Cluster Size Optimization in Gaussian Distributed Wireless Sensor Networks

Vinay Kumar [#1], Sanjay B. Dhok [#2], Rajeev Tripathi [*3] and Sudarshan Tiwari [&4]
[#]Department of Electronics Engineering, Visvesvaraya National Institute of Technology,
Nagpur, Maharashtra, India
[1]vk@ece.vnit.ac.in
[2]sbdhok@ece.vnit.ac.in
[*]Department of Electronics & Communication Engineering, Motilal Nehru National Institute of Technology,
Allahabad, Uttar Pradesh, India
[3]rt@mnnit.ac.in
[&]National Institute of Technology, Raipur, Chhattisgarh, India
[4]director@nitr.ac.in

Abstract: To deal with sensor network limitations such as limited energy and short range communication, sensor nodes are grouped into mostly non overlapping subsets called clusters. Choosing optimal number of clusters provides benefits such that limited resources can be utilized more efficiently and network lifetime is improved. Many of the existing researches provided the cluster size optimization in Wireless Sensor Networks (WSNs), in which nodes are uniformly and randomly placed in the sensing field (e.g. controllable WSN). Deployment of sensor nodes affects the energy consumption of WSNs along with individual nodes because the distance between nodes and Base Station (BS) is different due to different node position; consequently nodes have different energy loss. The energy efficient way of sensor deployment in sensing field is controlled node deployment with uniform distribution. However this procedure for node deployment may not be practically possible for some applications like, in large WSNs, locations of the sensing field may not be physically accessible because of geographical constraints. In this paper, we provide an analytical framework for the cluster size optimization of WSNs that follow Gaussian node deployment. This type of node deployment reduces energy hole problem, provides enhanced intrusion detection capability and support realistic applications. We have provided expression for optimal number of clusters using circular sensing model of nodes for square sensing field with consideration of boundary effect. We have also compared the cluster size optimization for uniform and Gaussian distributed sensor network.

Keywords: Gaussian Node Distribution, Energy Efficiency, Optimal Clustering, Network lifetime, Wireless Sensor Networks.

I. INTRODUCTION

Energy efficiency and coverage are crucial quality of service management for different applications in WSNs**.** In WSN, sensor nodes are energy constrained because most of the times sensor nodes left unattended in hostile environment [1, 2, 3]. If WSNs follow random and uniform placement of nodes, due to short transmission range of sensor node data traffic flow is carried out through multi hopping. In this case a single sensor node works both as data originator as well as data router. Thus sensor node closer to the BS takes heavier traffic load via multi hop transmission leading to energy holes around the BS. Formation of energy holes means that data traffic can no longer be delivered to the BS on a specific path and the entire sensor network may not function properly [4]. So we can say that problem of unbalanced energy consumption and energy hole problem in sensor network are due to random and uniform placement. Gaussian distribution with proper adjustment of mean distance can provide more energy balance with removal of energy hole problem in sensor networks. Random WSNs may follow a Gaussian distribution or uniform distribution depending on strategy of sensor node placement. Gaussian distribution provides enhanced intrusion detection capability, support real time applications and in large scale WSNs it reduces the energy hole problem [5, 6, 7].

To support better data aggregation with high network scalability, sensor nodes are grouped into mostly non overlapping subsets called clusters. In the clustering technique, each cluster has a leader, which is called the Cluster Head (CH) and it performs the tasks like fusion and aggregation of data [8, 9, 10]. The major thought behind an optimal clustering (selecting the optimal number of clusters or Cluster Heads) is to determine a clustering of the network such that the entire energy required for aggregating data from the whole network is minimized as compared with other possible clustering patterns [12]. If the clusters are not constructed in an optimal way, the total consumed energy of the sensor network per round is increased exponentially either when the number of clusters that are created is greater or when the number of the constructed clusters is less than the optimal number of clusters [11]. Selecting an optimal number of clusters in WSNs provide greater improvement in terms of energy efficiency, system scalability, network lifetime, and latency [13]. The concepts of cluster size and number of clusters are used interchangeably in this paper.

The major contributions of the paper are summarized as follow:

- Developing an analytical model for cluster size optimization in Gaussian distributed sensor networks.
- Comparing of Gaussian and uniformly distributed sensor network based on cluster size optimization.
- Providing the modeling and analysis by MATLAB programming and demonstrating the effectiveness of Gaussian distribution for different type of radio model.
- Providing expression for optimal number of clusters for square sensing field with consideration of boundary effect.

To the best of our knowledge, there is no published work providing the cluster size optimization for Gaussian distributed sensor networks. The subsequent sections of this paper are organised as follows: In Section II, we describe the network model that is used in this work and some background information. This includes radio energy dissipation model and data aggregation. Section III, discusses a review of related works with cluster size optimization. Section IV presents proposed cluster size optimization model for Gaussian distributed WSNs and describing procedure for finding analytically optimal number of clusters. Section V presents results and analysis. Section VI concludes the paper along with direction of future work.

II. PRELIMINARIES AND NETWORK MODEL

In our model, we consider homogeneous sensor nodes that follow Gaussian distribution over the sensing field. The BS is placed at centre of sensing field. Fig. 1 shows the sensor nodes which follow uniform random and Gaussian random distributions over the sensing field [7].

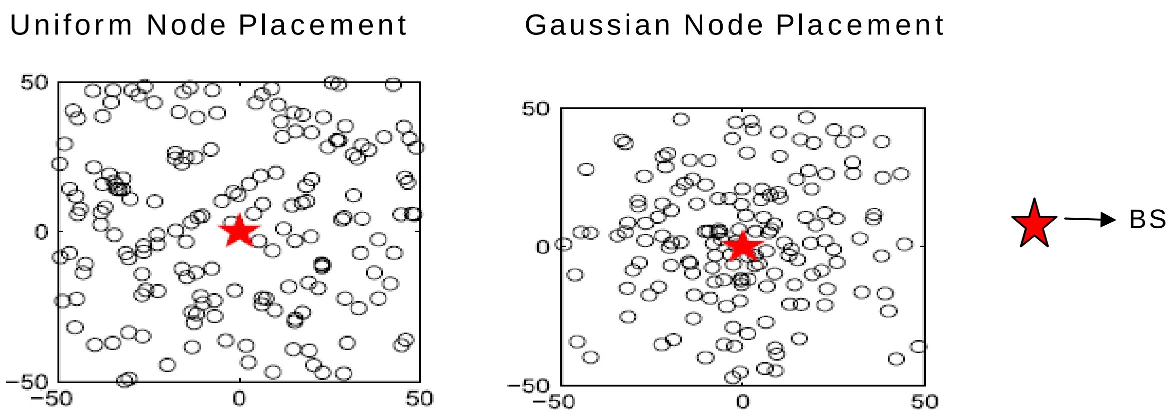

Fig. 1. Uniform vs Gaussian distributed sensor networks [7]

