A Novel Chain Formation Scheme for Balanced Energy Consumption in WSN-based IoT Network

Se-Jung LIM, Byeong-Mu SHIN*

Abstract: In the Internet of Things (IoT) technologies, wireless sensor networks (WSNs) are one essential part. The IoT network commonly consists of WSNs, where hundreds or even thousands of small sensors are capable of sensing, processing, and sending environmental phenomena in the targeted region. The energy consumption imbalance of sensors becomes the cause of the network performance decrement, as sensor nodes have limited energy available for operation after being randomly deployed. Therefore, more research is necessary for the design of energy-efficient routing algorithms in energy-constrained WSNs. This paper focuses on the chain-based routing algorithm, which is a popular algorithm for achieving energy efficiency in WSN-based IoT network. Chain-based routing algorithms offer numerous advantages for WSNs, such as energy conservation and extended lifetime of WSNs. However, they face challenges due to the issue of internal communication imbalance. The objective of our study is to design a novel chain formation scheme that improves the energy consumption imbalance caused by internal communication in WSN-based IoT network. The proposed scheme is categorized in three phases (initial communication phase, chain formation phase, and data collection phase). In the first phase, the sink acquires their location information from sensors deployed in the sensing region and then the sensing region is separated into sub-regions, and with the number of sensor nodes is balanced employing the concept of the k-dimensional binary tree (K-D-B-tree). The sub-regions are organized into a binary tree structure, which is then formed into a chain. Lastly, data is collected along the chain, and the selected representative sensor transmits the collected data to the sink. We utilized the OMNET++ simulator and demonstrated effective simulation results in terms of network lifetime and average residual energy. In the simulation results, a novel chain formation scheme outperforms the power-efficient gathering in sensor information systems (PEGASIS) and the concentric clustering scheme for efficient energy consumption in the PEGASIS (CCS).

Keywords: chain formation; energy-balanced routing; internet of things (IoT); k-dimensional binary tree (K-D-B-tree); wireless sensor networks (WSNs)

1 INTRODUCTION

The realm of wireless sensor networks (WSNs) has rapidly developed as an essential part of the Internet of Things (IoT) [1-5]. The IoT network commonly consists of WSNs, where hundreds or even thousands of small sensors can do so interconnected without requiring human interaction [6]. WSN-based IoT networks are being applied various applications such as environmental monitoring, home automation or smart building (domotics), agriculture (smart farming), aquaculture, smart healthcare in medical field, intelligent transportation and logistics, and so on [7]. WSNs have several features such as easy deployment, self-organization, lower cost, and fault tolerance. In WSNs, a small sensor is capable of sensing, processing, and sending environmental phenomena within the targeted region and made up of various hardware components. Firstly, a controller processes all data to execute the code. Secondly, different types of memory are used to store programs and data, respectively. Thirdly, sensors and actuators can observe or control the environment. Lastly, a communication device is a device for transmitting and receiving data over a wireless channel. Among these components, the controller is the center of a wireless sensor node since it gathers data, processes it, sends and receives data, and determines the actuator’s behavior. Sensors can be integrated by sensor technology, and sensing capabilities can be sophisticated. The memory component is fairly straightforward. Random Access Memory (RAM), Read-Only Memory (ROM), or, more commonly, Electrically Erasable Programmable Read-Only Memory (EEPROM), flash memory, and other types of memory are used [8]. Memory size, particularly RAM, can be critical in terms of manufacturing costs and power usage. The communication devices are used to interchange data between sensors according to the transmission medium (radio frequencies, optical communication, and ultrasound, etc.). The power supply is a crucial system component for wireless sensor nodes. Traditionally, batteries are used to store energy. Since the battery cannot be changed in constrained areas, sensor nodes often function with limited energy resources after deployment. Hence, the energy of sensor nodes may be running out, and the wireless communication between sensors can be permanently interrupted. Particularly during the internal communication of WSNs, the energy consumption imbalance of sensors is one of the reasons that the energy of some sensors might quickly deplete [9]. Since the network performance (connectivity, coverage, and lifetime, etc.) is affected by the node’s death [10, 11], the rapid energy consumption of some sensors leads to a decrease in the network performance. Therefore, more research is necessary for the design of energy-efficient routing algorithms for energy constrained WSNs. To improve the network performance of WSNs, improved chain-based algorithms, which are popular among them, have been studied. In [12-14], the merit of chain-based routing algorithms can be checked as mentioned below:
- The chain-based routing algorithm consumes less energy than topologies based on the cluster.
- The distribution of energy consumption is uniform.
- It provides a longer network lifetime through low energy consumption.
- The algorithm reduces the dynamic cluster formation overhead of cluster-based routing algorithms.

