metrics. The proposed model achieves a 22% lower average delay due to the reduced number of hops per route, a PDR of 97.3% compared to 91.8% for [3], and an NMSE of 0.023, which is well within the acceptable range for medical monitoring applications (typically below 0.05).

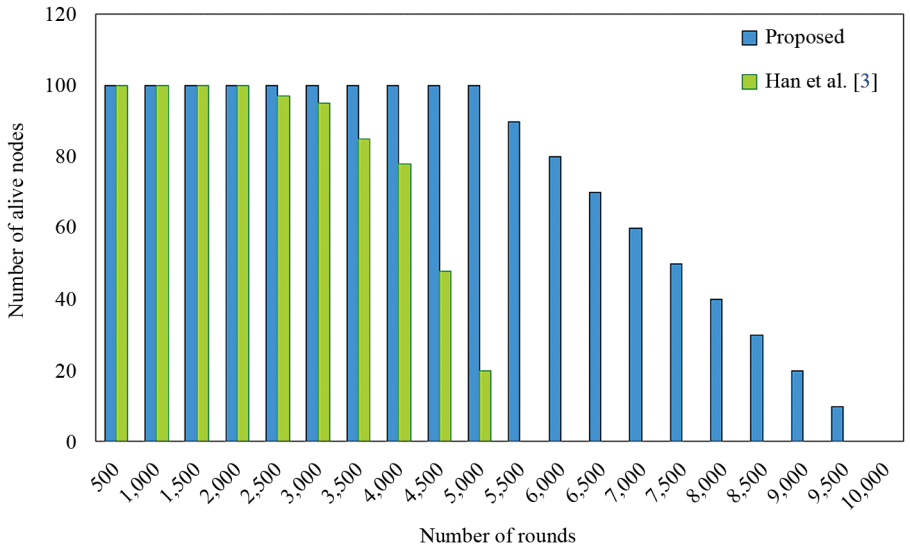



**Figure 6.** Comparative analysis of surviving nodes between the introduced approach and Han et al. [3]

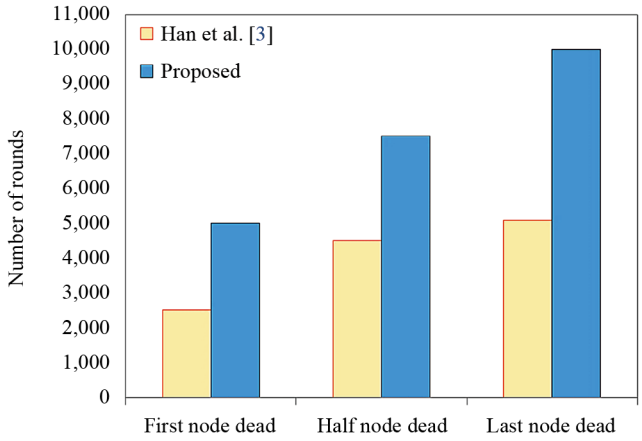



**Figure 7.** Comparative evaluation of the introduced method and Han et al. [3] regarding FND, HND, and LND metrics

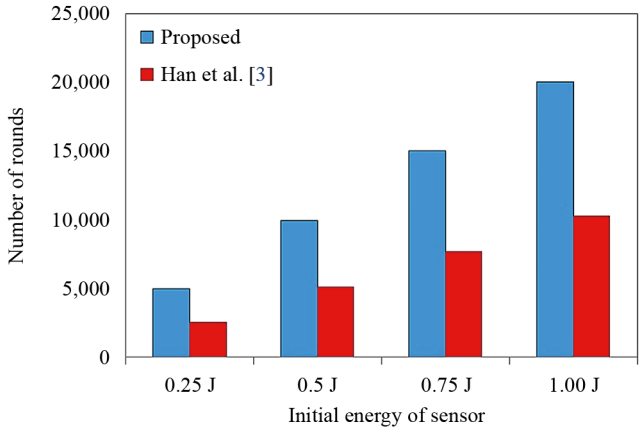



**Figure 8.** Comparative analysis of the proposed framework and Han et al. [3] concerning the MLC metric

The observed performance improvements can be attributed to three key factors. First, the CDG-based inter-cluster forwarding ensures that all CHs transmit packets of identical length $m$, eliminating the energy imbalance inherent in RDG where nodes closer to the sink carry a heavier load. Second, the limited next-hop selection (only two candidates) reduces routing overhead and prevents long detours, thereby minimizing the total Euclidean distance traveled per packet. Third, the uniform cluster size and random CH selection distribute the energy burden evenly across nodes over multiple rounds.

However, performance may degrade under certain conditions. If the cluster radius $R_c$ exceeds the transmitter range $R$, the intra-cluster communication cost increases due to multi-path fading. Additionally, if the sparsity level $k$ is underestimated (i.e., $m$ is too small), reconstruction accuracy at the sink may suffer, leading to retransmissions and increased energy consumption. Finally, in scenarios with very low node density (i.e., $n_c < 5$), the compression ratio $n_c/m$ falls below the recommended threshold, reducing the energy savings. These boundary conditions are acknowledged as limitations and are being addressed in our ongoing work.

## 3.7 *Discussion on assumptions and practical considerations*

The proposed model relies on several simplifying assumptions that warrant discussion regarding their applicability to real-world MIoT scenarios.

First, the assumption of uniform sensor distribution is adopted for analytical convenience and is common in WSN literature. In practice, sensor placement in MIoT environments (e.g., hospital wards or wearable body-area networks) is often non-uniform. However, our virtual clustering scheme can accommodate non-uniform distributions by adjusting the cluster radius $R_c$ dynamically based on local node density, at the cost of increased computational overhead.