A. Gaussian distribution

In Gaussian distribution, the probability density function that a sensor node resides at point (x, y) with respect to deployment point (x_0, y_0) [14].

$$f_i(x, y) = \frac{1}{2\pi\sigma_x\sigma_y} e^{-\left(\frac{(x-x_0)^2}{2\sigma_x^2}+\frac{(y-y_0)^2}{2\sigma_y^2}\right)} \tag{1}$$

Where σ_x *and* σ_y are standard deviation to x and y co-ordinate

For one dimensional Gaussian distribution with mean distance y_0, the PDF can be defined as follow:

$$f\ (y) = \frac{1}{\sqrt{2\pi}\sigma_y} e^{-\left(\frac{(y-y_o)^2}{2\sigma_y^2}\right)} \tag{2}$$

Where y_0 is the mean distance and σ_y represent standard deviation.

B. Network Assumptions

In this work, the following assumptions have been made [15, 16]:

1. BS that ultimately processes the collected data is located at the centre of the sensing field
2. All the sensor nodes are identical and are stationary after deployment.
3. Sensing field is square
4. Nodes follow Gaussian distribution
5. The data aggregation efficiency of Cluster Heads (CHs) is 100%.
6. The propagation channels are symmetric.
7. Nodes are not equipped with GPS unit and therefore, they are not location-aware.

C. Radio Energy Dissipation Model

Throughout this paper we are using the simple energy consumption model as introduced in [17, 18, 19] and shown in fig. 2. In this energy model, the total energy consumption of the which is given by: $E_c = E_{Tx}(p,d) + E_{Rx}(p)$. The energy consumption at the transmitter ($E_{Tx}(p,d)$) is divided into the energy consumption in transmit electronics and transmitter amplifier while the receiver energy consumption ($E_{Rx}(p)$) depends on only the receiver electronics. Then, the transmitter and receiver energy consumptions are $E_{Tx}(p,d) = pE_{elec} + p\varepsilon_{amp}d^n$ and $E_{Rx}(p) = pE_{elec}$ respectively. We are assuming that: $E_{elec} = E_{tx\_elec}(p) = E_{rx\_elec}(p)$ which is the energy being dissipated to run the transmitter or receiver circuitry to transmit or receive one bit of the data packet and $\varepsilon_{amp} = E_{tx\_amp}(p,d)$ is energy dissipation of the transmission amplifier to convey one bit of data packet to the receiver node with a distance of *d* as 1m away. *p* is length of transmitted/ received message in bits and *d* is distance between transmitter and receiver node, *n* is path loss exponent. *n* equal to 2 is used for free space model ($\varepsilon_{amp} = \varepsilon_{fs}$, when $d < d_0$) and *n* equal to 4 for multipath model ($\varepsilon_{amp} = \varepsilon_{mp}$ when $d > d_0$) $p\varepsilon_{fs} \times d_0^2 = p\varepsilon_{mp} \times d_0^4$ i.e. $d_0 = \sqrt{\frac{\varepsilon_{fs}}{\varepsilon_{mp}}}$

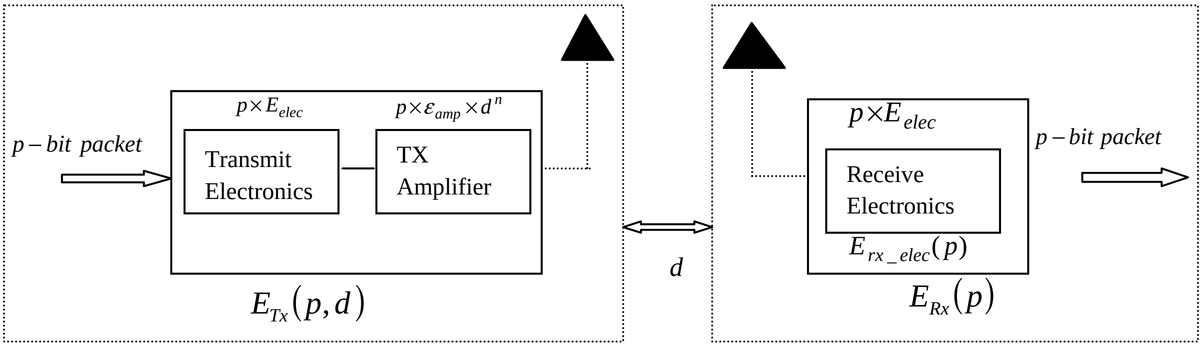

Fig. 2. Radio Energy Dissipation Model [21]

D. Data Aggregation

The CHs are responsible for aggregating their cluster members data signals to produce a single representative signal. Let $E_{agg}(s,p)$ be energy spent in aggregating "s" streams of "p" bits information into a single stream of p bits of aggregated information. Then $E_{agg}(s,p) = \gamma \times p$ Where γ is the energy required to aggregate one bit of data [20].

III. RELATED WORK & CLUSTER SIZE OPTIMIZATION

Mohammed et al. proposed collaborative beam forming technique to increase the transmission range of individual sensor nodes. The author found that the main challenge in using collaborative beam forming in WSNs is the uncertainty regarding the sensor node locations. However, the sensor node distribution can be modeled by a properly selected Probability Density Function (PDF). Authors stated that Gaussian PDF, is more suitable in many WSN applications than, for example, uniform PDF which is commonly used for ad-hoc networks [5]. Gaussian deployment gives wider main lobe and has lower chance of large side lobes.

Wang et al. found that a uniform random WSN is not able to detect moving intruder if it starts inside the network area and is close to the target. Gaussian-distributed WSNs can provide differentiated detection capabilities at different locations. Authors compared performance of Gaussian-distributed WSNs with uniformly distributed WSNs. This work analytically formulates detection probability in a random WSN and provides guidelines in selecting an appropriate deployment strategy and determining critical network parameters [6].

Hasbullah et al. considered the problem of energy holes and unbalanced energy consumption in many-to-one sensor network due to uniform deployment the nodes closer to a sink carrying heavy traffic loads via multi-hop transmission. This results into energy depletion within the area at an increased rate and ultimately leads to energy holes around the sink. Such hot-spots are more likely to occur closer to the sink instead of any other geographical area spanned by the network [4].

