

# A Review Study on Analytical Estimation of Optimal Number of Clusters in Wireless Sensor Networks

Vinay Kumar<sup>1</sup>, Sanjay B. Dhok<sup>2</sup>, Rajeev Tripathi<sup>3</sup> and Sudarshan Tiwari<sup>4</sup>

<sup>1,2</sup>*Department of Electronics Engineering, Visvesvaraya National Institute of Technology, Nagpur, Maharashtra, India,*

<sup>3</sup>*Department of Electronics & Communication Engineering, Motilal Nehru National Institute of Technology, Allahabad, Uttar Pradesh, India,*

<sup>4</sup>*National Institute of Technology, Raipur, Chhattisgarh, India*

[vk@ece.vnit.ac.in](mailto:vk@ece.vnit.ac.in); [sanjaydhok@gmail.com](mailto:sanjaydhok@gmail.com); [rt@mnnit.ac.in](mailto:rt@mnnit.ac.in); [stiwari@mnnit.ac.in](mailto:stiwari@mnnit.ac.in)

## ABSTRACT

To improve energy efficiency, total network scalability and data aggregation in Wireless Sensor Networks (WSNs), sensor nodes are often grouped into disjoint and mostly non-overlapping subsets called clusters. To provide an energy-efficient system by clustering, there are three main challenges. First is to find the optimum number of nodes in a specific cluster, second is to find the optimum number of clusters in the network and the third is to find the optimum position of Cluster Head (CH) in a specific cluster. Selecting an optimum number of clusters in WSNs provide greater improvement in terms of system scalability, energy efficiency, collision reduction, network lifetime, latency, and efficient routing backbone in the network. Selection of optimal number of clusters in WSNs is affected by level at which WSNs is modeled viz. Radio Energy Model Level, Network Model Level and Clustering Level. The objective of this paper is to present a state-of-the-art survey of distinct analytical methods used to calculate the optimum number of clusters, and its time-line comparative analysis based on network type, mathematical formula for an optimal number of clusters, base station positioning, energy model, strengths, weaknesses and applications of WSNs. We have also discussed the impact of different parameters on selecting the optimal number of clusters in WSNs.

**Keywords:** Wireless Sensor Networks, Clustering, Optimal Clustering, Energy Efficient WSNs, Optimal Number of Clusters, Algorithms for Optimal Number of Clusters

## 1 Introduction

WSNs are large-scale networks of small embedded devices, each with sensing, computation and communication capabilities and have been widely discussed in [1, 2, 8, 20]. In WSNs, sensor nodes have limited processing power, communication bandwidth, and storage space, which demand very efficient resource utilization. The sensor nodes are often grouped into individual disjoint sets called cluster [3, 4]. Clustering is used in WSNs [5, 6], as it provides network scalability, resource sharing and efficient use of constrained resources that give network topology stability and energy-saving attributes. Clustering schemes, offer reduced communication overheads, and effective resource allocations thus decreasing the overall energy consumption and reducing the interferences among sensor nodes. In sensing field if

we have more number of clusters while maintaining the same load per Cluster Heads (CHs), 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 Base Station (BS) will include more cluster heads to the cluster head hops, which mean higher overall energy consumption. Accordingly, finding the optimal number of clusters is a very crucial point in the system [57]. Selecting an optimal number of clusters in WSNs provide greater improvement in terms of energy efficiency, system scalability, network lifetime, and latency. Cluster optimization does not play a significant role for moderate size sensor networks if free space fading energy is low, but for large networks, cluster size optimization is still important even if free space fading is low [51]. The optimal number of clusters is very sensitive to energy model and sensing model of the nodes used in the system. The objective of this paper is to present a state-of-the-art survey of distinct analytical methods used to calculate the optimal number of clusters and its comparison based on, network lifetime, expression for an optimal number of clusters, base station position, energy model used, advantages, disadvantages and applications in WSNs. We have also discussed the impact of different parameters on selecting the optimal number of clusters in WSNs.

Though there are number of survey papers on the topics of WSNs and clustering in WSNs, but none of the research papers surveyed algorithms for an optimal number of clusters analytically or theoretically in WSNs. To the best of our knowledge, this is the first survey paper which carried out a review study on different algorithms for finding optimal number of clusters analytically in WSNs. The rest of the paper is organized as follows: in section 2, we provide a basic idea for clustering, cluster characteristics and the need for clustering in WSNs. Section 3 presents optimal clustering & factors effecting the optimal number of clusters in WSNs. Section 4 presents a survey on state-of-art of different algorithms for finding the optimal number of clusters analytically, reported in the literature. The open issues and challenges in WSNs are discussed in 5 and finally paper is concluded in section 6.