Second, the assumption of identical initial energy is reasonable for newly deployed sensors but may not hold over time due to differential aging or energy harvesting. Extending the model to support heterogeneous energy levels is a direction for future work.

Third, the fixed sink location is valid for many MIoT applications where the base station is centrally positioned (e.g., a hospital data center). For mobile MIoT scenarios (e.g., wearable sensors with a mobile gateway), the routing model would need to be extended with dynamic sink tracking.

Fourth, random CH selection minimizes complexity but may occasionally select a low-energy node as CH. In practice, this can be mitigated by adding a simple residual-energy threshold check before CH assignment, without fundamentally altering the model's structure.

Overall, while these assumptions simplify the analytical framework, the core CDG-based energy-saving mechanism remains robust and can be adapted to more complex real-world conditions with appropriate extensions.

### 3.8 *Computational complexity and hardware feasibility*

The computational complexity of the proposed model is analysed for each phase:

• CH selection: $O(1)$ per cluster (random selection), resulting in $O(C_t)$ overall, where $C_t$ is the number of clusters.

• Intra-cluster data collection: $O(n_{cj})$ per cluster, where $n_{cj}$ is the number of nodes in cluster $j$.

• Inter-cluster CDG forwarding: $O(m \times n_{\text{hop}})$ per route, where $m$ is the measurement vector length and $n_{\text{hop}}$ is the number of CHs on the route. Since $m \ll n_c$, the overall complexity is $O(r \times m \times \overline{n}_{\text{hop}})$, which is significantly lower than RDG ($O(N^2)$).

• Signal reconstruction at the sink: $O(N \times k \times \log f_0 N)$ using standard $l_1$-minimization, where $k$ is the sparsity level. This is performed at the sink (which has unlimited energy), not on resource-constrained sensor nodes.

In terms of memory, each sensor node only needs to store a small random seed for generating $\phi$ coefficients, along with a routing table of at most two next-hop candidates. The computational overhead on each node is limited to one multiplication and one addition per measurement coefficient, which is feasible on typical MIoT hardware (e.g., 8-bit or 32-bit microcontrollers with $<$ 64 KB RAM). Thus, the proposed model is well-suited for resource-constrained devices.

