Exploring Wireless Sensor Network Configurations for Prolonging Network Lifetime

Zhi Zhai† and Yuhua Chen††,

Department of Electrical and Computer Engineering, University of Houston, Houston, TX, USA

Summary

In this paper, we propose to control the distances between the cluster heads, as well as the cluster sizes to extend the network lifetime in wireless sensor networks. To the authors' best knowledge, this is the first paper that uses unequal transmission distances between cluster heads to solve the "hot spot" problem. More specifically, we take advantage of the fact that the transmission energy varies greatly as the transmission distance change. By using unequal transmission distance between cluster heads, as well as unequal cluster sizes, we are able to obtain sensor network configurations that significantly extend the network lifetime.

Key words:
Sensor network, Genetic Algorithms, network lifetime, energy

1. Introduction

To keep specific areas under surveillance, wireless sensor networks (WSNs) deploy hundreds or thousands of integrated sensor nodes to sample data from observed environment. In practice, due to the large quantity of sensor nodes or harsh working conditions, it is infeasible or unadvisable to recharge the batteries in WSNs. Therefore, sensor network lifetime is a primary concern in sensor network design.

Clustering [1] is a commonly adopted approach in sensor networks to manage power efficiently. In clustering, sensors in the monitoring area are grouped into clusters, as illustrated in Fig. 1. All sensor nodes within the same cluster send their data to the cluster head, which then forwards possibly aggregated data to the base station. Depending on the capability of sensor nodes and cluster heads, we can categorize the sensor networks into homogeneous sensor networks and heterogeneous sensor networks [2]. In homogeneous sensor networks, all sensors are equipped with the same hardware and software, in which case cluster heads can be rotated among different sensor nodes [3][4]. However, this approach requires all sensor nodes use expensive hardware to perform the cluster head functionality. In heterogeneous sensor networks, the cluster heads have more sophisticated hardware and software, and more battery power than normal sensor nodes [2][5][6]. Due to the potentially large quantity of sensor nodes dispersed in the monitoring areas, we believe that the heterogeneous sensor network model is more practical, therefore, is used for the rest of the paper.

The communications between the cluster heads and the base station (sink) can be classified into single-hop and multi-hop communication [2]. In single-hop communication, the cluster heads can reach the base station directly. In multi-hop communications, the cluster head transmits data to the neighboring cluster head, which relays data to the base station. Because multi-hop transmission is more energy efficient than single-hop transmission [7], we will use multi-hop communication between cluster heads as our network model.

To form clusters, most literatures divide the monitoring area into clusters of approximately the same size, which is called *Equal Clustering Size* (ECS) [2]. For example, in [4], the sensing area is divided into equal size clusters called cells. Based on the network density and transmission parameters, optimal cell size was derived. However, due to the unbalanced load among clusters in forwarding data to the base station, the cell size optimized for the total network energy consumption does not necessarily lead to the maximum network lifetime. The reason is that in multi-hop communication, the cluster heads closer to the base station relay more data and consume more energy. As soon as these cluster heads use up their energy, the whole sensor network loses its coverage. This is called the "hot spot" problem.

Instead of dealing with the problem of unbalanced load in sensor networks, authors in [2] tackles the problem using an approach called *Unequal Clustering Size* (UCS), which partitions the sensing area into clusters of unequal sizes, in order to balance the energy consumed by different

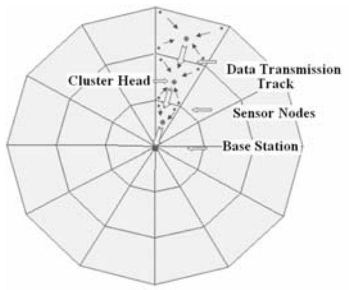

Fig. 1 Sensor network clustering

Manuscript received August 5, 2007
Manuscript revised August 20, 2007

cluster heads. In [2], cluster sizes are determined by the intra-cluster communication traffic relative to the expected inter-cluster communication traffic. The authors concluded that in order to balance the energy dissipation among cluster heads, the cluster heads closer to the base station should handle less intra-cluster communication, and therefore, should have smaller cluster size. By partitioning the sensing area into two layers, and setting the constraint of equal energy consumption for all clusters, the authors can mathematically determine the size of the clusters.

However, the approach in [2] has the following limitations. First, it only solved the problem for a two-layer sensor networks. Due to the maximum transmission range of the cluster heads, the two-layer approach might not be practical for large monitoring areas. Secondly, authors in [2] focused on the energy consumed on intra-cluster communication with respect to inter-cluster communication. They did not consider one important fact that the energy spent on inter-cluster communication is heavily influenced by the distance between the adjacent cluster heads.

To the authors' best knowledge, it is the first paper that extends the Unequal Clustering Size approach to multi-layer (>2) sensor networks. More importantly, we take advantage of the fact that the transmission energy varies greatly as the transmission distance change. By using unequal transmission distance between cluster heads, we are able to balance the energy consumption among different cluster heads. In particular, we use *Genetic Algorithm* (GA) to solve the multi-layer cluster partitioning problem. By controlling the cluster sizes, as well as the distances between cluster heads, we are able to obtain sensor network configurations that significantly extend the network lifetime.

The remainder of paper is organized as follows: In Section 2, we establish our network model. Section 3 describes the background of GA. In Section 4, we use the GA method to obtain sensor network configurations that prolong the network lifetime. We conclude our work in Section 5.

2. Network Model

The monitoring area is assumed to be a circled area with the base station in the center. We use multi-hop communication between cluster heads and the base station. Intermediate cluster heads act as relaying nodes. The assumptions in the proposed model are listed below:

1) Sensor nodes are uniformly distributed in the observed area with node density *D*.

2) Cluster heads are charged with the same amount of initial energy.

3) Cluster heads are approximately located in the center of each cluster.

4) We restrict the minimum and maximum distances between the cluster heads to be 5 m, and 300 m, respectively.

5) We assume all sensor nodes have the same duty cycle (sleep mode vs. active mode).

6) The network configuration is determined prior to the deployment.

