Abstract—Most of researches about the design of the topology and the operation of a wireless sensor network are focusing on how to take advantage of all sensor nodes to collect information that covering the whole sense area. A criticism has been ignored that the goal of lengthening lifespan is conflict with the complete coverage strategy due to the overlapping of sensor area among nodes result in unnecessary waste of energy. In this study, we propose another proposition from the idea of partial coverage strategy: a near perfect coverage rate can be achieved by least nodes to save the energy of the whole network. The proposed topology is constructed by clustered structure. According to the simulation results that made by comparing with the classic clustered topology, LEACH, and other related methods proposed recently, the study find that the method proposed in this paper can double the lifespan of a wireless sensor network under the same condition. Furthermore, the construction method can efficiently achieve graceful degradation to maintain the usage of a sensor network in a long period.

Keywords— wireless sensor network; clustered topology; partial coverage; hierarchical routing protocol; leach protocol

I. INTRODUCTION

Wireless sensor networks (WSNs) are composed of low-power, low-cost, small-sized and multifunctional sensor nodes. Each sensor node is capable of communication, data collection, and processing. The sensor nodes collaborate among each other to establish a sensor network for collecting critical information from the surrounding environment. These nodes are densely deployed either inside or very close to the phenomenon that is being monitored, every node has the ability to sense, process and transmit data to a base station (BS). WSNs have shown their suitability to various kinds of applications [1].

Since the sensor nodes in a WSN are usually equipped with non-rechargeable batteries. The network will lose its functionality as any node exhausts its energy. Due to the limitations of energy and cost of sensor nodes, it is crucial to minimize the energy consumption to prolong network lifespan [2]. For achieving the goal of minimization of energy consumption, most of the related researches focus on providing an energy-efficient routing protocol for network [3]. These studies were conducted on design algorithms to completely cover an area, such as [4] and [5], etc. We can call the issue as “complete coverage” strategy. However, there is a criticism been ignored that the goal of lengthening lifespan is conflict with the complete coverage strategy due to the overlapping of sensor area among nodes result in unnecessary waste of energy.

Besides the main viewpoint of the coverage issue which is to increase coverage rate of a wireless sensor network to achieve a most sensitive wireless sensor network, there is another viewpoint about coverage that is being proposed by recent studies: Can we limit the power consumption by guaranteeing an ideal coverage with less sensor nodes? One of the techniques which recently attract researchers’ attention is to reduce the coverage quality to trade for network lifespan. For example, mudflows monitoring applications may only require part of the area to be covered in sunny days. Thus, to extend network lifespan, we can lower the coverage quality if it is acceptable. The problem of covering only a portion of an area is referred to as the “partial coverage” strategy[6]. In this study, we propose another proposition from the idea of partial coverage strategy: a near perfect coverage rate can be achieved by least nodes to save the energy of the whole network.

Based on the characteristics of WSN, Al-Karaki and Kamal categorized routing protocols into flat routing protocol, location-based routing protocol and hierarchical routing protocol [7]. Flat routing protocol [8] is a data-centric routing protocol. The advantage of a flat routing protocol is that each sensor node does not need to store much route information; the disadvantage is that if the required data is returned by several sensor nodes simultaneously, it may cause network congestion or a broadcast storm. Consequently, it is not applicable to large-scale networks. In the location-based routing protocol [9], each sensor node is equipped with a Global Positioning System (GPS) to distinguish its own geographical position from others and figure out the best transmission path for itself. The location-based routing protocol can reduce transmitting unnecessary packets and is best for network topology that changes frequently. However, the cost is much higher than other protocols, too.

In the hierarchical routing protocol [10], the geographical region of a WSN is divided into several clusters, where each cluster selects a cluster head (CH) responsible for collecting the data from cluster members and transferring data to the BS via hierarchical routing. Utilizing clusters in hierarchical routing protocol has its advantages because it allows less power consumption in each node and the CH is capable of processing data aggregation. However, it imposes a larger load on the CH,
as a CH must manage not only data collection but also data relay.