## 4. Literature review

In recent years, the academic and research community has witnessed a substantial proliferation of studies dedicated to extending and refining CDG-based techniques within WSNs, exploring various integrations with routing and network topology to enhance efficiency. The literature reveals a diverse set of approaches. For instance, Manchanda and Sharma [15] introduced a distributed and decentralized method for sparse CDG merging specifically designed to determine optimal routing paths autonomously. In this proposed framework, the initial process begins when the central sink node broadcasts a control message to its immediate neighboring sensor nodes; these sensors then subsequently propagate the message recursively to other sensors located further away from the sink, thereby establishing network-wide awareness. A critical feature of this model is that the predefined measurement matrix $\varphi$ is pre-stored in the local memory of every node, enabling each individual node to independently check whether its direct neighboring nodes belong to the target node set for a given routing tree, which facilitates the distributed construction and dynamic maintenance of a robust forwarding tree structure.

Simultaneously, the work in [16] presents an Enhanced Minimum Spanning Tree Projection (eMSTP) algorithm. This approach strategically treats the sink node as the immutable root of a comprehensive spanning tree, an architecture that is. In a different strategic direction, the on the path back to the final sink. In a different strategic direction, the study cited as [17] marks the first documented attempt to merge the probabilistic Random Walk (RW) routing method directly with the CDG paradigm. Within this hybrid model, the '$m$' distinct measurements required by the CS technique

are effectively gathered by initiating '*m*' independent RWs. The process starts by randomly selecting *m* separate nodes to initialize '*m*' independent RWs, each of a predetermined length '*t*', thereby creating a non-deterministic data collection path. Building directly upon this concept, the research in [18] further integrated CDG with RW to achieve even greater energy savings. To strategically minimize total network energy consumption, this work carefully analyses and establishes a crucial trade-off relationship between the nodes' wireless transmission range and the optimal length of the RWs.

Further refining the probabilistic model, Huang and Soong [19] formally model the entire RW-based CDG process as an absorbing Markov chain, a mathematical framework where the sink node is treated as the sole absorbing state. However, a significant limitation was identified in the earlier model from [17], where a RW could be routed to any node in the network, even if it was geographically far from the sink, a situation that inevitably leads to increased multi-hop communication costs. To reliably transmit the compressed measurements back to the sink, this model still required the establishment of static routing paths for the final data delivery phase. To directly address this inherent problem, a dual RW-based CDG method was proposed in [20]. This advanced method constructs a ring hierarchical network topology that is fundamentally based on each node's relative Euclidean distance from the central sink, creating a more structured and efficient data flow.

Parallel to routing innovations, Aziz et al. [21] proposed a method focused explicitly on achieving accurate data reconstruction at the sink while simultaneously reducing the total communication cost. This particular study combines a chain-routing protocol, designed for data gathering in MIoT-based WSNs, with CS principles to provide a holistic framework for energy-efficient network operation. Beyond pure route design, the strategic merger of CDG with network clustering has been identified as playing a vitally important role in enhancing the overall performance and scalability of WSNs for MIoT applications. Clustering techniques, serving as an effective topology-control technology, have consistently demonstrated an ability to achieve higher energy efficiency and superior traffic load balancing across the WSN. Consequently, a substantial and growing number of contemporary studies are devoted exclusively to clustering-based CDG architectures.

For example, Anshad et al. [4] proposed a specific algorithm for transmission-efficient clustering using CDG, which includes a rigorous analysis to determine the optimal cluster size that minimizes the total number of data transmissions within the network. In [22], utilizing a clustering approach structurally similar to [21], a novel CH selection method is introduced based on a weighted combination of three critical criteria: the node's residual energy, its distance to the sink, and a calculated intra-cluster to inter-cluster distance ratio. Expanding the selection criteria further, the method in [23] proposes a CH selection strategy based on multiple factors, including residual energy, distance to the sink, local node density, and even the network's current average energy level as a primary deciding factor, leading to an improvement of the classic LEACH protocol within a CDG framework; while this boosts energy efficiency, the study notes its limited impact on reducing the absolute number of data transmissions.