Energy consumption on cluster heads can be divided into three categories: 1) *Aggregation Energy Consumption* E_a; 2) *Reception Energy Consumption* E_r; 3) *Transmission Energy Consumption* E_t [2]. For a given cluster head, its aggregation energy consumption E_a depends on how many sensor nodes it has and how much data each sensor node sends. E_a can be obtained as follows:

$$W_{a,i} = w * N_i \qquad 1 \le i \le n \tag{1}$$

$$N_i = A_i * D \qquad 1 \le i \le n \tag{2}$$

where $W_{a,i}$ is the amount of data (in bits) collected by cluster head i; w is the amount of data sent by one sensor node; N_i is the number of sensor nodes in the cluster i; D is the sensor node density; A_i is the area covered by cluster i; and n is the number of cluster heads in the transmission path. The area can be calculated by

$$A_1 = R_1^{\,2} * \frac{\theta_{rad}}{2} \tag{3.a}$$

$$A_i = (R_i^{\,2} - R_{i-1}^{\,2}) * \frac{\theta_{rad}}{2} \qquad 1 < i \le n \tag{3.b}$$

where R_i is the radius of the cluster and θ_{rad} is the radian of the sector, as shown in Fig. 2. So, the aggregation energy is:

$$E_{a,i} = e_a * W_{a,i} = e_a * w * A_i * D\,, \quad 1 \le i \le n \tag{4}$$

where e_a is the energy consumption for aggregating one bit of data. Similarly, the reception consumption E_r can be obtained as follows.

$$W_{r,i} = W_{t,i+1}\,,\ 1 \le i < n \tag{5.a}$$

$$E_{r,i} = e_r * W_{r,i}\,,\ 1 \le i \le n \tag{5.b}$$

where e_r is the energy consumption for receiving

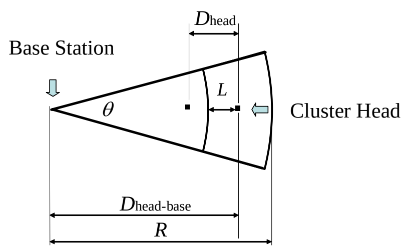

Fig. 2 Definition of distance variables

one bit of data; $W_{r,i}$ is the data amount transmitted by immediate upstream cluster head. Note that $W_{r,n}$ equals 0. The typical values for e_a and e_r are 5nJ/bit and 50nJ/bit respectively [2]. $E_{t,i}$ is related to the distance between clusters heads. According to [2] and [8], we obtain the expression of $E_{t,i}$ as follows:

$$W_{t,n} = W_{a,n} \qquad (6.a)$$

$$W_{t,n} = W_{a,n} \qquad (6.b)$$

$$E_{t,i} = (e_r + e_t * d_{heads}(h_{i-1}, h_i)^p) * W_{t,i} \qquad (6.c)$$

The value of constant e_t is given as 0.001pJ/bit/m^4, and p equals 4. The total energy consumption on a cluster head is:

$$E_i = E_{a,i} + E_{t,i} + E_{r,i},\ 1 \le i \le n \qquad (7)$$

The lifetime of cluster head i can be calculated as:

$$T_i = \frac{E_{initial}}{E_i * f},\ 1 \le i \le n \qquad (8)$$

where $E_{initial}$ is the initially charged energy, and f is the frequency of the data collection activity. In order to determine the distances between the cluster heads, and the size of the clusters, we first define the *Expected Energy Consumption* to be

$$E_{p,i} = \frac{1}{n} * \sum_{i=1}^{n} E_i * (V_l + \frac{V_h - V_l}{n-1} * (i-1)) \qquad (9)$$

where V_h is the highest energy dissipation rate, and V_l is the lowest energy dissipation rate among cluster heads in a given transmission trace. The assignment of these two parameters depends on the expiration sequence requirements in specific applications. The distance variables D_{heads}, $D_{head-base}$ and radius R are defined in Fig. 2. The distance variables have the following relationship:

$$D_{heads,1} = L_1 \qquad (10.a)$$

$$D_{heads,i} = L_{i-1} + L_i,\ 1 < i \le n \qquad (10.b)$$

$$D_{head-base,i} = 2 * \sum_{m=1}^{i-1} L_m + L_i \qquad (11)$$

$$D_{heads,i} = L_{i-1} + L_i,\ 1 < i \le n \qquad (12)$$

The angle θ can be calculated using

$$\theta = \frac{2 * L_n}{D_{node-base,n}} \qquad (13)$$

In order to keep the actual energy consumption as close to the expected energy consumption as possible, in this paper, we propose to use the variance of the actual energy consumption to the expected energy consumption as the objective function:

$$F = \sigma^2 = \frac{1}{n} * \sum_{i=1}^{n} [(E_i - E_{p,i})^2] \qquad (14)$$

E_i : tentative energy consumption on the i^{th} cluster head

$E_{p,i}$: expected energy consumption on the i^{th} cluster head

The total energy consumption can be calculated as the summation of the energy dissipation of all cluster heads:

$$E_{total} = \sum_{i=1}^{n} E_i \qquad (15)$$

3. Generic Algorithm

In this paper, we explore variable number of clusters in each layer, as well as variable distances between cluster heads, in which case, the optimal configuration cannot be easily mathematically computed. Therefore, we use *Genetic Algorithm* (GA) to obtain optimal multi-layer network configurations. GA belongs to stochastic search methods and its implementation is an iterative procedure. As illustrated in Fig. 3, there is an initial pool of possible solutions, whose fitness can be evaluated according to the *fitness function* (objective function). By using genetic operators such as *crossover* and *mutation*, parent solutions can proliferate and generate next generation solutions. From generation to generation, the quality of the candidate solutions is improved by the "*survival of the fittest*" scheme. This process is repeated until an optimal or near optimal solution is obtained. For a detailed description of GA, readers can refer to [9] and [10]. We describe some fundamental terminologies as follows.

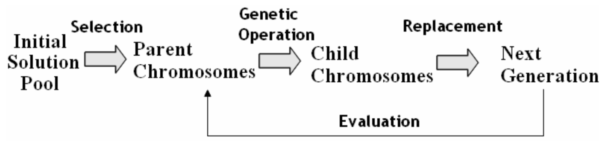

Fig. 3 Illustration of Genetic Algorithm

***Chromosome*:** Chromosome is a string representation of the candidate solution, as illustrated in Fig. 4.

***Crossover*:** Crossover is a genetic operator to generate children chromosomes, in which some portions of parents are exchanged. This operator simulates the

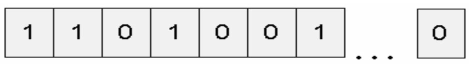

Fig. 4 Chromosome representation

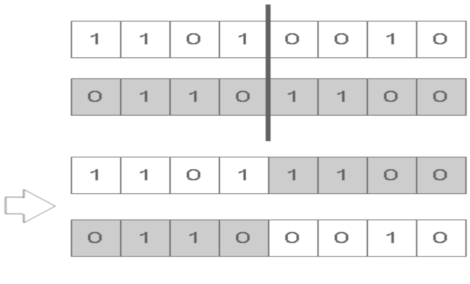

Fig. 5 Example of crossover operation

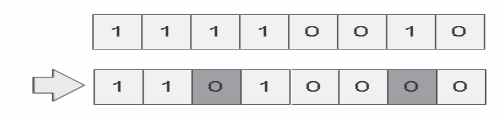

Fig. 6 Example of mutation operation

reproduction process in the biological evolution [10]. An example of crossover is shown in Fig. 5.