Heinzelman et al. proposed [21], Low Energy Adaptive Clustering Hierarchy protocol(LEACH) for finding optimal number of clusters in WSNs, where sensor nodes are deployed in random and uniform manner. LEACH uses a TDMA/CDMA MAC to reduce inter-cluster and intra-cluster collisions. Cluster formation is based on many properties such as the number and type of sensors, communication range and geographical location.

Kim et al. [22] has estimated the optimal number of clusters among random and uniform distributed sensors in a bounded sensing field. In this algorithm, optimal number of clusters-heads depends on the distance between the base station and sensor nodes.

Chan et al. [23] proposed a Fixed Optimal Cluster (FOC) numbers to examine the entire network. They have adapted two different ways for calculating optimal cluster numbers depending on the position of the base station and whether the BS is outside the sensing field or not.

Yang et al. [24] have proposed a more reasonable energy consumption model called Optimal Energy Consumption Model (OECM) in a homogeneous network with random and uniform node distribution. It shows that the optimal number of cluster heads not only depends on node density, but also depends on size of sensing field, circuit energy dissipation and packet length.

Navid et al. [25], analytically provides the optimal number of clusters that minimizes the total energy expenses in the uniform and random node distributed networks, where all sensor nodes communicate data through their elected CHs to the BS in a distributed fashion. The results show that the energy consumption of the transmitter circuitry has no impact on the optimal number of clusters and the energy consumption of the receiver electronics can substantially change the optimal number of clusters and more importantly it can decide whether or not it is worth performing clustering.

Chen et al. [26] proposed CH optimization based on energy. In this, the authors considered a threshold value and the residual energy of node, to optimize the selection of a cluster head. Results show that this algorithm can prolong the network lifetime efficiently compared with LEACH [21] protocol.

The algorithms based on random and uniform node distribution have following problems:

- Cluster head selection is uncontrolled.
- Cannot be applied to all the practical cases

Tripathi et al. [27] introduced clustering of non-uniform random distributed nodes in a circular ring, and calculated the optimal number of clusters in WSN. Results show that there is balanced energy expenditure in this non-uniform clustering.

So we can say that Gaussian distribution is better than uniform distribution for realistic applications and optimal number of clusters depends on different power of d_{toBS} , d^2_{toBS} , d^4_{toBS} . Where d_{toBS} denotes distance between CH and BS

A. Cluster Size Optimization

In sensing field, choosing more clusters while maintaining the same load per CH, the communication distance from a sensor node to its own CH is reduced. Therefore, the overall energy consumption is also reduced. On the other hand, increasing the number of clusters means that the communication path between a sensor and the BS will include more CH to CH hops, which mean higher overall energy consumption. Therefore, finding the optimal number of clusters is a crucial point for the WSNs [28].

Let us assume square shape sensing field of area A, (side M in square sensing field) with N nodes which follow Gaussian placement with optimal number of clusters K_{opt}. It means that each cluster contains

$$\frac{N}{K_{opt}} \text{nodes} \left\{ \left(\frac{N}{K_{opt}} - 1 \right) \text{non - cluster head nodes and one CH node} \right\}.$$

The energy consumption for non-CH nodes follows the free-space model and can be represented as

$E_{non-CH} = pE_{tx\_elec} + p\varepsilon_{fs} d^2_{toCH}$, where d_{toCH} is the distance between the non-CH node and its CH.

Expected value of square of d_{toCH} is given by:

$$E[d_{toCH}^2] = \frac{A}{2\pi K_{opt}} \tag{3}$$

Finally we can say the energy consumed in each non-CH node per round is given by:

$$E_{non-CH} = pE_{tx\_elec} + \frac{p\varepsilon_{fs}A}{2\pi K_{opt}} \tag{4}$$

So, the energy dissipated in an entire cluster during a single round:

$$E_{cluster} = E_{CH} + (\frac{N}{K_{opt}} - 1)E_{non-CH} \tag{5}$$

Total energy consumption of the system can be represented as

$$E_T = K_{opt}\ E_{cluster} \tag{6}$$

IV.

$$E_T = NpE_{rx\_elec} - K_{opt}\,pE_{rx\_elec} + NpE_{agg}(s,p) + K_{opt}\,p\varepsilon_{amp}d_{toBS}^n + NpE_{elec}^{Tx} + \frac{Np\varepsilon_{fs}A}{2\pi K_{opt}} - \frac{p\varepsilon_{fs}A}{2\pi} \tag{7}$$

Taking derivative of equation (7) with respect to K_{opt} and the optimal number of clusters can be calculated [20]:

$$K_{opt} = \sqrt{\frac{N\varepsilon_{fs}A}{2\pi(\varepsilon_{amp}d_{toBS}^n - E_{rx\_elec})}} \tag{8}$$

The equation 8 represent optimal number of clusters which depends on node density, area of sensing field, free space energy, amplifier energy, power of distance between BS and CHs and energy consumed by receiver circuitry. The expected values of different power of distance between CHs and BS (d_{toBS} , d_{toBS}^2 , d_{toBS}^4) has been derived in section IV.

V. PROPOSED ANALYTICAL MODEL FOR CLUSTER SIZE OPTIMIZATION IN GAUSSIAN DISTRIBUTED WSN

Let the shape of a sensing field is the square of side M and assume that the BS is located at the centre of the sensing field as shown in fig.3. The probability P that the distance between a randomly chosen point and the BS located at the center of the square is less than y should be obtained. The minimum and maximum value of angle δ are 0 radian and $\frac{\pi}{4}$ radian respectively. where $\delta = \tan^{-1}\left(\frac{\sqrt{y^2 - \frac{M^2}{4}}}{\frac{M}{2}}\right)$

The probability density function is given by:

$$f(y) = \begin{cases} \frac{\sqrt{2\pi}\, y}{M^2\sigma_y}\, e^{-\left(\frac{(y-y_0)^2}{2\sigma_y^2}\right)} & if \quad 0 \le y \le \frac{M}{2} \\ \frac{1}{\sqrt{2\pi}\,\sigma_y}\left\{\frac{2\pi y}{M^2} - \frac{8y}{M^2}\delta\right\} e^{-\left(\frac{(y-y_0)^2}{2\sigma_y^2}\right)} & if \quad \frac{M}{2} \le y \le \frac{M}{\sqrt{2}} \end{cases} \tag{9}$$

Assume that $\sigma_y = \sigma$ for simplicity.