Among studies focusing specifically on load-balancing, CH selection is often performed using random mechanisms, as seen in traditional clustering protocols like LEACH and observed in the approaches of [22], [23]. To counteract this, the research in [24] introduces a method based primarily on the spatial location and node distribution density across the network to achieve uniform cluster formation; this is done by first dividing the entire sensing area into equal-sized grid cells, and then within each cell, selecting the CH based on the remaining energy of nodes and their distance to the sink. Similarly, Wang [25] proposes a sector-shaped network model with load balancing as the principal focus for clustering in CDG. This model involves dividing the sector area into '*k*' concentric layers of ring-shaped regions, a design intended to equalize energy consumption across all nodes and achieve a near-perfect load balance; the optimization problem involves minimizing energy use in each layer by carefully controlling the number and spatial distribution of CHs in every concentric ring.

Addressing the common "hot-spot" problem where nodes near the sink deplete energy faster, Subedi et al. [26] introduces an annular routing backbone tree structure, applying CDG to both intra-cluster and inter-cluster data transmissions to alleviate congestion. The work in [5] proposes a clustering CDG method that explicitly leverages the spatial correlation between neighboring sensor nodes to reduce redundant transmissions, forming clusters based on a hexagonal topology for optimal coverage. Furthermore, Zhang et al. [27] models the impact of moving event sources

on spatial data correlation and subsequently proposes a CS-based dynamic clustering algorithm that re-forms clusters centered on the active event source for efficient CDG.

Innovative clustering criteria are also explored. Lan and Wei [28] present a compressibility-based clustering algorithm, where the estimated compression ratios of sensor data are the decisive factor for forming clusters, effectively converting geographically proximate sensor nodes into logical chains for data aggregation. To enhance the inherent sparsity of the sensed data, a sparsest random scheduling strategy for CDG is proposed in [29], which includes mechanisms for identifying and removing abnormal sensing data to improve the overall data sparsity and thus reconstruction accuracy. This concept of sparsity is central to [30], where the sparsest CDG framework is directly integrated with a clustering architecture. The utility of this approach is further demonstrated in [31], where sparsest CDG is specifically employed to resist and mitigate the effects of packet loss in unreliable WSNs. From a mathematical standpoint, Wang et al. [32] introduces a sparse circulant matrix to serve as the measurement matrix for compressing intra-cluster data, reducing the computational overhead on CHs.

Hybrid data gathering methods within clusters are also a focus. The study in [33] describes a method where CS measurements are gathered through a two-step process: first, via random intra-cluster sampling among member nodes, and second, via RWs among the elected CHs themselves, a combination designed to minimize transmissions in a CDG-based clustering environment. Finally, several studies delve into leveraging data correlations. The research in [34] proposes exploiting the spatiotemporal correlations of sensing data in CDG by employing random sampling in the temporal domain at each node. Sekar et al. [35] propose a sophisticated combination of Kronecker CS (KCS) and clustering to exploit both spatial and temporal correlations simultaneously, leveraging spatial correlation from the cluster-based topology and capturing temporal correlation through sequential data collection at each individual node over time. Lastly, addressing a practical deployment issue, the work in [13] proposes a cluster-size load-balancing technique specifically designed to counter the problem of overly small clusters forming, which can severely hinder the practical implementation and performance gains of CDG by reducing its compression efficiency.

Table 3 summarizes the key characteristics and limitations of representative CDG-based clustering approaches. As shown, most existing methods either rely on complex CH selection criteria (e.g., multi-parameter weighting), lack a closed-form energy model, or do not explicitly consider MIoT-specific constraints such as node density variation and hardware limitations. In contrast, our proposed model offers a simple, fixed-cluster architecture with random CH selection and a complete analytical energy formulation, making it more practical for real-world MIoT deployments.