***Mutation*:** Mutation operator generates new possible solutions by introducing small disturbance to the chromosomes. For example, in Fig. 6, the '1's in the 3rd and 7th bit are flipped to '0's while the remaining bits are kept untouched.

***Fitness function*:** Fitness function (objective function) is a metric to evaluate the quality of the candidate solutions. The most promising solutions are selected from the pool of candidate solutions according to the fitness function.

Population size: Population size is the total number of candidate solutions in one generation**.**

As mentioned above, the strength of the GA lies in its "survival of the fittest" scheme. The GA tends to keep the most promising solutions and prune low-quality ones. Therefore, it is especially suitable for solving multi-variable problems, where there is often a complicated interference among multiple variables.

In this paper, we demonstrate that we are able to find high quality wireless sensor network configurations in large search space using the evolutionary selection process.

4. Genetic Algorithm Results

Based on the network model we have developed in Section 2, we can use GA to obtain sensor network configurations which allow the cluster heads to use up their energy at approximately the same time. In this section, we obtain the network configuration in seven scenarios, and compare the sensor network lifetime among different configurations.

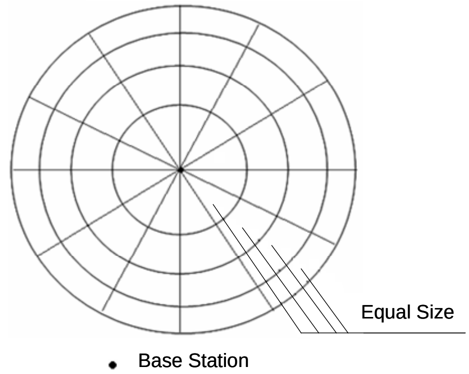

Fig. 7 Scenario 1, traditional equal clustering size (ECS)

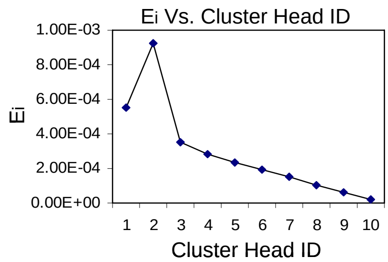

Fig. 8 Scenarios 1, actual energy consumption (ECS reference system)

The following values apply to all configurations unless specified otherwise: The monitoring area is divided into 10 layers. The total number of sensor nodes is 8000. Every sensor node sends 80 bits of data per second to its cluster head. The initial energy storage charged in each cluster head is 100 J and the radius of surveillance circle is 500 m. In GA, the population size in one generation is 1000 and the number of generation is also set to 1000.

4.1 Scenario 1 (Reference Configuration)

In Scenario 1, we use the traditional *Equal Clustering Size* model shown in Fig. 7 as our reference configuration. In this scenario, all parameters are deterministic, and are mathematically computed as follows.

Since all clusters have equal sizes, the distance between cluster heads and base station can be obtained as follows.

$$A_i = A/10 \text{ , } 1 \le i \le 10 \tag{16}$$

$$R_i = \sqrt{i * \frac{2}{\theta} * \frac{A}{10}} \text{ , } 1 \le i \le 10 \tag{17}$$

Using the value R_i, we can get values of L_i by

$$L_1 = \frac{R_1}{2} \tag{18.a}$$

$$L_i = \frac{R_i - R_{i-1}}{2} \text{ , } 1 < i \le 10 \tag{18.b}$$

Formula (10) and (11) are used to calculate D_{heads} and $D_{head-base}$.

The energy consumption in Scenario 1 is plotted in Fig. 8. As expected, in Scenario 1, the cluster heads closer to the base station consume much more energy than the cluster heads further away. As a result, cluster heads closer to the base station will use up their stored energy sooner, resulting in the "hot spot" problem.

4.2 Scenario 2

In Scenarios 2, each layer contains the same number of clusters as illustrated in Fig. 9. We use GA to determine the optimal distance between cluster heads.

Fig. 10 shows the results obtained by GA. The results show that in order to balance the energy consumption, cluster heads closer to the base station should be placed closer to each other, while cluster heads further away from the base station should be placed further apart. The results match our intuition. Note that the shorter the distance, the less energy spent on transmission. Since the cluster heads closer to the base station have to relay more data, in order to balance the energy consumption, they should have shorter transmission distance so that the energy spent on transmitting each bit of data is less.

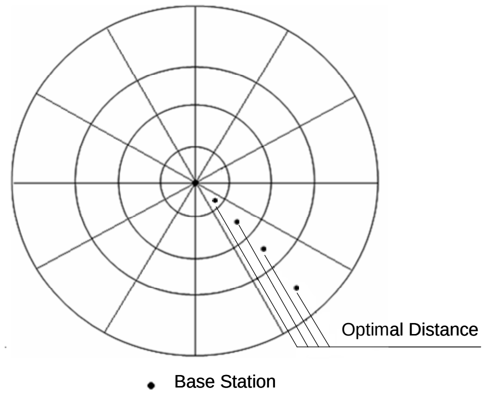

Fig. 9 Scenario 2, variable distance between cluster heads

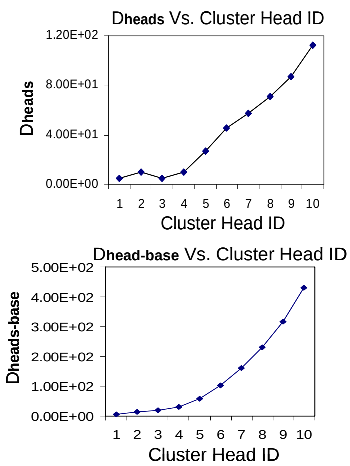

Fig. 10 Scenario 2, distances between cluster heads and base station

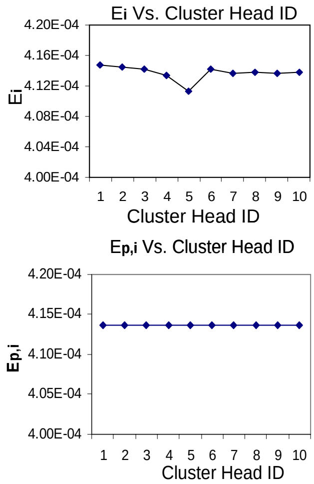

Fig. 11 Scenario 2, actual and expected energy consumptions, same expiration time

In Fig.11, we can see that the actual energy consumption matches well with the expected energy consumption. In this case, the cluster heads will expire at approximately the same time.

In the second case, we set the expected energy consumption to have gradient values among cluster heads as shown in Fig. 12, where V_h and V_l in (9) are set as 1.05 and 0.95 respectively. As expected, the network loses its coverage gradually from periphery to center.