However, existing chain-based routing algorithms still have some drawbacks, such as unbalanced energy consumption and long-distance data transmission in internal communication. The purpose of our study is to design a novel chain formation scheme to improve the energy unbalanced routing in chain-based routing algorithms of WSNs. Our scheme uses the k-dimensional binary tree (K-D-B-tree) and the in-order traversal of a binary search tree for our novel chain formation. Our novel chain formation scheme is categorized as set forth in Sections 3.1 (initial communication phase), 3.2 (chain...
formation phase), and 3.3 (data collection phase) and has the following significance for WSNs. First, we use chaining by the concept of the K-D-B-tree and the in-order traversal of a binary search tree for data transmission in WSNs. Second, we overcome the imbalance of energy consumption and avoid long-distance data transmission in internal communication. We present performance results (network lifetime, average residual energy) that support our scheme using the OMNET++ simulator [15, 16], and demonstrate that our research may greatly alter the imbalance in energy use. The remaining sections of the article are organized as follows. The next section reviews chain-based routing algorithms and the K-D-B-tree structure. Then, Section 3 goes into great depth about our novel chain formation scheme. Meanwhile, Section 4 investigates the performance of simulation findings. Finally, section 5 brings this paper's investigation to a close.

2 RELATED WORKS

2.1 Chain-based Routing Algorithms

Many researchers have developed and improved chain-based routing algorithms aimed at prolonging the network lifetime of WSNs. The power-efficient gathering in sensor information systems (PEGASIS) [17] is the most well-known routing algorithm based on a greedy chain formation approach in WSNs. In PEGASIS, sensors are randomly distributed in the network range and then the chain is organized using the greedy algorithm. Each sensor fuses its data with that of its neighbours. The fused data is transmitted to a leader sensor along the chain. Then the leader sensor transmits the fused data to the base station (BS). To collect data, the token-passing mechanism is used. In [18], multiple levels and chains are used. A chain is formed at each level by using the greedy algorithm. A leader sensor is selected at each level. The leader sensor receives the data from the lower level and then forwards the reception data and the data from its members towards the higher level. At the higher level, the leader sensor sends the collected data to the sink. In the concentric clustering scheme for efficient energy consumption in the PEGASIS (CCS) [19], the authors proposed a concentric clustering scheme based on multiple chains. The signal strength from the BS decides the level of sensors. A head sensor in the different levels sends its location to both of its upper and lower levels. Finally, at the lowest level, a head sensor sends the gathered data to the base station. In CCS, the relay communication between head sensors plays a crucial role in minimizing the communication distance from the head sensors to the BS. In [20], one or two heads are selected at each level. In the levels with two heads, the collected data is sent to the lower level with one head sensor, and then the head sensor sends data to the next level cluster with two head sensors. Data is sent to the base station along with the head sensors based on diamond-shaped structures. In [17], since the leader sensor transmits the aggregated data to the base station for each round, the failure of a leader sensor leads to data loss. In [21], the data loss problem is addressed. The leader's neighbour, which is collected on the basis of residual energy, sends the fused data to the base station. In [22], the network field is split into sub-areas of equal size. In each sub-area, the chain structure is organized and is connected using bridge nodes. In [23], authors adopted the multi-hop transmission for conserving energy. Optimal communication and hop counts are considered to decrease time of forwarding the data and optimize energy consumption. Among algorithms mentioned above, PEGASIS employs the greedy algorithm to create a chain within the network, resulting in long communication distances between sensor nodes. Consequently, the increased energy consumption associated with these long distances leads to a rapid depletion of node energy, ultimately impacting the overall performance of the network. In CCS, once a sensor is assigned a level in the network, it remains at that level unless there is a change in the BS's location. As a result, early congestion near the BS may occur, leading to potential network performance issues in terms of WSNs.

2.2 The k-dimensional Binary Tree (K-D-B-tree)

The k-dimensional binary tree (K-D-B-tree) was initially created in 1981 to answer the problem of retrieving multi-key records using range queries [24]. It is one of traditional data structures for indexing and searching multidimensional data efficiently and is a multi-way tree with fixed-size pages that branches each page into regions. Starting from the root, which covers the entire space, it recursively subdivides the entire space into the k-dimensional space by hyper-rectangles to create an index. In the K-D-B-tree, since a search space is partitioned into two subspaces, the number of pages accessed on a path from the root page to a leaf page is the same for all leaf pages. There are pages of two types in the K-D-B-tree, as follows:

- Region pages: every internal page is a region page, and the elements in entries store data pointers and subspace description (region, page ID) pairs.
- Point pages: point pages, which are all leaf pages, are an index record that consists of the point and location pairs. Since the location of the database record is given as the location, they are actual data or references to them.