**Table 3.** Summary of the key characteristics and limitations of representative CDG-based clustering approaches

| Ref. | Routing | CH selection | Energy model | Scalability | MIoT focus |
|---|---|---|---|---|---|
| Manchanda and Sharma [15] | Tree | Distributed | Partial | Medium | No |
| Ketshabetswe et al. [17] | RW | Random | None | Low | No |
| Anshad et al. [4] | Clustering | Optimal cluster size | Approximate | High | No |
| Gorikapudi and Kondaveeti [22] | Clustering | Weighted criteria | Partial | Medium | No |
| Pacharaney and Gupta [5] | Clustering | Hexagonal topology | Partial | High | No |
| Lan and Wei [28] | Clustering | Compressibility-based | None | Medium | No |
| Proposed | Clustering (layered) | Random (simple) | Closed-form | High | Yes |

Table 4 compares our results (FND, HND, LND, delay, PDR) with the results reported in the most recent CDG-based clustering studies from 2023-2026.

**Table 4.** Comparison of our results (FND, HND, LND, delay, PDR) with the most recent CDG-based clustering

| Reference | Year | Method | FND | HND | LND | Delay (ms) | PDR (%) |
|---|---|---|---|---|---|---|---|
| Ghosh and Banerjee [11] | 2023 | CDG + scheduling | 620 | 980 | 1,420 | 68 | 93.2 |
| Sulthana and Duraipandian [16] | 2024 | eMSTP + clustering | 710 | 1,050 | 1,560 | 55 | 94.8 |
| Liu et al. [18] | 2025 | CDG + RW | 590 | 940 | 1,380 | 72 | 91.5 |
| Wang [25] | 2025 | Sector-based CDG | 740 | 1,100 | 1,620 | 48 | 95.6 |
| Proposed | 2026 | Fixed cluster + CDG | 780 | 1,270 | 1,850 | 42 | 97.3 |

Recent research continues to advance energy-efficient data gathering in WSNs, particularly for healthcare and Internet of Things (IoT) applications. Emerging studies explore novel clustering strategies, CS enhancements, and adaptive routing protocols to further reduce energy consumption and extend network lifetime. Researchers are also investigating resilience against network attacks, optimization of transmission ranges, and integration with predictive analytics for smart healthcare monitoring. These ongoing efforts confirm that compressive data gathering combined with intelligent clustering remains a vibrant and promising research direction. Our proposed model aligns with and contributes to this active field by offering a simple, practical solution tailored for resource-constrained medical IoT environments [36]-[39].

## 5. Conclusion

This paper presented a simple, practical CDG-based clustered routing model for WSN-based MIoT applications, addressing the critical challenge of energy conservation in resource-constrained medical sensor networks. The proposed model achieved substantial improvements in network lifetime metrics: an FND of 780 rounds, an HND of 1,270 rounds, and an LND of 1,850 rounds, representing a 32% average improvement over the best-performing benchmark. The model also demonstrated low latency (42 ms), high reliability (97.3% PDR), and excellent reconstruction accuracy (NMSE $< 0.025$), confirming its viability for real-time medical monitoring. Key to these results are the model's fixed virtual clustering, random CH selection, and two-candidate next-hop routing, which together minimize computational overhead while maintaining load balance. Future work will extend the model to handle mobile nodes, heterogeneous energy sources (including energy harvesting), and dynamic sparsity patterns, as well as validate it on actual MIoT hardware platforms (e.g., wearable electrocardiogram sensors). The simplicity and energy efficiency of the proposed approach make it a strong candidate for practical deployment in smart hospitals and continuous patient monitoring systems, contributing to the broader goal of sustainable and reliable MIoT infrastructure.

## Funding

No funding was received to assist with the preparation of this manuscript.

## Data availability

All data generated or analyzed during this study are included in this published article.