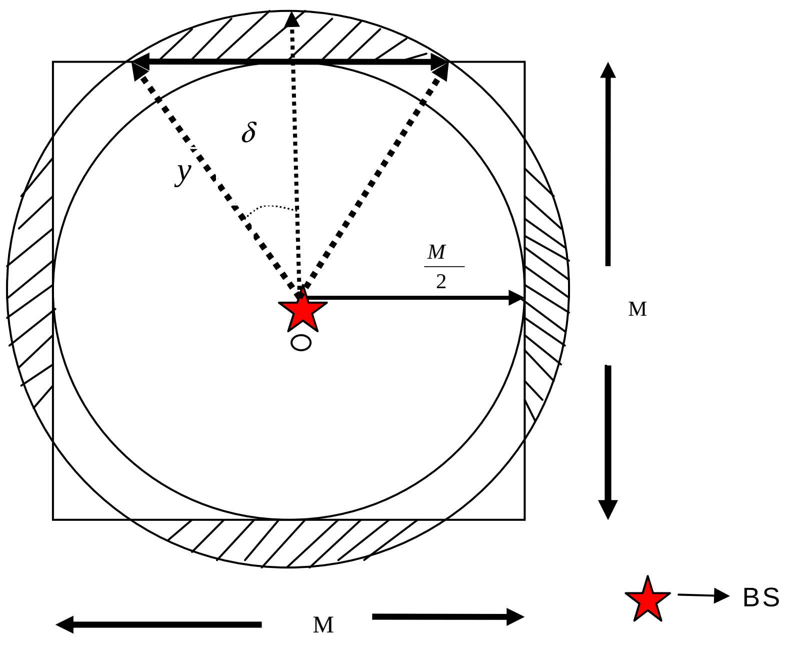

Fig. 3. Squared shaped Sensing field with BS at the centre

A. Calculation of Expected value of d_{toBS}

To find the expected value of the distance from the centre to the entire area of the square, one must integrate $y\,f(y)$ in the interval of $[0, \frac{M}{\sqrt{2}}]$. The values for all combination of standard deviation of Gaussian distribution (σ) and mean distance (y_0) are shown in TABLE I.

TABLE I: Values of d_{toBS} for Gaussian Node distribution WSN for all combination of y_0 and σ

σ \ y_0	0.10M	0.20M	0.30M	0.40M	0.50M	0.60M	0.70M	0.80M	0.90M
0.10M	0.1205	0.3120	0.6080	0.9017	0.9284	0.6054	0.6054	0.0509	0.0054
0.20M	0.2311	0.3835	0.5386	0.6369	0.6299	0.5177	0.3514	0.1958	0.0892
0.30M	0.2731	0.3583	0.4284	0.4663	0.4617	0.4155	0.3396	0.2519	0.1694
0.40M	0.2629	0.3096	0.3446	0.3622	0.3597	0.3373	0.2987	0.2497	0.1971
0.50M	0.2389	0.2661	0.2856	0.2951	0.2936	0.2814	0.2596	0.2306	0.1973
0.60M	0.2139	0.2309	0.2426	0.2483	0.2474	0.2400	0.2267	0.2086	0.1868
0.70M	0.1916	0.2028	0.2104	0.2140	0.2134	0.2086	0.2000	0.1879	0.1732
0.80M	0.1725	0.1802	0.1854	0.1878	0.1874	0.1842	0.1783	0.1699	0.1595
0.90M	0.1566	0.1621	0.1658	0.1675	0.1672	0.1649	0.1607	0.1547	0.1471

B. Calculation of Expected Value of d_{toBS}^2

To find the expected value of the distance from the centre to the entire area of the square, one must integrate $y^2\,f(y)$ in the interval of $[0, \frac{M}{\sqrt{2}}]$. The values for all combination of standard deviation of Gaussian distribution (σ) and mean distance (y_0) are shown in TABLE II.

TABLE II: Values of d_{toBS}^{2} for Gaussian Node distribution WSN for all combination of y_0 and σ

d_{toBS}^{2}									
σ \ y_0	**0.10M**	**0.20M**	**0.30M**	**0.40M**	**0.50M**	**0.60M**	**0.70M**	**0.80M**	**0.90M**
0.10M	$0.0256M$	$0.0872M$	$0.2152M$	$0.3814M$	$0.4467M$	$0.3224M$	$0.1366M$	$0.0312M$	$0.0034M$
0.20M	$0.0775M$	$0.1417M$	$0.2164M$	$0.2750M$	$0.2891M$	$0.2504M$	$0.1778M$	$0.1030M$	$0.0485M$
0.30M	$0.1055M$	$0.1443M$	$0.1794M$	$0.2023M$	$0.2069M$	$0.1919M$	$0.1612M$	$0.1226M$	$0.0844M$
0.40M	$0.1070M$	$0.1289M$	$0.1466M$	$0.1574M$	$0.1594M$	$0.1523M$	$0.1373M$	$0.1167M$	$0.0936M$
0.50M	$0.09964M$	$0.1126M$	$0.1225M$	$0.1283M$	$0.1293M$	$0.1255M$	$0.1172M$	$0.1054M$	$0.0912M$
0.60M	$0.09042M$	$0.0985M$	$0.1045M$	$0.1080M$	$0.1086M$	$0.1063M$	$0.1013M$	$0.0940M$	$0.0849M$
0.70M	$0.0816M$	$0.0870M$	$0.0909M$	$0.0931M$	$0.0935M$	$0.0920M$	$0.0888M$	$0.0840M$	$0.0778M$
0.80M	$0.0738M$	$0.0775M$	$0.0802M$	$0.0817M$	$0.0820M$	$0.0810M$	$0.0788M$	$0.0755M$	$0.0712M$
0.90M	$0.0672M$	$0.0699M$	$0.0718M$	$0.0729M$	$0.0731M$	$0.0724M$	$0.0708M$	$0.0684M$	$0.0653M$

C. Calculation of expected value of d_{toBS}^{4}

To find the expected value of the distance from the centre to the entire area of the square, one must integrate $y^4 f(y)$ in the interval of $[0, \frac{M}{\sqrt{2}}]$. The values for all combination of standard deviation of Gaussian distribution (σ) and mean distance (y_0) are shown in TABLE III