To summarize: the region and point pages contain the region and page ID, and point and location pairs, respectively. In the (region, page ID) pair, the page ID represents the page number assigned the child that covers the corresponding region. As for (point, location), the location refers to the disk page where the record containing the given point is stored. In each dimension, a region is given, and the region represents the space that satisfies all these regions. If it is a 2-dimensional space, regions are in the shape of regular rectangular blocks. So, there is no overlapping between the regions. The properties of the K-D-B-tree are followed.

- If each page is considered a node, then in a region page, each page ID becomes a node pointer, and the graph structure of the tree with the root ID is created.
- All leaf pages have the same path length from the root page.
- On the page, the regions are disjointed.
- The union of the root page and a region page is \( \text{domain}_0 \times \text{domain}_1 \times \ldots \times \text{domain}_{k-1} \).
- A region is the coalition of the regions on a region page that is the child page referred to by the child ID.
3 PROPOSED SCHEME

In this section, we describe in detail our novel chain formation scheme to achieve energy efficiency via balanced energy consumption in WSN-based IoT network. The key contribution of our study is to design a novel chain formation using the K-D-B-tree and the in-order traversal of a binary search tree. Our novel chain formation scheme consists of an initial communication, a chain formation, and a data collection phase. We use the centralized method to form the chain-based network structure. In the initial communication phase, a sensor transmits its location via the Global Positioning System (GPS) signal and then transmits its location information to a sink. The sink obtains its location information from all sensors. After the information transmission is finished, the algorithm enters the chain formation phase.

3.1 Initial Communication Phase

In the network region, when sensors are randomly deployed, the first initial communication phase commences. In this phase, the chain-based network structure has not been established. Each sensor can get its location via the Global Positioning System (GPS) signal and then transmit its own location information to a sink. The sink obtains its location information from all sensors. After the information transmission is finished, the algorithm enters the chain formation phase.

3.2 Chain Formation Phase

In this phase, all sensors are organized into a single chain. Fig. 3 explains how to divide the sensing region while cyclically and sequentially iterating through the x and y domains based on 2 - dimensions with odd overall sensors using the concept of the K-D-B-tree. The set of sensors in the sensing region is made into a binary search trees by the K-D-B-tree. First, sensor nodes are sorted as
the coordinates \((x, y)\) of them and then the region is split into two sets, \(L_0\) (left region) and \(R_0\) (right region), by the \(x\) domain based on the division number \(\text{div}_0\) by the concept of the K-D-B-tree. If the number of sensors is an odd number, a splitting node becomes root, otherwise the root node is decided within the left division region. Until all sensors become nodes of the binary search tree, this process is performed cyclically and sequentially iterating via the \(x\), and \(y\) domains. The process of constructing binary trees based on a 2 dimensions is illustrated in Fig. 3. In the proposed scheme, the final chaining result is formed using the in-order traversal of a binary search tree, as described in Fig. 4. Once the sink calculates the chaining results, it broadcasts them to each sensor in the network. This broadcasting ensures that all sensors receive the necessary information to construct the chain-based network. Upon receiving the chaining results, each sensor utilizes the information to establish connections with their respective neighboring sensors, forming the chain-based network.

### Table 1 Construction of binary trees

<table>
<thead>
<tr>
<th>Input:</th>
<th>Output:</th>
</tr>
</thead>
<tbody>
<tr>
<td>- List of sensors in left or right regions</td>
<td>- List of sensors in left and right regions</td>
</tr>
<tr>
<td>- A splitting node of the current level - 1</td>
<td>- A splitting node of the current level</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>1: Set ( n ) = count of sensors</td>
<td>1: Set ( \text{div}_n ) = 0</td>
</tr>
<tr>
<td>2: Set ( \text{div}_n ) = 0</td>
<td>3: if ( 0 &lt; n \mod d ) then</td>
</tr>
<tr>
<td>4: ( \text{div}_n = (n / 2) + 1 )</td>
<td>5: else</td>
</tr>
<tr>
<td>6: ( \text{div}_n = (n / 2) )</td>
<td>7: end if</td>
</tr>
<tr>
<td>8: Sort {list of sensors} as ascending of x- or y-coordinates</td>
<td>9: Select a splitting node of the current level</td>
</tr>
<tr>
<td>10: Set List of sensors in left and right regions</td>
<td>11: Set a splitting node of the current level - 1</td>
</tr>
</tbody>
</table>