## **2 Clustering, Cluster Characteristics and the need for Clustering in WSNs**

### **2.1 Clustering in WSNs**

To support high scalability and better data aggregation, sensor nodes are often grouped into disjoint and mostly non-overlapping subsets called clusters. In the clustering, each cluster has a leader, which is called the CH and it performs the tasks like fusion and aggregation of data. Figure 1 represents clustering in sensor networks along with inter cluster and intra cluster communications, CHs and sensor nodes.

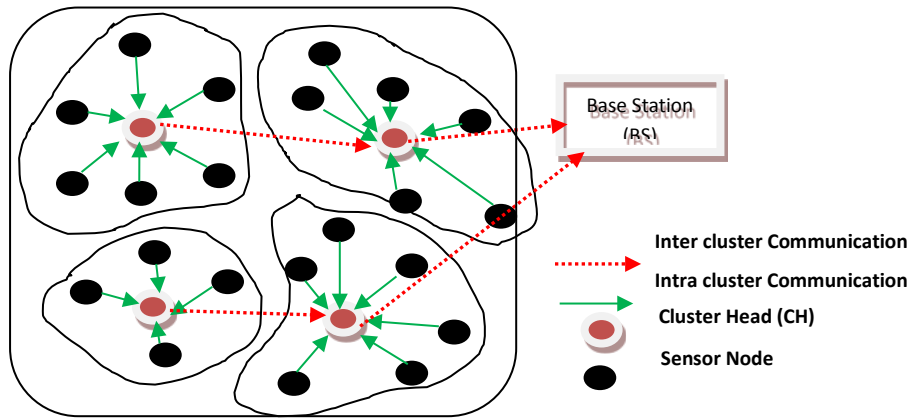


Figure 1: Clustering in WSNs with inter and intra cluster communication

The sensor nodes in a particular cluster periodically transmit their data to the CH nodes. CH nodes aggregate the data and transmit to the Base Station (BS) either using single-hopping or multi-hopping communication. The communication between CH and nodes is called intra-cluster communication and communication between CHs and base station is called inter-cluster communication [9]. The definitions of some terms are as follows: **Cluster head (CH)**: CH aggregates the data sensed by the cluster members (sensor nodes) in a particular cluster and aggregated data will be transmitted to BS. **Base station (BS)**: It has high processing capabilities and high level of energy. BS is the co-ordinator of the network where all the aggregated data from CHs are processed. **Sensor node**: Most of the nodes in the network, which are neither CHs nor BS, are considered simple sensor nodes.

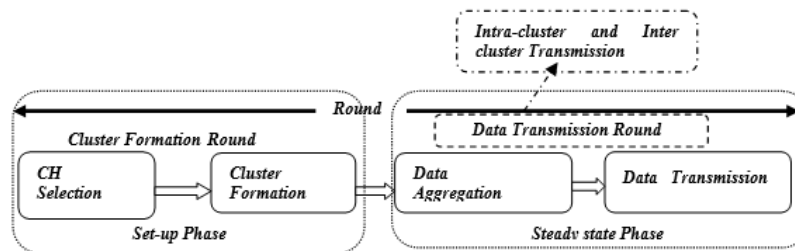


Figure 2: Phases in single round of clustering techniques

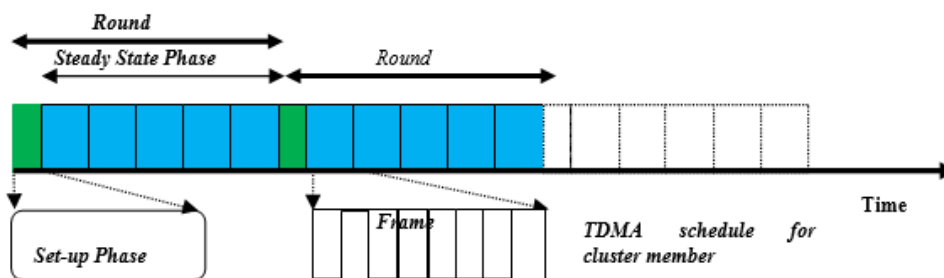


Figure 3: Timing diagram of phases in single round of clustering techniques

Cluster-based algorithms work in four stages: CH selection, cluster formation, data aggregation and data transmission. Most of the clustering protocols, divide the schedule of the network into different rounds of fixed duration. Figure 2, shows that each round consists of a setup phase and steady-state phase. During the set-up phase some sensor nodes elect themselves as CHs. The steady state phase, which is

sub divided into data aggregation and data communication. Steady state phase divided into different frames. During the steady-state phase, within each frame the cluster heads receive sensor data from cluster members (according to some multiple access technique and MAC protocol), and transfer the aggregated data to the BS [15]. Figure 3 represent timing diagram of phases in single round of clustering techniques.