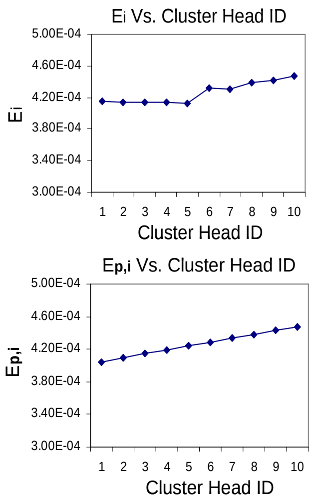

Fig. 12 Scenario 2, actual and expected energy consumptions, gradient expiration time

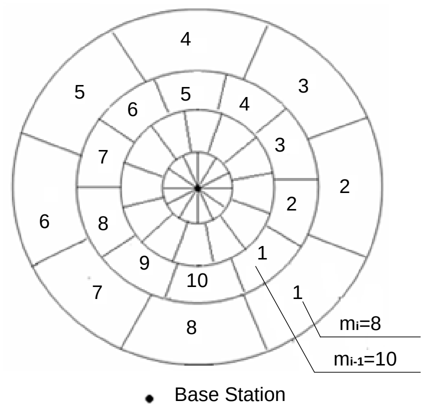

Fig.13 Scenario 3, variable number of clusters per layer, variable distance between cluster heads

4.3 Scenario 3

In Scenario 3, we allow variable number of clusters in each layer, as well as variable distance between the cluster heads, as shown in Fig. 13. Let m_i be the number of cluster heads in Layer *i*, we have

$$m_i = Round(N_{total} * Ratio_i) \qquad (19)$$

$$\sum_{i=1}^{10} Ratio_i = 1 \qquad (20)$$

where $Ratio_i$ is the ratio of the number of cluster heads in the i^{th} layer to the total number of cluster heads, and N_{total} is the total number of cluster heads.

Since the cluster number *m* in each layer can be different, the calculation for $W_{r,i}$ in (5.a) is replaced by

$$W_{r,i} = W_{r,i+1} * \frac{m_{i+1}}{m_i} \qquad (21)$$

In GA, a total of 20 variables are deployed, where the first 10 variables represent the number of clusters in each layer and the next 10 variables represent the distance between the cluster heads. In particular, the first 10 variables are assigned as the ratio of the number of cluster

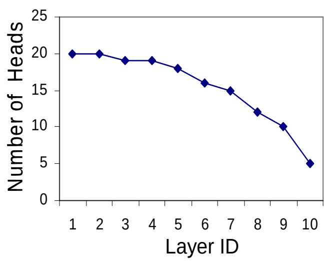

Fig: 14 Scenarios 3, number of cluster heads in each layer

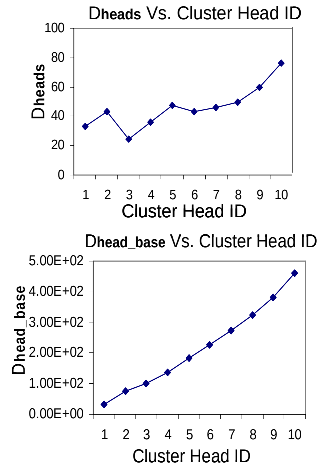

Fig. 15 Scenario 3, distance between cluster heads and base station

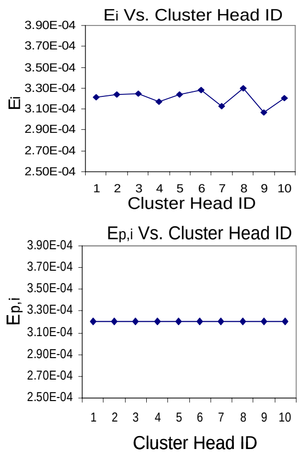

Fig. 16 Scenario 3, actual and expected energy consumption

heads in each layer to the total number of cluster heads. To avoid impractical solutions, we limit the number of clusters in each layer to be in the range of [5, 20]. Fig. 14 shows the resulting number of clusters in each layer. The distances between the cluster heads are shown in Fig. 15. Note that in Fig. 15, the increase in distance between cluster heads and the base station is more balanced compared to the results in Scenario 2. In addition, Fig. 16 shows that the actual energy consumption is also more balanced among cluster heads compared to that in Scenario 1. As a result, there is less residual energy in the network when the network expires. The average energy consumption of the cluster heads in Scenario 3 is about 23% less than that in Scenario 2, which means the network is Scenarios 3 has a longer lifetime compared to Scenario 2.

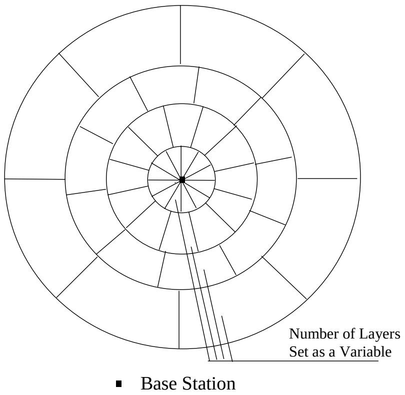

Fig. 17 Scenario 4, variable number of layers

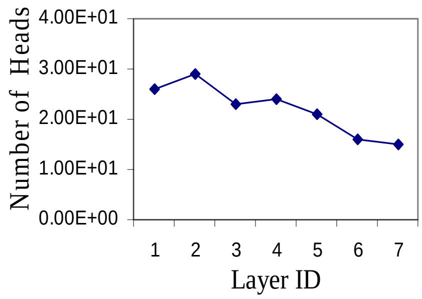

Fig. 18 Scenarios 4, number of cluster heads in each layer

4.4 Scenario 4

In the previous scenario, we set the distance between cluster heads and the number of clusters in each layer as variables. In Scenario 4, we set the number of layers in the surveillance area as a variable as well, as illustrated in Fig.17.

The results obtained by GA are presented in Fig. 18-20. Fig.18 shows the number of clusters in each layer. Fig. 19 shows the distance between the cluster heads and the base station. The actual and the expected energy consumption in the cluster heads are shown in Fig. 20. We notice that the worst energy dissipation rate has been reduced and the average energy consumption in cluster heads has been decreased too, which is better compared to results in Scenario 3.