### 3.3 Data Collection Phase

To collect data in the \(r\)th round, a leader node is elected by \(r \mod n\) (\(r\): round, \(n\): number of sensors). As shown in Fig. 4, the leader node uses the token-passing mechanism and takes turns transmitting the aggregated data to a sink. The small token does not affect the simulation results. In each round, a sensor node transmits its data, which is aggregated with the information of death nodes toward a leader node. Each sensor performs the data aggregation. The sink calculates the chain-based network structure for the next round via the information of death sensors and then broadcasts it. If the first sensor among sensors dies, the chain formation phase is repeated to start the next round. In our scheme, because sensor nodes are made into a binary tree using the concept of the K-D-B-tree and then the chain-based network structure is calculated by the in-order traversal of a binary tree, sensor nodes can communicate with their near neighbors for collecting data. Therefore, our scheme can improve the internal communication overhead issues by long-distance data transmission and unbalanced binary search tree.
3.4 Research Method

To analyze and evaluate the performance of our study, simulation experiments are carried out until the first sensor dies in the same network environment using the OMNET++ simulator. The performance of our scheme is compared and analyzed with the existing PEGASIS and CCS in terms of three metrics (longest-distance communication, network lifetime, and average residual energy). For the simulation, assumptions and conditions are defined as follows:
- Sensors are randomly deployed in the network region.
- They are homogeneous.
- They are conscious of their stationary location information via GPS reception.
- They can transmit their data to a sink.
- The sensor adjusts its radio range.
- There are restrictions on their energy, and they cannot be recharged.

In Tab. 2, the energy consumption formulas are defined. We increased the reliability of the simulation by applying the same energy model used in PEGASIS and CCS.

4 SIMULATION RESULTS ANALYSIS

4.1 Simulation Parameters

Tab. 3 shows the simulation parameters. Sensors are randomly deployed in a 100 meter × 100 meter region. The initial energy of each sensor is 1 J. The packet length (k) is 2000 bit. E_{elec} per bit is dissipated at 50 nJ/bit to run the transmitter or receiver circuitry. The power of the transmitter amplifier is 100 pJ/bit/m². Energy (5 nJ/bit/message) is spent on aggregating data.

4.2 Simulation Scenarios

In this experiment, we divided the simulation scenarios into two aspects and conducted simulations for evaluating the performance of PEGASIS, CCS, and the proposed scheme. In the first aspect of the experiment, we generated scenarios with different numbers of sensor nodes (N = 50, 100, 150, 200) within a 100 m × 100 m simulation area, as detailed in Tab. 4. The sensor nodes were deployed randomly, as illustrated in Fig. 6. Indeed, since WSNs often consist of hundreds or even thousands of sensor nodes, and the density of these nodes plays a crucial role in determining how efficiently the network operates, the deployment density of WSNs can significantly impact network performance. Careful consideration of the deployment density is crucial to ensure optimal performance in WSNs. In the second aspect, we located the sink to (50, 200) and (50, 300), outside the sensing region. Typically, the sink is located far from sensor nodes, resulting in significant energy consumption to transmit data to the sink. Conducting experiments to measure network performance by varying the location of the sink in WSNs can be helpful in predicting and optimizing network coverage in real-world operating environments.
4.3 Simulation Metrics

In the scenarios as described in Tab. 4, the simulation experiments are carried out until the first sensor dies in the same network environment. We measure three metrics (longest-distance communication, network lifetime, and average residual energy). The performance of our scheme is verified via comparison with PEGASIS and CCS. Existing chain-based routing algorithms still have the energy consumption imbalance. Some sensors consume much more energy and die due to long-distance transmitting in internal communication. Our scheme is proposed to reform long-distance transmission in internal communication. The longest-distance communication is the longest distance among communication distances for receiving and sending data packets in the chain. The network lifetime is the total number of rounds until the first sensor dies. When a chain leader sends the collected data and the sink receives it, the round increments the count by 1. The average residual energy is measured as the average value for the residual energy of all alive sensors when the network lifetime is measured by the death of the first sensor.