TABLE III: Values of d_{toBS}^{4} for Gaussian Node distribution WSN for all combination of y_0 and σ

d_{toBS}^{4}									
σ \ y_0	**0.10M**	**0.20M**	**0.30M**	**0.40M**	**0.50M**	**0.60M**	**0.70M**	**0.80M**	**0.90M**
0.10M	$.0016M^3$	$.0087M^3$	$.0316M^3$	$.0754M^3$	$.1106M^3$	$.0961M^3$	$.0473M^3$	$.0120M^3$	$.0014M^3$
0.20M	$.0119M^3$	$.0249M^3$	$.0428M^3$	$.0604M^3$	$.0696M^3$	$.0654M^3$	$.0499M^3$	$.0308M^3$	$.0153M^3$
0.30M	$.0204M^3$	$.0295M^3$	$.0387M^3$	$.0458M^3$	$.0491M^3$	$.0475M^3$	$.0415M^3$	$.0328M^3$	$.0234M^3$
0.40M	$.0224M^3$	$.0279M^3$	$.0326M^3$	$.0360M^3$	$.0375M^3$	$.0368M^3$	$.0340M^3$	$.0296M^3$	$.0243M^3$
0.50M	$.0217M^3$	$.0250M^3$	$.0277M^3$	$.0295M^3$	$.0303M^3$	$.0299M^3$	$.0284M^3$	$.0260M^3$	$.0228M^3$
0.60M	$.0201M^3$	$.0222M^3$	$.0238M^3$	$.0249M^3$	$.0254M^3$	$.0252M^3$	$.0243M^3$	$.0228M^3$	$.0208M^3$
0.70M	$.0183M^3$	$.0197M^3$	$.0208M^3$	$.0215M^3$	$.0218M^3$	$.0217M^3$	$.0211M^3$	$.0201M^3$	$.0188M^3$
0.80M	$.0167M^3$	$.0177M^3$	$.0184M^3$	$.0189M^3$	$.0191M^3$	$.0190M^3$	$.0186M^3$	$.0180M^3$	$.0171M^3$
0.90M	$.0153M^3$	$.0160M^3$	$.0165M^3$	$.0169M^3$	$.0170M^3$	$.0170M^3$	$.0167M^3$	$.0162M^3$	$.0156M^3$

VI. RESULTS AND ANALYSIS

In this section, analytical results of optimal number of clusters for square sensing fields with Gaussian distributed sensor nodes are discussed as follows: The value of d_{toBS}, d_{toBS}^{2}, d_{toBS}^{4} with different values of Gaussian standard deviation and mean distance are shown in TABLE I, II and III respectively. The analytical results presented in the section IV is further validated through MATLAB programming on Gaussian distributed sensor networks. TABLE VII shows the simulation parameters. It can be inferred from TABLE IV that the optimal number of clusters (Optimal cluster size) depends on the node density (N) as well as dimension of the sensing field (M) under some constraint. Such constraint implies that electronics energy of the receiver should be less and BS should be located on centre of the sensing field for both free space and two ray radio model. The value of d_{toBS}^{2}, d_{toBS}^{4} from TABLE II & III shows that the value of optimal number of clusters will be least, when mean distance is half of the sensing field and Gaussian standard deviation is 10% of the size of the sensing field for all combination of Gaussian standard deviation and mean distance.

TABLE IV: Optimal Number of clusters for free space and two ray radio model

	K_{opt} *for* $n=2$		K_{opt} *for* $n=4$	
Standard Deviation σ with y_0 =0.50M	Optimal number of Clusters	Optimal number of cluster for small E_{rx_elec}	Optimal number of Clusters	Optimal number of cluster for small E_{rx_elec}
0.10M	$\sqrt{\dfrac{N\varepsilon_{fs}M^2}{2\pi(0.4467M\ \varepsilon_{fs}-E_{rx\_elec})}}$	$0.5970\sqrt{NM}$	$\sqrt{\dfrac{N\varepsilon_{fs}M^2}{2\pi(0.1106M^3\varepsilon_{fs}-E_{rx\_elec})}}$	$1.199\sqrt{\dfrac{N}{M}}\sqrt{\dfrac{\varepsilon_{fs}}{\varepsilon_{mp}}}$
0.20M	$\sqrt{\dfrac{N\varepsilon_{fs}M^2}{2\pi(0.2891M\varepsilon_{fs}-E_{rx\_elec})}}$	$0.7421\sqrt{NM}$	$\sqrt{\dfrac{N\varepsilon_{fs}M^2}{2\pi(0.0696M^3\varepsilon_{fs}-E_{rx\_elec})}}$	$1.512\sqrt{\dfrac{N}{M}}\sqrt{\dfrac{\varepsilon_{fs}}{\varepsilon_{mp}}}$
0.30M	$\sqrt{\dfrac{N\varepsilon_{fs}M^2}{2\pi(0.2069M\ \varepsilon_{fs}-E_{rx\_elec})}}$	$0.8772\sqrt{NM}$	$\sqrt{\dfrac{N\varepsilon_{fs}M^2}{2\pi(0.0491M^3\varepsilon_{fs}-E_{rx\_elec})}}$	$1.802\sqrt{\dfrac{N}{M}}\sqrt{\dfrac{\varepsilon_{fs}}{\varepsilon_{mp}}}$
0.40M	$\sqrt{\dfrac{N\varepsilon_{fs}M^2}{2\pi(0.1594M\ \varepsilon_{fs}-E_{rx\_elec})}}$	$0.999\sqrt{NM}$	$\sqrt{\dfrac{N\varepsilon_{fs}M^2}{2\pi(0.0375M^3\varepsilon_{fs}-E_{rx\_elec})}}$	$2.060\sqrt{\dfrac{N}{M}}\sqrt{\dfrac{\varepsilon_{fs}}{\varepsilon_{mp}}}$
0.50M	$\sqrt{\dfrac{N\varepsilon_{fs}M^2}{2\pi(0.1293M\varepsilon_{fs}-E_{rx\_elec})}}$	$1.109\sqrt{NM}$	$\sqrt{\dfrac{N\varepsilon_{fs}M^2}{2\pi(0.0303M^3\varepsilon_{fs}-E_{rx\_elec})}}$	$2.292\sqrt{\dfrac{N}{M}}\sqrt{\dfrac{\varepsilon_{fs}}{\varepsilon_{mp}}}$
0.60M	$\sqrt{\dfrac{N\varepsilon_{fs}M^2}{2\pi(0.1086M\varepsilon_{fs}-E_{rx\_elec})}}$	$1.210\sqrt{NM}$	$\sqrt{\dfrac{N\varepsilon_{fs}M^2}{2\pi(0.0254M^3\varepsilon_{fs}-E_{rx\_elec})}}$	$2.503\sqrt{\dfrac{N}{M}}\sqrt{\dfrac{\varepsilon_{fs}}{\varepsilon_{mp}}}$
0.70M	$\sqrt{\dfrac{N\varepsilon_{fs}M^2}{2\pi(0.0935M\varepsilon_{fs}-E_{rx\_elec})}}$	$1.305\sqrt{NM}$	$\sqrt{\dfrac{N\varepsilon_{fs}M^2}{2\pi(0.0218M^3\varepsilon_{fs}-E_{rx\_elec})}}$	$2.7026\sqrt{\dfrac{N}{M}}\sqrt{\dfrac{\varepsilon_{fs}}{\varepsilon_{mp}}}$
0.80M	$\sqrt{\dfrac{N\varepsilon_{fs}M^2}{2\pi(0.0820M\ \varepsilon_{fs}-E_{rx\_elec})}}$	$1.393\sqrt{NM}$	$\sqrt{\dfrac{N\varepsilon_{fs}M^2}{2\pi(0.0191M^3\varepsilon_{fs}-E_{rx\_elec})}}$	$2.880\sqrt{\dfrac{N}{M}}\sqrt{\dfrac{\varepsilon_{fs}}{\varepsilon_{mp}}}$
0.90M	$\sqrt{\dfrac{N\varepsilon_{fs}M^2}{2\pi(0.0731M\varepsilon_{fs}-E_{rx\_elec})}}$	$1.475\sqrt{NM}$	$\sqrt{\dfrac{N\varepsilon_{fs}M^2}{2\pi(0.0170M^3\varepsilon_{fs}-E_{rx\_elec})}}$	$3.060\sqrt{\dfrac{N}{M}}\sqrt{\dfrac{\varepsilon_{fs}}{\varepsilon_{mp}}}$