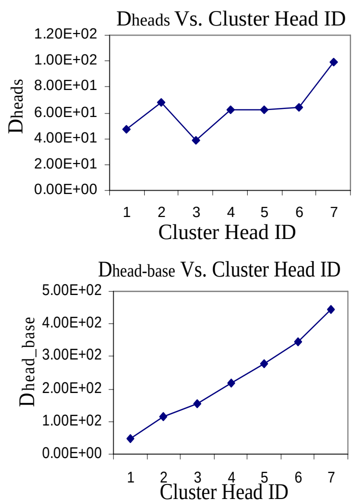

Fig. 19 Scenarios 4, distances between cluster heads and base station

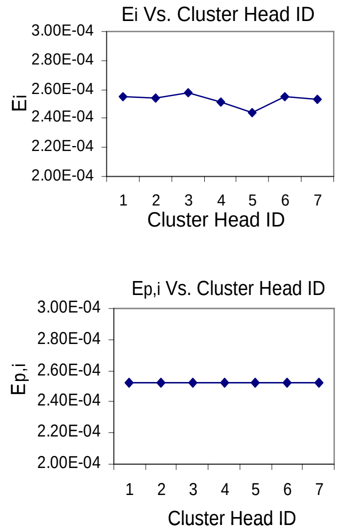

Fig. 20 Scenarios 4 actual and expected energy consumption

4.5 Scenario 5

In Scenarios 5, we investigate the quadrangle partition,

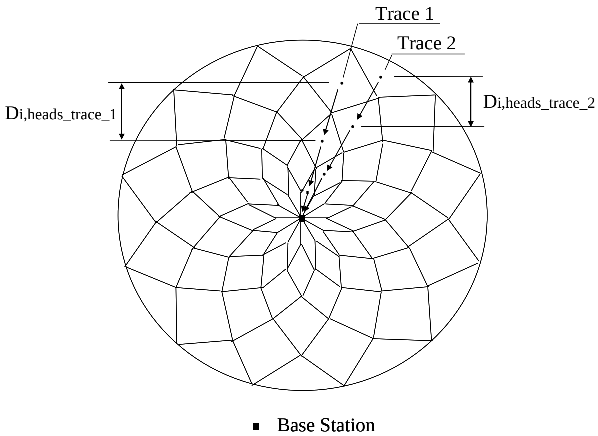

Fig. 21 Scenario 5, quadrangle partition

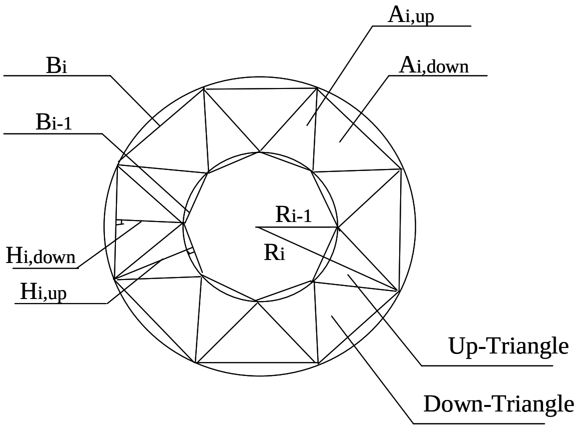

Fig. 22 Up-triangle and down-triangle configuration

where each irregular quadrangle shape consists of two component triangles as shown in Fig. 21. In this topology, each layer contains the same number of clusters and there are two paths that a cluster head can use to reach the base station. Since there are more nodes near the base station in this topology, the transmission burden on the cluster heads near to the base station can be reduced. In addition, since the number of clusters in each layer is no longer a variable in GA, we have greatly reduced the search space compared to Scenario 4.

Since the network configuration in Scenario 5 is different from the previous four scenarios, we use following formulas to calculate the transmission distance and the size of the area covered by the clusters.

$$C = Round(N_{total} / N_L) \quad (22)$$

$$\theta = 2\pi / C \quad (23)$$

N_L: Number of layers.
N_{total}: Total number of cluster heads.
C: Number of clusters in each layer.

Here, we define two types of triangles: if the vertex is closer to the base station than the base line in the triangle, it is called *Down-Triangle*; otherwise it is called *Up-Triangle* as shown in Fig 22.

$$H_{1,down} = R_1 * \cos(\theta/2) \quad (24.a)$$

$$H_{i,up} = R_i - R_{i-1} * \cos(\theta/2),\ 1 < i \le n \quad (24.b)$$

$$H_{i,down} = R_i * \cos(\theta/2) - R_{i-1},\ 1 < i \le n \quad (24.c)$$

$H_{i,down}$: Height of the i^{th} Down-Triangle
$H_{i,up}$: Height of the i^{th} Up-Triangle
R_i: Radius of the i^{th} layer

$$B_i = 2 * R_i * \sin(\theta/2),\ 1 \le i \le n \quad (25)$$

$$A_{i,down} = (H_{i,down} * B_i)/2,\ 1 \le i \le n \quad (26.a)$$

$$A_{i,up} = (H_{i,up} * B_{i-1})/2,\ 1 < i \le n \quad (26.b)$$

$$A_{n,down} = (H_{n,down} * B_n)/2 + \pi * R_n^2 / C - [R_n * \cos(\theta/2) * B_n]/2 \quad (26.c)$$

$A_{i,,down}$: Area of Down-Triangle in the i^{th} layer.
$A_{i,,up}$: Area of Up-Triangle in the i^{th} layer.
B_i : Base line of triangles in the i^{th} layer.

The area of an irregular quadrangle is the sum of the covered area in two sub-triangles as follows.

$$A_{i,trace\_1} = A_{2i,up} + A_{2i-1,down},\ 1 \le i \le \lceil n/2 \rceil \quad (27)$$

$$A_{i,trace\_2} = A_{2i,down} + A_{2i+1,up},\ 1 \le i \le \lfloor n/2 \rfloor \quad (28)$$

$A_{i,trace\_1}$: Area of the i^{th} cluster along trace 1.
$A_{i,,trace\_2}$: Area of the i^{th} cluster along trace 2.
n : Number of layers in the surveillance area

The distance between the cluster heads is the sum of the heights in the two triangles sharing the same vertex, which can be obtained by calculating the difference between the corresponding radius in (29)-(32). Here, the transmission distance in quadrangle configurations is the distance between the base line of the two triangles, whereas in next scenario, the transmission distance is between the two triangle centers.

$$D_{1,heads\_trace\_1} = R_1 * \cos(\theta/2) \quad (29.a)$$

$$D_{i,heads\_trace\_1} = D^1_{i,after} - D^1_{i,before},\ 1 < i \le \lceil n/2 \rceil \quad (29.b)$$

$D_{i,heads\_trace\_1}$: The length of the i^{th} transmission distance along trace 1

$$D^1_{i,before} = R_{2i-3} * \cos(\theta/2) \quad (30.a)$$

$$D^1_{i,after} = R_{2i-1} * \cos(\theta/2),\ 2i-1 \ne n \quad (30.b)$$

$$D^1_{i,after} = R_n - L_n,\ 2i-1 = n \quad (30.c)$$

The notation '1' in (29) and (30) means that we are computing the cluster head distance along trace 1, while '2' in