## 2.2 Cluster characteristics

Cluster characteristics are very important in the clustering process of WSNs. Figure 4 shows the taxonomy of cluster characteristics in WSNs. Some cluster characteristics are defined as follows: **Cluster Changeability**: Clustering techniques can be classified into two types: fixed and variable ones. In the fixed techniques, the set of CHs are predetermined and the number of CHs is fixed. In a variable technique number of CHs is variable, in which CHs are selected randomly, from the deployed sensor nodes [15]. **Cluster Sizes**: It can be classified into two types: uniform (same size clusters) and non-uniform (different size clusters), in the network. **Intra-Cluster connectivity**: This characteristic classified based on basis of communication inside a particular cluster; it includes two classes: single-hop and multiple-hop intra cluster connectivity and **Inter-Cluster Connectivity**: this characteristic classified on the basis of communication between the base station and cluster heads; its include two classes namely single hop and multi-hop inter cluster communication.

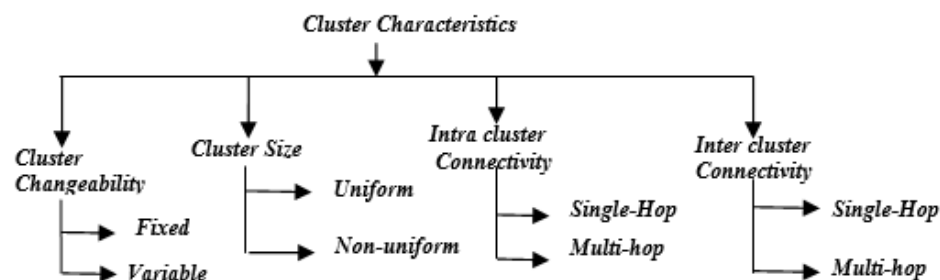


Figure 4: Taxonomy of Cluster characteristics

## 2.3 Why clustering in WSNs?

1. The advantages of clustering are many and are listed as follows:
2. It enables bandwidth reuse, thus can improve the system capacity within a cluster.
3. All the normal nodes send their data to the CHs so that energy saving is achieved by absence of flooding, multiple routes, routing loops [9, 14]
4. Clustering enables efficient resource allocation and thus helps in better designing of power control. Clustering facilitates data aggression/data fusion.
5. Any changes of node's behavior within a cluster affect only that cluster, but not the entire network, which make it robust to these changes [25].
6. Since the backbone network consists only the CHs, which are fewer in number than all the sensor nodes in the entire network. It requires less storage of routing information.
7. Clustering schemes make it easier in responding to changes caused by network dynamics, node mobility, unpredicted failures and local changes. Since these changes need to be managed and detected within individual clusters only [11, 12].

8. If the sensor nodes are mobile in nature, then nodes residing in concern clusters only need to update the information. Thus, local changes need not be updated by whole network, and this reduces the information processed and stored by each mobile sensor node [13].
9. In clustering, only CHs are responsible for transmitting data to base station this will reduce data collisions between the nodes [77, 78].
10. Generally, sensor network uses multi-hopping to transfer data to the BS. In this, the traffic transmitted by each node includes both self and relayed data. The sensor nodes closer to the BS have to transmit more data than those far away from the BS. As a result, the nodes closest to the BS heavily deplete their energy, creating a hole near the BS. So with the help of clustering hole problem can be reduced [17, 18, 19,79].

### 3 Optimal Clustering in WSNs

The main idea behind 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 collecting data from the whole network is minimized as compared to other possible clustering patterns [67].

#### 3.1 Why Optimal Clustering in WSNs?

If the clusters are not constructed in an optimal way, the total energy consumed by the sensor network per round is increased exponentially when the number of clusters created is greater especially when the number of the constructed clusters is less than the optimal number of clusters [30].

- i. In sensing field, choosing more clusters while maintaining the same load per Cluster Heads (CHs), 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 [57].
- ii. In WSNs from the **Physical (PHY) layer** point of view, using a large number of clusters can reduce energy consumption because the communication distance between CHs can be reduced. From the **Medium Access Control (MAC)** layer point of view, using a lesser number of clusters can reduce energy consumption because it decreases the average possibility of being a cluster head for each sensor node and from **Network Layer** point of view, lesser clusters yield fewer hop counts to the data sink, and result in less energy consumption. So for a cross-layer trade-off design issue among the required communication power in the physical layer, the possibility of being CHs in the MAC layer and the number of hops in the relay path in the network layer. So we have to optimize the number of clusters in WSNs [53].
- iii. Consider the possibility of processing data inside the cluster. After processing of the data, it will be transferred to CHs. Energy consumption decreases with increasing cluster sizes because data traffic decreases and data aggregation rate grows. However, for very large cluster, the performance is rather irrelevant, so optimal number of clusters should be selected [76].