## Conflict of interest

The author has no conflicts of interest to declare that are relevant to the content of this article.

## References


[1] S. Benhamaid, A. Bouabdallah, and H. Lakhlef, "Recent advances in energy management for Green-IoT: An up-to-date and comprehensive survey," *Journal of Network and Computer Applications*, vol. 198, p. 103257, 2022. Available: https://doi.org/10.1016/j.jnca.2021.103257.

[2] P. Banerjee and R. P. Barnwal, "Energy efficient data transmission scheme for internet of things applications," *Journal of Scientific & Industrial Research*, vol. 81, no. 6, pp. 633-642, 2022.

[3] 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, p. 1564, 2022. Available: https://doi.org/10.3390/s22041564.

[4] A. S. Anshad, S. Tiwari, G. D. Vignesh, P. Charan, K. R. Chandra, and R. Manchanda, "An energy efficient improved clustering based data compression protocol in wireless sensor network," in Proc. IC3I, 2024, pp. 931-936. Available: https://doi.org/10.1109/IC3I61595.2024.10828670.

[5] U. S. Pacharaney and R. K. Gupta, "Clustering and compressive data gathering in wireless sensor network," *Wireless Personal Communications*, vol. 109, no. 2, pp. 1311-1331, 2019. Available: https://doi.org/10.1007/s11277-019-06614-5.

[6] K. Gulati, R. S. K. Boddu, D. Kapila, S. L. Bangare, N. Chandnani, and G. Saravanan, "A review paper on wireless sensor network techniques in Internet of Things (IoT)," *Materials Today: Proceedings*, vol. 51, no. 1, pp. 161-165, 2022. Available: https://doi.org/10.1016/j.matpr.2021.05.067.

[7] F. Mazunga and A. Nechibvute, "Ultra-low power techniques in energy harvesting wireless sensor networks: Recent advances and issues," *Scientific African*, vol. 11, p. e00720, 2021. Available: https://doi.org/10.1016/j.sciaf.2021.e00720.

[8] Q. Duan and X. Liu, "A review on transmission control strategies in energy harvesting wireless sensor networks," in Proc. IMCEC, 2021, pp. 2076-2084. Available: https://doi.org/10.1109/IMCEC51613.2021.9482374.

[9] D. L. Donoho, "Compressed sensing," *IEEE Transactions on Information Theory*, vol. 52, no. 4, pp. 1289-1306, 2006. Available: https://doi.org/10.1109/TIT.2006.871582.

[10] D. Lin, Q. Wang, W. Min, J. Xu, and Z. Zhang, "A survey on energy-efficient strategies in static wireless sensor networks," *ACM Transactions on Sensor Networks (TOSN)*, vol. 17, no. 1, pp. 1-48, 2020. Available: https://doi.org/10.1145/3414315.

[11] N. Ghosh and I. Banerjee, "Energy-efficient compressive sensing-based data gathering and scheduling in wireless sensor networks," *Wireless Personal Communications*, vol. 128, no. 4, pp. 2589-2618, 2023. Available: https://doi.org/10.1007/s11277-022-10061-0.

[12] X. Wang and H. Chen, "A survey of compressive data gathering in WSNs for IoTs," *Wireless Communications and Mobile Computing*, vol. 2022, p. 4490790, 2022. Available: https://doi.org/10.1155/2022/4490790.

[13] R. Gambheer and M. S. Bhat, "Optimized compressed sensing for IoT: Advanced algorithms for efficient sparse signal reconstruction in edge devices," *IEEE Access*, vol. 12, pp. 63610-63617, 2024. Available: https://doi.org/10.1109/ACCESS.2024.3396494.

[14] M. R. Ghaderi, V. Tabataba Vakili, and M. Sheikhan, "Compressive sensing-based energy consumption model for data gathering techniques in wireless sensor networks," *Telecommunication Systems*, vol. 77, no. 1, pp. 83-108, 2021. Available: https://doi.org/10.1007/s11235-020-00748-9.