The classical hierarchical protocol such as LEACH [11] has better performance in saving the energy consumption. LEACH divides the communication process into rounds with each round including a set-up phase and a steady-state phase. In the set-up phase, some sensor nodes are selected as cluster heads (CHs) according to certain rules and other nodes join in the clusters as member nodes. In the steady-state phase, the CHs collect and aggregate the data coming from their own cluster members and then transmit them to a BS. There are many improved versions based on LEACH have been proposed in the last decade. Improved-LEACH, abbreviated as ILEACH in this paper, is an example that proposed most recently [12]. ILEACH improved energy efficiency by the modification to choosing of cluster heads formula. Meanwhile, the work establish a vice cluster head for each cluster during the communication process, which aims to diminish the energy consumption spent on the re-clustering and prolong the time of being in a steady-state phase.

In this study, we focus on how to build an effective clustered WSN based on a partial coverage strategy proposed in previous work[13]. The proposed topology control method, named as Keeping Desired Partial Coverage rate (KDPC), can keep predefined Quality of Service (QoS) of coverage as long as the sensor nodes can stand. This research modifies the KDPC by distributing sensor nodes with uneven deployment, and dynamically redefining the expected QoS to save the waste of unnecessary sensor nodes. The lifespan of the WSN can then be further lengthened over 10% than original design.

This paper is organized as follows: The related definitions about coverage estimation used for the method are defined first as the basis of this research in Section II. The idea and detail algorithm of KDPC accompany with the performance evaluation by comparing with other similar method is then introduced in the Section III. In Section IV, we first immigrate KDPC to play as a NEAR "complete coverage" strategy and then improve the method step by step. The improved lifespan are shown in the simulation results exhibited in the section. Finally, Section V is a conclusion of this study.

II. DEFINITIONS ABOUT COVERAGE ESTIMATION

A. Adaptive Random Clustering Protocol

KDPC is a topology control method modified by a partial coverage algorithm named Adaptive Random Clustering (ARC) [14]. There are several novel features been exploited in ARC: First, instead of using location information, ARC forms a multi-hop cluster network with a required connectivity by using a novel cluster head competition scheme and proper transmission of power settings. Second, the lifespan of a wireless sensor network is prolonged through balancing energy consumption by updating cluster heads periodically, and reducing energy consumption by collision avoidance mechanism. Finally, ARC is suitable for practical applications of large-scale or high-density wireless sensor networks due to its distributed processing, scalable cluster topology, and easy management.

KDPC periodically resets the network to reorganize the construction of cluster structure. Each round of operation consists of a setup phase and a steady phase. The setup phase is composed of three steps: CHN(Cluster Head Node) selection, route setup, and NCHN(Non-Cluster Head Node) activation. In the steady phase, every active NCHN sends data packets to its cluster head using specified band and time slot in a TDMA manner, and sleeps in other time slots. Every CHN sends data packets to its upstream CHN according to its route table established in the stage of route setup. CHNs operate in a CSMA/CA protocol for inter-cluster communication.

It is assumed that there are several channels (bands) available, and different clusters use different channels to prevent collisions of data packets. Besides, there is a primary band assigned for transmitting all kinds of data packets in setup phase and inter-cluster data packets in steady phase. The transmit power for packet transmissions between a CHN and an NCHN is denoted as P1, corresponding to the intra-cluster communication radius is denoted as RC1, and the transmit power for packet transmissions from a CHN to another CHN or to the sink is denoted as P2, corresponding to the inter-cluster communication radius is RC2. P1 is less than P2 in order to reduce energy consumption and to inhibit mutual interference of data packets among different clusters.

B. Estimation of Coverage Intensity

For evaluating the coverage and connectivity of a wireless sensor network, the coverage intensity C is considered as below: if active nodes are independently and uniformly distributed in a deployment region, and each node can connect to the BS through a certain route, then network coverage ratio can be calculated from the number of active nodes n, the sensing area S, and the sensing radius R, as shown in following equation:

\[
C = 1 - (1 - q)^n, \quad q = \frac{\pi R^2}{S}
\]

In the equation, q is the probability that a point in the deployment region is covered by a single active node, and n is total number activated in a round. Thus \((1 - q)^n\) is the probability that the point is not covered by any active nodes.