TABLE V: Comparison of Optimal Number of clusters for uniform and Gaussian node distributed WSNs

Sensing field	Radio Model	Location of BS	Uniform Node Distribution	Gaussian Node Distribution
			Optimal number of cluster K_{opt} **[20]**	Optimal number of cluster K_{opt} $(\sigma = 0.10M, y_0 = 0.50M)$ **[Proposed]**
Square	Free space Model	Centre of sensing field	$0.977\sqrt{N}$	$0.60\sqrt{NM}$
	Two ray Model		$2.023\sqrt{\dfrac{N}{M^2}}\sqrt{\dfrac{\varepsilon_{fs}}{\varepsilon_{mp}}}$	$1.19\sqrt{\dfrac{N}{M}}\sqrt{\dfrac{\varepsilon_{fs}}{\varepsilon_{mp}}}$

From TABLE IV, optimal number of clusters for free space radio model with Gaussian distributed sensor network under given all assumptions can be represented by $K_{opt(n=2)} = \alpha\sqrt{N\,M}$ where $0.59 \le \alpha \le 6.84$. Optimal number of clusters for two ray radio model with Gaussian distributed sensor network under given assumptions can be represented by $K_{opt(n=4)} = \beta\sqrt{\dfrac{N}{M}}$ where $104 \le \beta \le 927$. Fig. 4 &5 represents comparative analysis between the uniform and Gaussian distributed sensor network for free space and two ray path model respectively. Fig. 4 shows that the optimal number of clusters increases in Gaussian distributed WSNs as dimension of sensing field increases but for uniform WSNs it is independent of dimension of the sensing field. For two ray path model as shown in fig. 5, the optimal number of clusters decreases as dimension of the sensing field increases. Fig. 6 shows Optimal number of clusters vs standard deviation of gaussian distributed sensor networks. From this figure we can say that for fixed dimension, as number of nodes increases the value of optimal number of clusters also increases. Optimal number of clusters also increases as the standard deviation of gaussian distribution increases. In this paper we have proposed the gaussian distribution with specific mean distance that will provide feasable network for realistic applications. Fig. 7 shows the optimal number of clusters vs mean distance in gaussian distributed sensor networks(n=2). From this figure we can say that for 50% of the mean distance, the optimal number of clusters will be minimum. Because at this value the power of distance between CH and BS will be maximum. For any application if we need minimum value of optimal number of clusters we have to keep value of standard devaition as 0.10M and mean distance should be 0.50M. TABLE V shows the comparative analysis between uniform and gaussian distributed sensor networks.

TABLE VI: Simulation parameter

Parameters	**Values**
Sensing field type	Square ($M \times M$)
Location of BS	Centre of sensing
Sensing Model	Circular Sensing
Energy model initial energy of each	2J
E_{tx_elec}	50nJ/bit
E_{rx_elec}	50nJ/bit
ε_{fs}	10pJ/bit/m^2
ε_{mp}	0.0013pJ/bit/m^4
Path loss exponent (n)	2, 4
$E_{agg}(p)$	5nJ/bit/signal

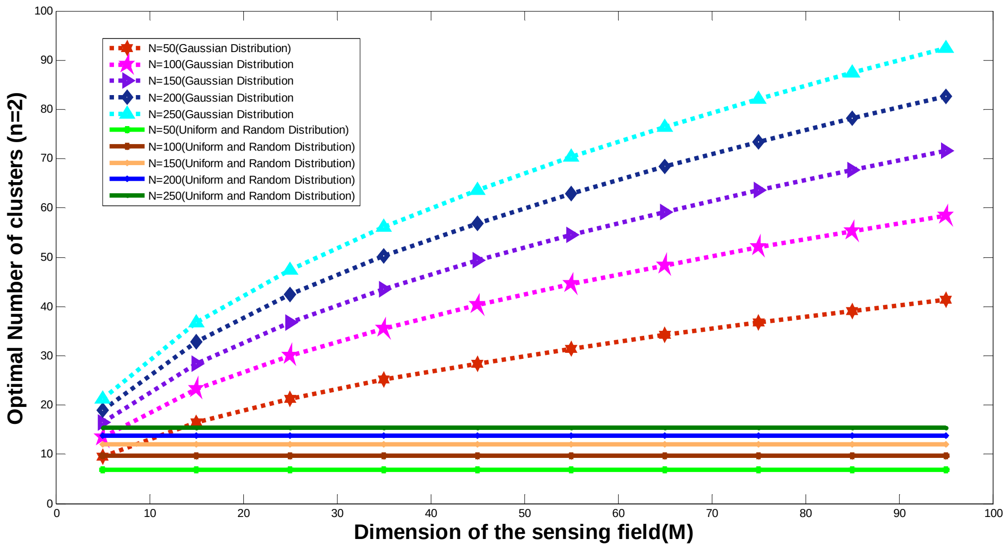

Fig. 4. Optimal number of clusters vs dimension of sensing field for uniform and gaussian distributed sensor networks for(n=2)