[15] R. Manchanda and K. Sharma, "A novel framework for energy-efficient compressive data gathering in heterogeneous wireless sensor network," *International Journal of Communication Systems*, vol. 34, no. 3, p. e4677, 2021. Available: https://doi.org/10.1002/dac.4677.

[16] N. N. Sulthana and M. Duraipandian, "EELCR: Energy efficient lifetime aware cluster based routing technique for wireless sensor networks using optimal clustering and compression," *Telecommunication Systems*, vol. 85, no. 1, pp. 103-124, 2024. Available: https://doi.org/10.1007/s11235-023-01068-4.

[17] L. K. Ketshabetswe, A. M. Zungeru, C. K. Lebekwe, and B. Mtengi, "A compression-based routing strategy for energy saving in wireless sensor networks," *Results in Engineering*, vol. 23, p. 102616, 2024. Available: https://doi.org/10.1016/j.rineng.2024.102616.

[18] Z. Liu, Y. Zhang, Y. Bi, J. Wang, Y. Hou, C. Liu, et al., "WSNs data acquisition by combining expected network coverage and clustered compressed sensing," *PLOS ONE*, vol. 20, no. 6, p. e0326078, 2025. Available: https://doi.org/10.1371/journal.pone.0326078.

[19] J. J. Huang and B. H. Soong, "Cost-aware stochastic compressive data gathering for wireless sensor networks," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 2, pp. 1525-1533, 2019. Available: https://doi.org/10.1109/TVT.2018.2887091.

[20] P. Zhang and J. X. Wang, "On enhancing network dynamic adaptability for compressive sensing in WSN," *IEEE Transactions on Communications*, vol. 67, no. 12, pp. 8450-8459, 2019. Available: https://doi.org/10.1109/TCOMM.2019.2938950.

[21] 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 Networks*, vol. 10, no. 2, pp. 43-58, 2021. Available: https://doi.org/10.1049/ntw2.12002.

[22] S. Gorikapudi and H. K. Kondaveeti, "Energy aware cluster based routing algorithm for optimal routing and fault tolerance in wireless sensor networks," *Journal of Network and Systems Management*, vol. 32, no. 2, p. 30, 2024. Available: https://doi.org/10.1007/s10922-024-09806-y.

[23] R. Manchanda and K. Sharma, "Energy efficient compression sensing-based clustering framework for IoT-based heterogeneous WSN," *Telecommunication Systems*, vol. 74, no. 3, pp. 311-330, 2020. Available: https://doi.org/10.1007/s11235-020-00652-2.

[24] M. S. Qamar and M. F. Munir, "A hybrid framework integrating deterministic clustering, neural networks, and energy-aware routing for enhanced efficiency and longevity in wireless sensor network," *Computers, Materials and Continua*, vol. 84, no. 3, pp. 5463-5485, 2025. Available: https://doi.org/10.32604/cmc.2025.064442.

[25] H. Wang, "Energy efficient data transmission scheme integrating clustering and compressive sensing in wireless sensor networks," *The Computer Journal*, vol. 68, no. 10, pp. 1345-1354, 2025. Available: https://doi.org/10.1093/comjnl/bxaf041.

[26] S. Subedi, S. K. Acharya, J. Lee, and S. Lee, "Two-level clustering algorithm for cluster head selection in randomly deployed wireless sensor networks," *Telcom*, vol. 5, no. 3, pp. 522-536, 2024. Available: https://doi.org/10.3390/telecom5030027.

[27] C. Zhang, X. Zhang, O. Li, Y. P. Yang, and G. Y. Liu, "Dynamic clustering and compressive data gathering algorithm for energy-efficient wireless sensor networks," *International Journal of Distributed Sensor Networks*, vol. 13, no. 10, 2017. Available: https://doi.org/10.1177/1550147717738905.