C. The Calculation of Backoff Time

Because the CHN should be determined in a local range to organize a cluster and collect all sensed data for inter-cluster transmission. A CHN should be the nodes with relatively more residual energy. All deployed nodes will compete for being CHNs in CSMA/CA protocol. We summarize two conditions that should be considered for the determination of backoff time: First, the power of an executed node will be consumed no matter whether the node is a CHN or not. Although a CHN will consume more energy, a node that acts as a NCHN many times should exhaust its energy, too. Second, two nodes may consume different power in the same round even though they both act as CHN. Different CHNs may consume different energy because different amounts of downstream packets are passed through the nodes. Based on the considerations, the equation to calculate the backoff time for Node i can be
expressed as (2), and the symbols appearing in the equation are defined below:

\[
T_e = T_i \times \frac{R_{current}}{R_{pass}} + T_i \times \frac{1}{\sum \frac{N_{pass,transmit(k - 1)}}{N_{pass,transmit(k - 1)} + N_{pass,transmit(k)}}} + T_i \times R_{\text{rand}[0,1]} \quad (2)
\]

- **R\text{execute}_i**: The number of rounds that the node i has ever been activated, whether it is acting as a CN or a NCHN.
- **R\text{Current}_i**: The total number of rounds that has been executed.
- **N\text{pass}_transmit(i, k)**: When node i acted as a CHN in the kth round, the number means the total number of CHNs that transmit packets through node i in the round; otherwise, the number is zero.
- **N\text{Cluster}_k**: The number of NCHNs required for a cluster in kth round.

There are three terms to make up the backoff time. The first term is weighed by the times that the node has been activated in the past rounds. The second term is weighed by the average amount of data that has been transmitted through the node. The third term is random in value introduced as an enhancement to the existing CSMA/CA algorithm. It keeps the nodes close to each other from sending CHN packets at the same time. T1, T2, and T3 are weights of these terms, respectively. Based on the equation, the differences among nodes will be more precisely determined by the status of actual energy consumption.

### D. NCHN Activation

If a node has received CHN packets from other CHNs before its timer expire, it will become a candidate of NCHN in the current round and try to join the cluster with strongest signal strength at the end of CHN selection step. At last, CHN selects its member randomly according to the number of NCHNs calculated in advance to achieve the required coverage.

The number of members deployed in a cluster is usually redundant in order to take duty in turn and prolong the network lifespan. For each cluster, only a part of NCHNs are activated for each round. For the sake of energy saving, the number of active nodes should be minimized while still satisfying the required coverage ratio. For the research proposed in this paper, the number of members in a cluster is determined by expanding the estimation of the relationship between the coverage ratio and the number of nodes into every cluster. By modifying the equation (1) to fit the condition of a cluster, the number of members in a cluster can be calculated individually as (3).

\[
N_{\text{cluster}} = \frac{\log(1 - C_0)}{\log(1 - \frac{\pi R_i^2}{\pi R_{\text{current}}^2})} \quad (3)
\]

In the equation, \( C_0 \) means the required coverage ratio provided via the message issued from BS. \( R_i \) and \( R_{\text{current}} \) are the sensing radius of sensor node and the radius of intra-cluster communication range respectively. According to the number of members in a cluster calculated by every CHN individually, each CHN can deploy its NCHNs concurrently then reduce the time for constructing clusters.

### III. KDPC Algorithm

The setup phase is primarily composed of two steps: Cluster assembling and Route setup. When clusters are assembled and route tables are established, then every active NCHN can send data packets to its cluster head using specified band and time slot in a TDMA manner. At the same time, every CHN sends data packets to its upstream CHN according to its route table as established in the stage of route setup. CHNs operate in a CSMA/CA protocol for inter-cluster communication.

#### A. Cluster Assembling

The algorithm of the cluster assembly is shown in Algorithm (1): In this stage, all deployed nodes compete for being CHNs in CSMA/CA protocol. At first, a node has to wait a period of backoff time as defined in equation (2). If no broadcast packets are received from other CHNs, it will declare itself as a CHN by broadcasting a CHN declaration packet with transmission power P1 to its one-hop neighboring nodes.