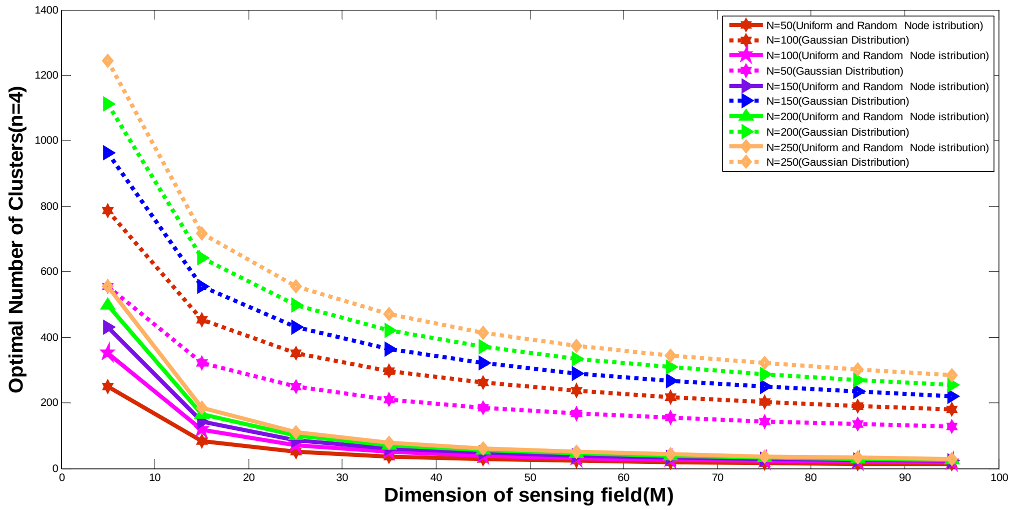

Fig.5. Optimal number of clusters vs dimension of sensing field for uniform and gaussian distributed sensor networks(n=4)

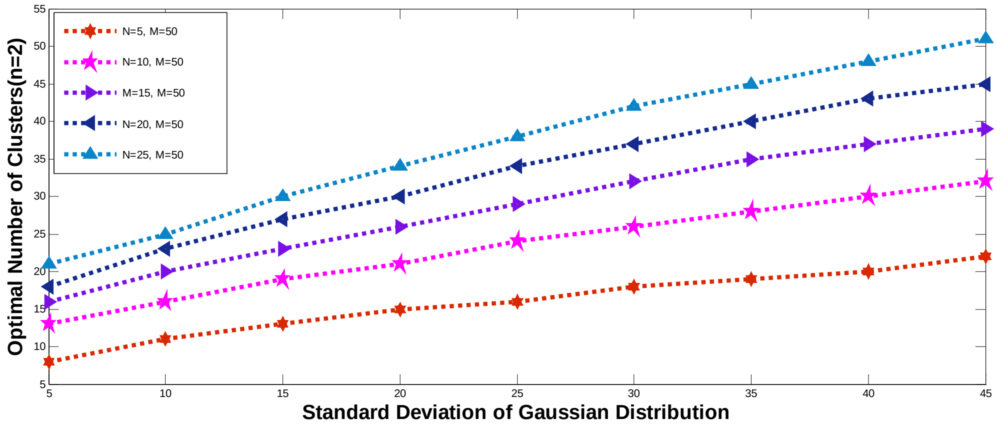

Fig. 6. Optimal number of clusters vs standard deviation of gaussian distributed sensor networks(n=2)

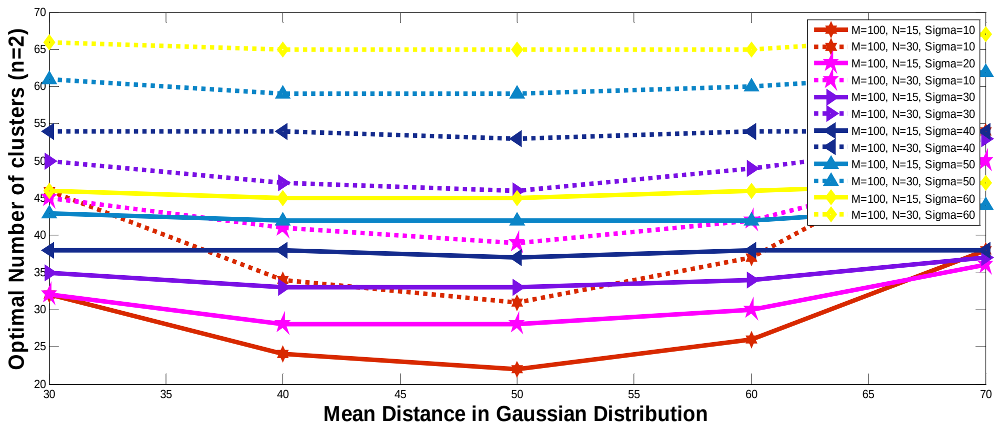

Fig. 7. Optimal number of clusters vs mean distance in gaussian distributed sensor networks(n=2)

VI. CONCLUSIONS

In this paper, we have provided analytical framework to determine the optimal number of clusters for Gaussian distributed sensor networks. The analytical results are summarized below:

1: The optimal number of clusters in Gaussian distributed sensor networks depends on node density(N) as well as dimension of the sensing field(M) i.e. $K_{opt} = f(N, M)$ under certain specific condition for both free space and two ray radio model. For uniform distributed sensor networks the optimal number of clusters only depends on node density (N) for free space radio model (n=2) i.e. $K_{opt} = f(N)$. Thus proposed Gaussian distributed based cluster optimization appears realistic in practice.

2: The value of optimal number of clusters will be least when mean distance is half of the sensing field (0.50M) and Gaussian standard deviation is 10% of the size of the sensing field for all possible combinations of Gaussian standard deviation (σ) and mean distance(y_0).

3: Optimal number of clusters for free space radio model with Gaussian distributed sensor network under given assumptions can be represented by $K_{opt(n=2)} = \alpha\sqrt{N\,M}$ where $0.59 \le \alpha \le 6.84$. Assuming BS at centre of the square sensing field

4: Optimal number of clusters for two ray radio model with Gaussian distributed sensor network can be represented by $K_{opt(n=4)} = \beta\sqrt{\frac{N}{M}}$ where $104 \le \beta \le 927$. Assuming BS at centre of the square sensing field.

A. Open Issues and Challenges

- With Gaussian node placement, utilizing a more practical radio energy model other than [22], by including energy required for modulation process in long range communication.
- Investigating optimal number of cluster for different shape of sensing field (Circular, hexagonal, equilateral triangle for maximum coverage etc) for dynamic position of BS with Gaussian distributed sensor network.
- Investigating optimal number of cluster for different shapes of sensing field (Square Circular, hexagonal, equilateral triangle for maximum coverage etc) considering different sensing model of node (Multilevel sensing model and Elfes sensing model) with Gaussian distributed sensor network.