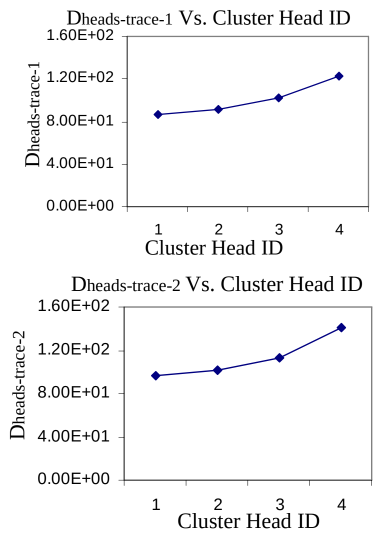

Fig. 23 Scenario 5, distances between cluster heads in trace 1 and 2

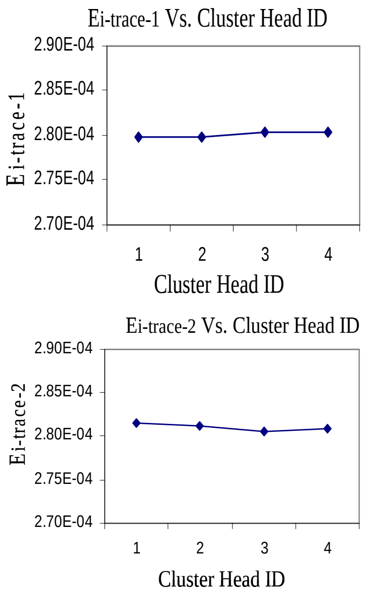

Fig. 24 Scenario 5, energy consumption in trace 1 and 2

(31) and (32) means trace 2.The definitions of R_n and L_n can be found in Fig. 2.

$$D_{1,heads\_trace\_2} = R_2 * \cos(\theta/2) \quad (31.a)$$

$$D_{i,heads\_trace\_2} = D^2_{i,after} - D^2_{i\ before}, 1 < i \le \lfloor n/2 \rfloor \quad (31.b)$$

$$D^2_{i\ ,before} = R_{2i-2} * \cos(\theta/2) \quad (32.a)$$

$$D^2_{i\ ,after} = R_{2i} * \cos(\theta/2),\ 2i \ne n \quad (32.b)$$

$$D^2_{i\ ,after} = R_n - L_n,\ 2i = n \quad (32.c)$$

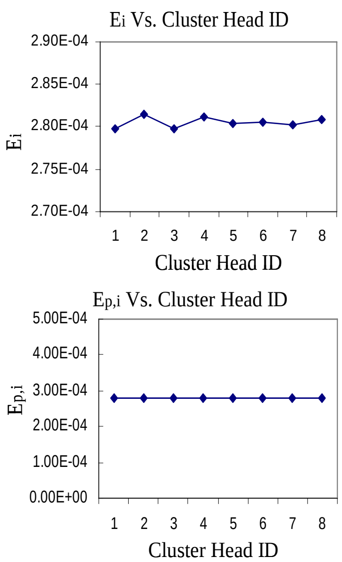

Fig.25 Scenario 5, actual and expected energy consumptions. (odd IDs: trace 1; even IDs: trace 2)

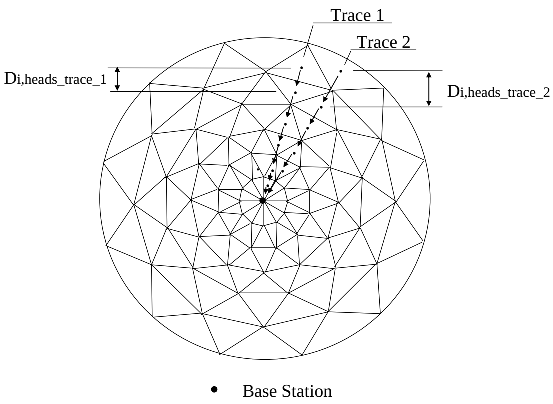

Fig. 26 Scenario 6, triangle partition

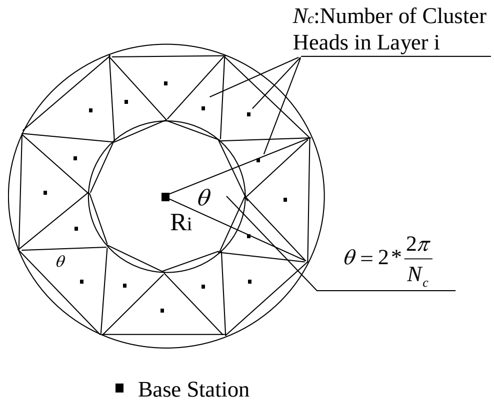

Fig. 27 Computation of θ in Scenario 6

$D_{i,heads\_trace\_2}$: The length of the i^{th} transmission distance along the trace 2.

Fig. 23 shows the distance between the nodes in trace 1 and 2 respectively. Fig. 24 shows the energy consumption along these two different paths. Fig. 25 shows the actual and the expected energy consumption of the cluster heads. Compared to Scenario 4, Scenario 5 has a much smaller search space, and therefore, requires a shorter time to obtain the solution.

4.6 Scenario 6

Based on the observations in Scenario 5, we can allocate more cluster heads around the base station to reduce the transmission burden on the cluster heads near the base station. We further investigate the following two configurations: the *Triangle Configuration* in Scenario 6 and the *Hexagon Configuration* in Scenario 7.

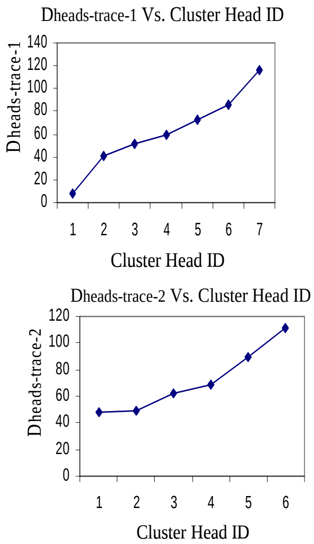

Fig. 28 Scenario 6, distances between cluster heads in trace 1 and 2

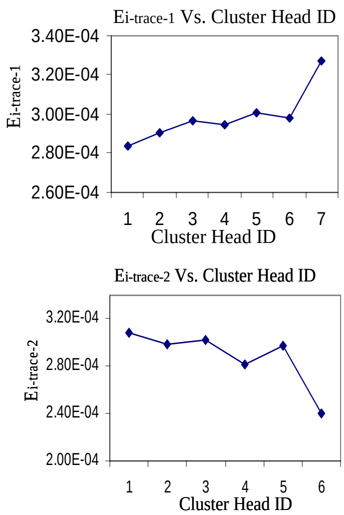

Fig. 29 Scenario 6, energy consumption in trace 1 and 2

In scenario 6, the basic shapes of the clusters are set as triangles, as shown in Fig. 26. When we configure the irregular quadrangle cluster shape in Scenario 5, we have already analyzed the height and the area of the triangle, which are also applicable in Scenario 6. The calculation of the cluster coverage area and the transmission distance in Scenario 6 is listed as follows:

$$A_{1,trace\_1} = A_{1,down} \quad (33)$$

If the cluster ID i is even, then

$$A_{i,trace\_1} = A_{i,up}\,,\ i > 1 \quad (34.a)$$

$$A_{i,trace\_2} = A_{i+1,up}\,,\ i \geq 1 \quad (34.b)$$

Otherwise,

$$A_{i,trace\_1} = A_{i,down}\,,\ i > 1 \quad (35.a)$$

$$A_{i,trace\_2} = A_{i+1,down}\,,\ i \geq 1 \quad (35.b)$$

$A_{i,trace\_1}$: Area of the i^{th} cluster along trace 1
$A_{i\ trace\_2}$: Area of the i^{th} cluster along trace 2