**Algorithm (1)-Cluster Assembly**

1. BS broadcasts a Round Beginning packet with the expected Coverage ratio to all nodes
2. For all nodes in the wireless sensor network
3. Calculate its backoff time for competing
4. Backoff time count down
5. While (Backoff time is not expired)
6. If (there is a CHN declaration packet received)
7. Save the CHN info into the Priority Queue using the signal strength
8. Endwhile
9. If (Backoff time expired AND there is no CHN declaration packet received)
10. Node announces itself as a CHN by broadcasting a CHN declaration packet with transmission power P1
11. Calculate the amount of NCHNs, say N, needed for the coverage ratio
12. while N ≠ 0 { 
13. If (there is a JOIN message sent from a Node, say Node,
14. Assign a time slot for the new NCHN
15. Reply with a PERMISSION message with the slot number to
16. N = 1
17. Endif
18. If (Route setup message sent from the BS received)
19. Break this routine
20. Endif
21. If (there is a CANCEL message sent from a NCHN)
22. Remove the slot number to delete the NCHN
23. N +1
24. Endif
25. }
26. Broadcast a COMPLETION message to declare the cluster is assembled
27. Else
28. While (Priority Queue is not empty)
29. Dequeue to get the info of CHN
30. Send a JOIN message to the node
31. Waiting time count down
32. While (Waiting time is not expired)
33. If (PERMISSION message is received)
34. Receive and record the time slot assigned
35. Break this routine
36. Endif
37. If (Time Out message sent from the BS received)
38. Sleep in this round
39. Endif
If a node received CHN declaration packet from other CHNs during the backoff time, it becomes a NCHN candidate in the current round. A NCHN candidate may receive several CHN declaration packets during the backoff time. It will store the messages into a Queue according to the order of message signal received. When the backoff time is expired, the NCHN candidate will request to join a cluster by sending a JOIN message to the CHN that recorded in the front of Queue.

On the other hand, the selection of NCHNs of a cluster begins by the sending of CHN declaration packet. The CHN will calculate the amount of NCHNs needed to achieve the required coverage at first. When the required number of NCHNs is achieved, the CHN will stop the collection process and broadcast COMPLETION message to notify other nodes that are still stuck in backoff time to abandon the CHN. Of course, a node receiving a COMPLETION message does not mean that the node cannot be a NCHN in the current round. It can still request to join other clusters when its backoff time expired. This method will guarantee that the members of a cluster are all recruited with the NCHNs with more energy that the range can provide.

B. Route Setup

The task of this step is to establish multi-hop route from every CHN to the BS so as to ensure network connectivity. In KDPC, multi-hop routes among CHNs are established based on the minimum hop count, and the band of each cluster for intra-cluster communication in the working phase is determined so that the bands of any neighboring clusters are different from each other. Here, neighboring clusters are defined as the clusters that their CHNs can communicate directly with transmission power P_2.

The process of route setup is described as follows: The BS broadcasts a HOP packet with transmission power P_2. The neighboring CHNs receive the HOP packet, which consists of the serial number of the CHN, the minimum hop count to the BS, the total number of nodes in the cluster, and the band for intra-cluster communication in the working phase. Each CHN waits a period of time after it receives the HOP packet. CHN may receive several HOP packets during the time. It will select randomly a CHN with minimum value of hop count as its own upstream CHN to BS. When the time is expired, the CHN rebroadcasts the HOP packet by adding one to hop count value to downstream CHNs for following hop connections.

In the working phase, every active NCHN sends data packets to its cluster head using specified band and time slot in a TDMA manner, and sleeps in other time slots. Every CHN sends data packets to its upstream CHN according to its route table established in the stage of route setup.

The major differences between KDPC and ARC algorithms include: The rule used for determining a node to active as a NCHN or just sleep in this round is different. The amount of NCHN for each CHN is different from ARC since the equations for calculating the amount are different. Furthermore, the equations that determine backoff time of each node are different, too. To verify the practicality of KDPC, both ARC and KDPC are simulated for comparison. For the calculation of backoff time in simulation, the weighted ratio, T1:T2, used in ARC is 7:3. For KDPC, the three weighted values, T1, T2, and T3 are 1:6:3. The simulation results reported in the study are all deduced by simulating each scenario 5 times to get the average values as the results. All simulations are designed by JAVA language with Eclipse development environment.

In the simulations, there are 1000 nodes uniformly and independently distributed in a deployment range. The sink node is located at the location of a corner with coordinate (0, 0). The sensing range is set as a circle with radius of Rs. The communication ranges for intra-cluster and inter-cluster are circles with radius of Rc1 and Rc2, respectively. We adopt the energy model in [15] and assume that the energy cost of overhearing a packet equals to that of receiving a packet. The simulation parameters are set the same as the ARC made in [14]. The related parameters are listed in Table I.

Figure 1 is the average coverage ratio achieved in the first 2000 rounds. The x-axis is the rounds simulated, and the y-axis is the average coverage rates achieved. We can observe from the figure that the coverage ratio of ARC begins to drop down around 1500 rounds. For the last round, the coverage ratio is dropped below 0.7. On the contrary, KDPC can prolong the required sensing quality until the last 100 rounds. The dropped ratio achieved is about 0.75 in the last round.
IV. MODIFICATION OF KDPC AND SIMULATION RESULTS

For evaluating the effect of KDPC over the traditional cluster based WSN, a preliminary simulation was made by comparing with LEACH and ILEACH. For the simulations for KDPC, the expected coverage rate is set as 99.9% to gain best coverage rate.