4.4 Simulation Results and Discussion

During data transmission of a WSN, as the energy is consumed in proportion to the square of the distance, as described in Tab. 2, the relationship between energy consumption and communication distance is closely related. Fig. 7 illustrates the measurement results for the longest distance among the various communication distances measured under different numbers of sensors and algorithms. The PEGASIS has longer communication distances compared to both the CCS and our scheme. These results do not vary significantly with different numbers of sensors, indicating that the longest-distance communication in the PEGASIS tends to remain consistent regardless of the sensor count. The CCS demonstrates a decreasing trend in results as the number of sensors increases. However, it still has higher values compared to our scheme across all the results. In the case of our scheme, the measurement results for the longest-distance communication, with the number of sensors being 50, 100, 150, and 200, are approximately 46, 61, 63, and 45 meters respectively.
From these results, it can be observed that the proposed scheme facilitates balanced communication among sensor nodes regardless of the density of the sensor nodes. Consequentially, in our scheme, sensor nodes can achieve a balanced energy consumption for data transmission.

In Fig. 8, the number of sensors is represented on the x-axis, while the network lifetime is represented on the y-axis. The network lifetime is determined by the number of data transmission rounds until the first sensor among the sensors dies. As shown in Fig. 8, the results of the proposed scheme exhibit approximately 2 - 3 times more efficient performance compared to PEGASIS and CCS. In particular, in our conducted simulations, we also measured the network performance based on the sink's location, (50, 200) and (50, 300). In both cases, we have observed that the proposed scheme outperforms PEGASIS and CCS in terms of network performance. In our study, the sensor nodes are characterized by restricted energy resources and are not rechargeable. Thus, the results obtained in our study carry significant importance as they demonstrate the potential of enhancing the operational lifetime of the network within these energy-constrained conditions.

Fig. 9 indicates the average residual energy when the first sensor node dies. Energy is a valuable resource in WSNs, as shown in the discussion by many researchers [25]. Since the balanced energy consumption of each sensor in the internal communication must be considered, an accurate measurement of the average residual energy cost is also important. In Fig. 9, the number of sensors and the average residual energy are represented in x-axis and y-axis coordinates, respectively. As shown in the results, the existing algorithms (PEGASIS and CCS) consumed less energy than the energy consumption of our scheme. If the energy consumption were to decrease, the lifetime of the network would decrease as well. These results indicate the energy imbalance caused by the abnormal long-distance transmission and reception between some sensors in both algorithms (PEGASIS and CCS).

5 CONCLUSION

In WSN, which is an essential part of the IoT, because sensor nodes must operate with only limited energy resources in the region of targeted environment, their operation as long as possible is critical. For this reason, the primary concern in WSNs lies in designing routing protocols that enable energy saving, and more research is necessary for the design of energy-efficient routing algorithms. In this paper, we present a novel chain formation scheme to draw forth the balanced consumption of energy to achieve energy efficiency. The objective of this study is to enhance the network lifetime by addressing the unbalanced energy consumption resulting from internal communication in WSN-based IoT networks. The proposed scheme consists of three phases. During the initial communication phase, data, such as location information, is exchanged between each sensor and a sink. In the chain formation phase, the K-D-B-tree is employed to construct a 2-dimensional binary tree. Subsequently, the binary tree is organized into a chain by utilizing the in-order traversal of the tree. Ultimately, by employing the K-D-B-tree and utilizing binary tree traversal in the divided regions, each sensor can efficiently establish communication with its neighboring sensors. These are the core of our scheme as outlined in this paper. Finally, in the data collection phase, data is gathered and transmitted to the sink. We compared the performance of our scheme and existing algorithms (PEGASIS and CCS) using the OMNET++ simulator. The performance of our study is better than PEGASIS and CCS algorithms regarding longest-distance communication, network lifetime, and average...
residual energy in homogeneous network models. Nevertheless, the issue of unbalanced energy consumption in WSNs is not entirely resolved. In future work, we aim to improve our scheme through a performance comparison analysis between the current state-of-the-art chain-based routing algorithms and the scheme proposed in this paper. Furthermore, because the results obtained from simulations can indeed vary based on various factors such as the parameters, measurement conditions, network settings, and specific performance metrics being evaluated, we would like to interpret the results obtained through the diversification of network parameters.

Acknowledgements

This study was supported by a research fund from Honam University, 2021.

6 REFERENCES


Contact information:

Se-Jung LIM
Liberal Arts & Convergence Studies,
Honam University,
120, Honamdae-gil, Gwangsan-gu, Gwangju-si, 62399, Republic of Korea
E-mail: limsejung@honam.ac.kr

Byeong-Mu SHIN
(Corresponding author)
Interdisciplinary program of Smart City,
Chonnam National University,
50 Daehak-ro, Yeosu, Jeollanam-do, 59626, Republic of Korea
E-mail: 217023@jnu.ac.kr