Here, the transmission distance in the triangle partition can be obtained by calculating the distance between the triangle centers, using the following formulas.

$$D_{1,heads\_trace\_1} = 0.5 * h_{1,down} \quad (36.a)$$

$$D_{1,heads\_trace\_2} = 0.5 * h_{2,down} + R_1 \quad (36.b)$$

If the transmission distance ID i is even, then

$$D_{i,heads\_trace\_1} = 0.5 * h_{i,up} + 0.5 * h_{i-1,down}\,,\ i > 1 \quad (37.a)$$

$$D_{i,heads\_trace\_2} = 0.5 * h_{i+1,up} + 0.5 * h_{i,down}\,,\ i > 1 \quad (37.b)$$

Otherwise,

$$D_{i,heads\_trace\_1} = 0.5 * h_{i,down} + 0.5 * h_{i-1,up}\,,\ i > 1 \quad (38.a)$$

$$D_{i,heads\_trace\_2} = 0.5 * h_{i,up} + 0.5 * h_{i+1,down}\,,\ i > 1 \quad (38.b)$$

$D_{i,heads\_trace\_1}$: Length of the i^{th} transmission distance along trace 1.

$D_{i,heads\_trace\_2}$: Length of the i^{th} transmission distance along trace 2.

In the triangle configuration, we use (39) instead of (23) to calculate θ as shown in Fig. 27:

$$\theta = 2 * (2\pi / C) \quad (39)$$

The GA results in this case are shown in Fig. 28-30. Fig. 28 shows the distance between nodes in the triangle configuration and Fig. 29 shows the energy consumption along different traces. Fig. 30 shows the actual and the expected energy consumption of the cluster heads in the surveillance area.

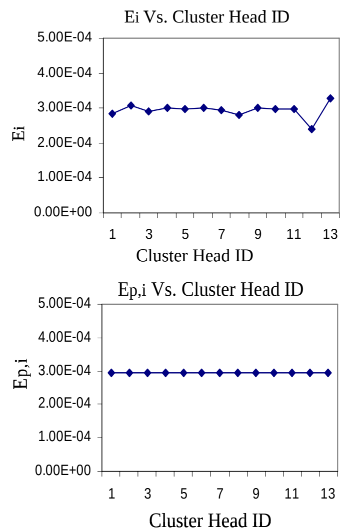

Fig. 30. Scenario 6, actual and expected energy consumption (odd IDs: trace 1; even IDs: trace 2)

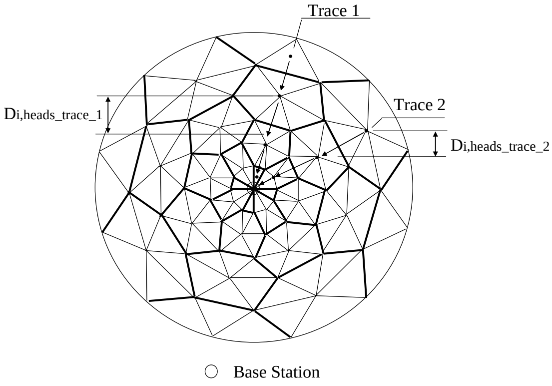

Fig. 31 Scenario 7, hexagon partition

4.7 Scenario 7

Based on the triangle partition in Scenario 6, the individual clusters can be combined together to form hexagons as shown in Fig. 31. The formulas to calculate the coverage area and the transmission distance in Scenario 7 are listed below.

$$A_{1,trace\_1} = 2 * A_{1,down} + 2 * A_{2,up} + A_{2,down} \qquad (40.a)$$

$$A_{i,trace\_1} = 2 * A_{2i-1,down} + 2 * A_{2i,up} + A_{2i-1,up} + A_{2i,down}, \quad 1 \le i \le \left\lceil n/2 \right\rceil \qquad (40.b)$$

$$A_{1,trace\_2} = A_{1,down} \qquad (41.a)$$

$$A_{i,trace\_2} = 2 * A_{2i-2,down} + 2 * A_{2i-1,up} + A_{2i-1,down} + A_{2i-2,up}, \quad 1 \le i \le \left\lfloor n/2 \right\rfloor + 1 \qquad (41.b)$$

$A_{i,trace\_1}$: Area of the i^{th} cluster along trace 1.
$A_{i,\, trace\_2}$: Area of the i^{th} cluster along trace 2.

The transmission distance in the hexagon case is the distance between the two hexagon centers.

$$D_{1,heads\_trace\_1} = R_1 \qquad (42.a)$$

$$D_{i,heads\_trace\_1} = R_{2i-1} - R_{2i-3},\ 1 < i \le \left\lceil n/2 \right\rceil \qquad (42.b)$$

$$D_{1,heads\_trace\_2} = 0.5 * R_1 * \cos(\theta/2) \qquad (43.a)$$

$$D_{2,heads\_trace\_2} = R_2 - 0.5 * R_1 * \cos(\theta/2), \quad 2 < i \le \left\lfloor n/2 \right\rfloor + 1 \qquad (43.b)$$

$$D_{i,heads\_trace\_2} = R_{2i-2} - R_{2i-4}, \quad 2 < i \le \left\lfloor n/2 \right\rfloor + 1 \qquad (43.c)$$

$D_{i,heads\_trace\_1}$: Length of the i^{th} transmission distance along trace 1.

$D_{i,heads\_trace\_2}$: Length of the i^{th} transmission distance along trace 2.

We set the hexagon as the basic shape of the clusters and apply the GA method to find optimal configuration. The number of layers and the width of the layers are set as

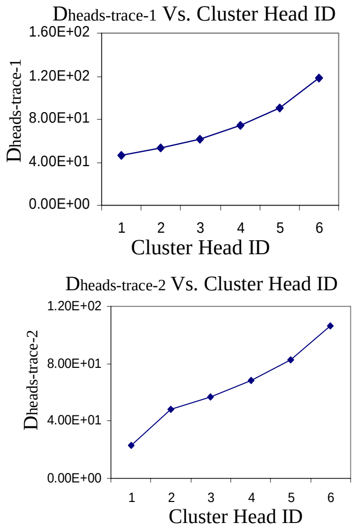

Fig. 32 Scenario 7, distances between cluster heads in trace1 and 2

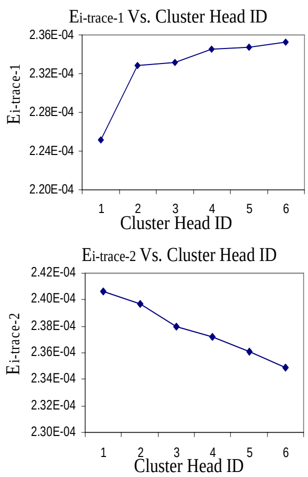

Fig. 33 Scenario 7, energy consumption in trace 1 and 2

variables. The formula to calculate the number of clusters in each layer is

$$\theta = 2\pi / (3 * C) \qquad (44)$$

The GA results are presented in Fig. 32-34. Fig. 32 shows the distance between the cluster heads. The energy consumption in different path is shown in Fig. 33. Fig 34 gives the energy consumption in each node.