REFERENCES

[1] G. Pottie and W. Kaiser “Wireless Integrated Network Sensors,” ACM Communications, Vol.43, Issue 5, pp. 51–58, 2000.
[2] M. Ilyas and I. Mahgoub, Handbook of Sensor Networks: Compact Wireless and Wired Sensing Systems, CRC PRES Boca Raton London New York Washington, D.C 2005.
[3] I. F. Akyildiz and M. C.Vuran, Wireless Sensor Networks, John Wiley and Sons, Ltd, Publication 2010.
[4] A. Ur-Rahman, H. Hasbullah and N. Sam “Impact of Gaussian Deployment Strategies on the Performance of Wireless Sensor Network,” International Conference on Computer & Information Science (ICCIS), 2012, pp.771-776.
[5] M. F. A. Ahmed, S. A. Vorobyov “Collaborative Beam forming for Wireless Sensor Networks with Gaussian Distributed Sensor Nodes,” IEEE transactions on Wireless Communications, Vol. 8, issue 2, February 2009.
[6] Y. Wang, W. Fu, and D.P. Agrawal “Gaussian versus Uniform Distribution for Intrusion Detection in Wireless Sensor Networks” IEEE Transactions on Parallel and Distributed Systems, Vol. 24, Issue No. 2, February 2013.
[7] Y. Wang, Z. Lun “Intrusion detection in a -Gaussian distributed wireless sensor network,” Journal of Parallel and Distributed Computing, Vol. 71, Issue 12, pp. 1598-1607, December 2011.
[8] A. A. Abbasi and M. Younis “A Survey on Clustering Algorithms for Wireless Sensor Networks,” International Journal on Computer Communications, Vol. 30, pp. 2826–2841, 2007.
[9] S. Kushwaha, V. Kumar, S. Jain “Node architectures and its deployment in wireless sensor networks: A survey” International Conference on High Performance Architecture and Grid Computing, published Springer Berlin Heidelberg, 2010, pp. 515-526.
[10] V. Kumar, S. Jain and S.Tiwari “Energy efficient clustering algorithms in wireless sensor networks: A survey” International Journal of Computer Science Issues pp. 1694-0814, vol.8, issue5, 2011.
[11] N. Tuah, M. Ismail, and K. Jumari, "Energy efficient algorithm for heterogeneous wireless sensor network," IEEE International Conference on Control System, Computing and Engineering, pp. 92-96, 2011.
[12] A. Dabirmoghaddam, M. Ghaderi and C. Williamson “On the optimal randomized clustering in distributed sensor networks” Journal of computer network, Vol. 59, Issue 11, pp. 17–32, February 2014.
[13] S. Naeimi, H. Ghafghazi, C. Chow, H. Ishii “Survey on the Taxonomy for Cluster-based Routing Protocols for Homogeneous Wireless Sensor Networks” Journal of Sensor, MDPI, Vol. 12, Issue 6, pp. 7350–7409, 2012.
[14] A. Leon-Garcia, Probability and Random Processes for Electrical Engineering, second ed., Addison-Wesley, 1993.
[15] L. Zhao, Q. Liang “An access-based low-energy hierarchy for sensor networks” Proceedings of the IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC), Barcelona, Spain, 2004.
[16] K. Yang, Y. Wu, and H. Zhou, “Research of optimal energy consumption model in wireless sensor network” 2nd International Conference on Computer Engineering and Technology Chengdu, China, April 2010, pp. V7-421 - V7-424.
[17] C. Alippi, R. Camplani, M. Roveri “An adaptive LLC-based and hierarchical power-aware routing algorithm”, IEEE Transactions on Instrumentation and Measurement Vol. 58, pp. 3347–3357, issue 9, 2009.
[18] B. Braem, B. Latre, I. Moerman, C. Blondia, E. Reusens, W. Joseph, L. Martens, P. Demeester “The need for cooperation and relaying in short-range high path loss sensor networks” Proceedings of International Conference on Sensor Technol. Appl. (SENSORCOMM 2007), 2007, pp. 566–571.

[19] Z. Cheng, M. Perillo, W.B. Heinzelman, "General network lifetime and cost models for evaluating sensor network deployment strategies", IEEE Transactions on Mobile Computing Vol.7, Issue 4, pp. 484–497, 2008.

[20] A. Navid, V. Alireza, X. Wenyao, G. Mario and S. Majid. "Cluster size optimization in sensor networks with decentralized cluster-based protocols" Journal of Computer Communications, Vol.35, Issue 2, pp. 207–220, January 2012.

[21] W. R. Heinzelman, A. Chandrakasan, and H. Balakrishnan, "An application-specific protocol architecture for wireless microsensor networks," IEEE Transaction on Wireless Communication, Vol. 1, No. 4, pp. 660– 670, October 2002.

[22] H. Kim, S.W. Kim, S.B. Lee, and B. Son "Estimation of the optimal number of cluster-heads in sensor network" Proceeding of KES, Melbourne, Australia, 2005, Vol. 3, pp. 87–94.

[23] T.J. Chan, M.C. Chen, Y.F. Huang, J.Y. Lin, and T.R. Chen "Optimal Cluster Number Selection in Ad-hoc Wireless Sensor Networks" WSEAS Transaction on Communication, Vol. 7, Issue 8 pp. 837-846, August 2008.

[24] H. Karl and A. Willig, Protocols and architectures for wireless sensor networks, A John Wiley and Sons, Ltd, Publication, 2005

[25] T. Banerjee, B. Xie, J.H. Jun, D.P. Agrawal "Increasing lifetime of wireless sensor networks using controllable mobile cluster heads" journal of Wireless Communications and Mobile Computing, 2009

[26] B. Chen, Y. Zhang, Y. Li, X. Hao and Y. Fang "A Clustering Algorithm of Cluster-head Optimization for Wireless Sensor Networks Based on Energy" Journal of Information & Computational Science, Vol.8, Issue 11, pp. 2129–2136, 2011.

[27] R.K.Tripathi, Y.N. Singh and N.K.Verma "Clustering algorithms for non-uniformly distributed nodes in WSNs" Electronics Letters, Vol.49 No.4, February 2013

[28] A. Dabirmoghaddam, M. Ghaderi and C. Williamson "On the optimal randomized clustering in distributed sensor networks" Journal of computer network, Vol. 59, Issue 11, pp. 17–32, February 2014.

[29] W. Li, P. Martins, and L. Shen "Determination method of optimal number of clusters for clustered wireless sensor networks" Journal Wireless Communications and Mobile Computing, Vol. 12, Issue 2, pp. 158 – 168, 2012

[30] Förster, A. Förster, and A. L. Murphy, "Optimal cluster sizes for wireless sensor networks: an experimental analysis," Ad Hoc Networks, Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering, Vol. 28, pp. 49–63, 2010.