Figure 2. exhibits the coverage rate of LEACH, ILEACH, and KDPC under the same assumptions. From the figure we can find that ILEACH can indeed improve the effectiveness of LEACH. However, the improvement is much less than the performance achieved by KDPC. Although KDPC cannot achieve 100% coverage rate in the beginning since the partial coverage strategy, the method can prolong effective service time more than twice than other two methods. By assuming 80% coverage rate as a threshold of effective service, readers can observe from the figure that the coverage rate of LEACH will drop down below 80% at 246th round, ILEACH begin drop below 80% at 386th rounds. On the other hand, the effective service can be prolonged to 852th rounds for KDPC. Furthermore, the drop curve of KDPC is much smooth than LEACH and ILEACH. The truth means that KDPC can effectively use the power potential of every sensor nodes to achieve a graceful degradation for long time usage.

![Figure 2. Coverage rates achieved by LEACH, ILEACH, and KDPC](image1)

By examining the status of each round in simulations, we found that the exhausting of energy always happens from the innermost nodes to outermost. In the end of simulations, we always find that most of outmost nodes are still alive. The situations are induced by two major reasons: First, since the nodes in the innermost layer will spend more energy to transmit the packets that came from the outer ranges because of multi-hop transmission. The truth makes the power exhausted rapidly for the innermost nodes when the node acts as a CHN. Second, the CHNs located in the middle range of the area will spend more energy to collect the data of its NCHNs since their responsibility area are bigger the CHNs which located nearby the boundary of the sensing area. For evaluating the improvement of KDPC for the conditions that the WSN can deploy nodes by uneven manner, we divide the area into three regions to deploy sensor nodes with different density. This study find that by distributing sensor nodes in the innermost area, the area in the area nearby the BS with distance less than Rc2, with 40% more than average distribution. By distributing 20% more sensor nodes in the region of middle area, the area far from the boundary with distance Rc2. The increasing of deployment density will lead the density of outer area to be decreased by the same number of sensor nodes used in simulation. The effective service time will be prolonged 7.4% more than even distribution as shown in figure 3.

![Figure 3. Improvement of KDPC by uneven deployment](image2)

For immigrating KDPC to play as a NEAR "complete coverage" strategy, the expected coverage rate is set to be 99.9% during the whole simulation process. The value will lead KDPC to collect sensor nodes as more as possible to fit the requirements of clusters. The condition may waste sensor nodes for useless overlapping since the actual coverage is drop down. Thus we further modify the version of KDPC with uneven deployment to change the expected coverage rate according to the actual coverage rate got by the last run. When the actual rate is drop below the expected value for 5%, the expected coverage rate used for next round will be decreased by 5%. The modification will make KDPC be graceful degradation for the coverage and lead the usage of nodes to be more realistic. Figure 4 shows the simulation results for the modification by comparing with the original design. The effective service time will be improved from 7.4% to 10.3% more than original KDPC.

![Figure 4. Improvement of KDPC by graceful degradation](image3)

V. CONCLUSION

This paper proposes a clustered topology control algorithm for large-scale wireless sensor networks with randomly deployed nodes. The algorithm is designed with novel feature: Less redundant nodes are activated than the existing algorithms since the number of sensor nodes used to achieve the required coverage ratio will be determined and limited by the algorithm. It is suitable for practical applications of large-scale or high-density wireless sensor networks due to its distributed processing and scalable cluster topology. According to the
simulation results that are made by comparing with the classic clustered topology, LEACH, and an improved method proposed recently. The study find that the method proposed in this paper can double the lifespan of a wireless sensor network under the same condition.

There are many issues worthy of investigation. For example, an interesting condition is found from the simulation that if the distribution of sensor nodes can be deployed unevenly. The lifespan of the scheme will be extended efficiently. As another example, the messages passed through CHNs to BS are transmitted by a data-centric routing paradigm, an energy-efficient routing protocol that is worth proposing for multi-source transmission scenarios. An efficient routing protocol can significantly reduce network traffic, and thus promote energy efficiency.

REFERENCES