4.8 Comparison of Network Lifetime

In this section, we compare the lifetime of the network in all 7 scenarios. We use the definition that the *lifetime* of the sensor network is the time period from the instant when the network starts to function to the instant when the first cluster head runs out of energy [11]. As shown in Fig. 35, Scenario 1 (traditional equal clustering size model) has the shortest lifetime and Scenario 7 has the longest

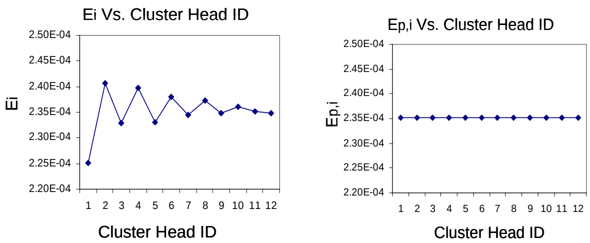

Fig.34 Scenario 7, actual and expected energy consumptions (odd IDs: trace 1; even IDs: trace 2)

lifetime. The lifetime in Scenario 2 is more than twice the lifetime in Scenarios 1. Compared to Scenario 2, Scenario 3 has 23.2% improvement. Scenario 4 has further 30.7% improvement compared to Scenario 3. The lifetimes in Scenario 5 and 6 are comparable to Scenario 4. The lifetime of the hexagon configuration in Scenario 7 is significantly longer than any other configurations.

5. Conclusion

In this paper, we have proposed to control the distances between the cluster heads, as well as the cluster sizes to balance the energy consumption in cluster heads. To the authors' best knowledge, this is the first paper that uses unequal transmission distances between cluster heads to solve the "hot spot" problem. In particular, we use *Genetic Algorithm* (GA) to minimize the variance of the actual energy consumption to the expected energy consumption to obtain optimal cluster configurations.

We have demonstrated that we can prolong the lifetime of the sensor networks by optimally configuring the network. The two main approaches are summarized as follows: 1) allowing variable size clusters, variable distance between the cluster heads, as well as variable number of layers (Scenario 1 – 3); 2) utilizing parallel paths to balance the work load in the cluster heads near the base station (Scenario 4 – 7). The lifetime of the hexagon configuration in Scenario 7 is more than five times longer than the traditional *Equal Clustering Size* model.

Acknowledgments

This work is partially supported by the University of Houston GEAR program.

References

[1] T. Kwon and M. Gerla, "Clutering with power control," *Proc. of IEEE Military Communications Conference (MILCOM)*, vol. 2, Atlantic City, NJ, Oct/Nov. 1999, pp. 1424-1428.
[2] S. Soro and W. Heinzelman, "Prolonging the Lifetime of Wireless Sensor Networks via Unequal Clustering," *Proc. of the 19th IEEE International Parallel and Distributed Processing Symposium (IPDPS'05)*, April 2005.

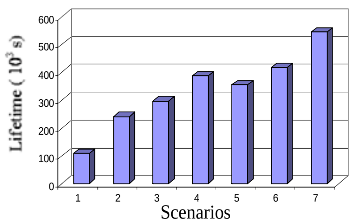

Fig.35 Network lifetime comparison

[3] W. Heinzelman, A. Chandrakasan, H. Balakrishman, "An Application-specific protocol architecture for wireless microsensor networks," *IEEE Transactions on Wireless Communications*, vol. 1, no. 4, Oct. 2002, pp. 660-670.
[4] Q. Xue, A. Ganz, "Maximizing sensor network lifetime: analysis and design guidelines," *Proc. of IEEE Military Communications Conference (MILCOM)*, vol. 2, Oct/Nov, 2004, pp. 1144 – 1150.
[5] V. Mhatre, C. Rosenberg, D. Kofman, R. Mazumdar and N. Shroff, "A Minimum cost heterogeneous sensor network with a lifetime constraint," *IEEE Transactions on Mobile Computing*, vol. 4, no. 1, Jan/Feb 2005, pp. 4 – 15.
[6] M. N. Halgamuge, K. Ramamohanarao, M. Zukerman, M., "High powered cluster heads for extending sensor network lifetime," *IEEE International Symposium on Signal Processing and Information Technology*, Aug. 2006, pp. 64 – 69.
[7] M. Perillo, Z. Cheng and W. Heinzelman, "On the problem of unbalanced load distribution in wireless sensor setworks," *Proc. of IEEE GLOBECOM Wireless Ad Hoc and Sensor Networks*, Nov. 2004, pp. 74-79.
[8] M. Bhardwaj and A. P. Chandrakasan, Bounding the lifetime of sensor networks via optimal role assignments, *Proc. of INFOCOM 2002*, vol. 3, New York, NY, 2002, pp. 1587-1596.
[9] M. Mitchell, *An Introduction to Genetic Algorithms*, MIT Press, reprint edition, February 6, 1998.
[10] D. Goldberg, *Genetic Algorithms in Search, Optimization, and Machine Learning*, Boston, MA: Addison-Wesley Professional, 1989.
[11] C. F. Chiasserini, I. Chlamtac, P. Monti and A. Nucci, "Energy efficient design of wireless ad hoc networks," *Proc. of European Wireless*, Florence, Italy, 2002.

Zhi Zhai received the B.S. degree in Electrical Engineering from Nankai University, China in 2005. He is working towards his M.S. degree in the Department of Electrical and Computer Engineering at the University of Houston, TX, USA. His research interests include sensor networks and optical networks.

Yuhua Chen received the B.S. degree in Electronic Engineering from Fudan University, China in 1995. She received the M.S. degree in computer science, the M.S. degree in Electrical Engineering, and the D.Sc. degree in Electrical Engineering from Washington University in St. Louis, USA in 1997, 1998, and 2003, respectively. She was a Research Associate at the Applied Research Laboratory (ARL), Washington University in St. Louis from 1999 to 2004. She is currently an Assistant Professor at the Department of Electrical and Computer Engineering, University of Houston, Houston, TX, USA. Her research interests include sensor networks, optical networks, and system-on-chip.