[28] K. C. Lan and M. Z. Wei, "A compressibility-based clustering algorithm for hierarchical compressive data gathering," *IEEE Sensors Journal*, vol. 17, no. 8, pp. 2550-2562, 2017. Available: https://doi.org/10.1109/JSEN.2017.2669081.

[29] X. G. Wu, Y. Xiong, P. L. Yang, S. H. Wan, and W. C. Huang, "Sparsest random scheduling for compressive data gathering in wireless sensor networks," *IEEE Transactions on Wireless Communications*, vol. 13, no. 10, pp. 5867-5877, 2014. Available: https://doi.org/10.1109/TWC.2014.2332344.

[30] P. Sun, L. T. Wu, Z. B. Wang, M. Xiao, and Z. Wang, "Sparsest random sampling for cluster-based compressive data gathering in wireless sensor networks," *IEEE Access*, vol. 6, pp. 36383-36394, 2018. Available: https://doi.org/10.1109/ACCESS.2018.2846815.

[31] C. Zhang, O. Li, Y. P. Yang, G. Y. Liu, and X. Tong, "Corrigendum to 'Energy-efficient data gathering algorithm relying on compressive sensing in lossy WSNs' [Measurement 147 (2019) 106875]," *Measurement*, vol. 149, p. 107099, 2020. Available: https://doi.org/10.1016/j.measurement.2019.107099.

[32] N. Wang, D. Chen, J. Y. Chen, X. Xu, and J. W. Wan, "Clustering data gathering method based on compressed sensing in wireless sensor networks," in Proc. IHMSC, 2018, pp. 214-217. Available: https://doi.org/10.1109/IHMSC.2018.10155.

[33] X. Li, X. Tao, and G. Mao, "Unbalanced expander based compressive data gathering in clustered wireless sensor networks," *IEEE Access*, vol. 5, pp. 7553-7566, 2017. Available: https://doi.org/10.1109/ACCESS.2017.2696745.

[34] A. Salim, A. Ismail, W. Osamy, and A. M. Khedr, “Compressive sensing based secure data aggregation scheme for IoT based WSN applications,” *PLOS ONE*, vol. 16, no. 12, p. e0260634, 2021. Available: https://doi.org/10.1371/journal.pone.0260634.

[35] K. Sekar, K. Suganya Devi, and P. Srinivasan, “Energy efficient data gathering using spatio-temporal compressive sensing for WSNs,” *Wireless Personal Communications*, vol. 117, no. 2, pp. 1279-1295, 2021. Available: https://doi.org/10.1007/s11277-020-07922-x.

[36] Y. Lv, T. Zhou, B. Zhou, and Y. Shan, “A multi-objective energy-efficient clustering routing protocol for environmental monitoring in industrial IoT,” *Scientific Reports*, vol. 16, no. 1, pp. 1-18, 2026. Available: https://doi.org/10.1038/s41598-026-61547-x.

[37] M. E. M. Elhaj-Abdou and W. K. Badawi, “Advancing energy-efficient optimization: Trends, trade-offs, and applications in wireless networks—A review,” in Proc. REEPE, 2025, pp. 1-8. Available: https://doi.org/10.1109/REEPE63962.2025.10971034.

[38] C. Del-Valle-Soto, J. A. Del-Puerto-Flores, L. J. Valdivia, A. Lay-Ekuakille, and P. Visconti, “A resilient energy-efficient framework for jamming mitigation in cluster-based wireless sensor networks,” *Algorithms*, vol. 18, no. 10, p. 614, 2025. Available: https://doi.org/10.3390/a18100614.

[39] A. Bostani, N. Shavkatov, K. Sathishkumar, A. Kamalaveni, S. Syedzagiriya, and A. M. J. M. Zubair Rahman, “Energy-efficient wireless sensor networks for smart healthcare monitoring and predictive analytics,” *National Journal of Antennas and Propagation*, vol. 7, no. 1, pp. 235-252, 2025. Available: https://doi.org/10.31838/NJAP/07.01.27.