With the help of above given statements, we can say that optimal clustering in WSNs plays a great role. It provides benefits like limited resources can be utilized more efficiently, overall energy efficiency is improved and sensor network lifetime is improved.

### 3.2 Parameters affecting an optimal number of clusters at different levels in WSNs

The parameters affecting an optimal number of clusters in WSNs are divided into three levels: Radio model level, Network level and clustering level. Table 1 represents all the factors affecting the optimal number of clusters in WSNs.

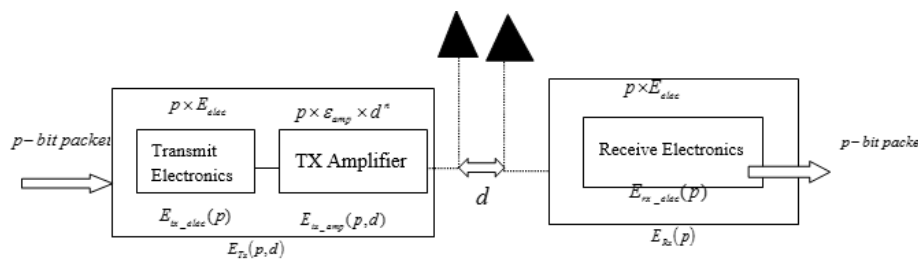
**Table 1: Parameters Affecting Optimal number of clusters at different levels in WSN**

Level	Parameters
Radio Model Level	Energy Models
	Sensing Model
	Shadowing and path loss Exponent
Network Level	Node Density
	Transmitter and Receiver Circuit
	Size of Sensing Fields
	Number of Base Stations
	Base Station Positioning
	Single and Multi-Hopping
Clustering Level	Distance Between Cluster heads and Position of Cluster Heads
	In-network Processing
	Data Correlation and Distortions

The parameters affecting optimum number of clusters are described below:

#### 3.2.1 Radio Model Level Parameters

- i. **Energy Model:** Optimal number of clusters depends highly on the type of energy model used. Therefore, it is important to use the right energy model. In WSNs there are four energy model which are generally used [10, 23, 24, 25]. Most of the surveyed algorithms in this paper using energy model represented and explained by figure 5 [10].



**Figure 5: Radio Energy Dissipation Model**

The energy consumption model can be simplified for a transmitter–receiver pair at distance  $d$  apart as Follows  $E_c = E_{Tx}(p, d) + E_{Rx}(p)$ . Where  $E_{Tx}(p, d)$  and  $E_{Rx}(p)$  are the energy consumption of the transmitter and the receiver, respectively.  $E_{elec} = E_{tx\_elec}(p) = E_{rx\_elec}(p)$  = Energy dissipated to run the transmitter or the receiver circuitry to transmit or receive one bit of the data packet.

$\epsilon_{amp} = E_{tx\_amp}(p, d)$  is energy dissipation of the transmission amplifier to convey one bit of data packet to the receiver node at distance of  $d=1m$  away

The energy consumption at the transmitter is divided into the transmit electronics and transmitter amplifier while the receiver energy consumption depends only on the receiver electronics. Then, the transmitter and receiver energy consumptions are:

$$E_{Tx}(p, d) = pE_{elec} + p\epsilon_{amp}d^n \text{ and } E_{Rx}(p) = pE_{elec}$$

where  $p$  = Length of transmitted/ received message in bits,  $d$ =distance between transmitter and receiver node,  $n$ =path loss exponent.

$n=2$  for free space model ( $\epsilon_{amp} = \epsilon_{fs}$ , when  $d < d_0$ ) and  $n=4$  for Multipath Model ( $\epsilon_{amp} = \epsilon_{mp}$  when  $d > d_0$ )

$$l\epsilon_{fs} \times d_0^2 = l\epsilon_{mp} \times d_0^4 \quad d_0 = \sqrt{\frac{\epsilon_{fs}}{\epsilon_{mp}}}$$

**Table 2: This data is well taken from [51], optimal number of clusters with different energy models, for 100 numbers of sensor nodes(N) and side of sensing area  $M= 100$ , when the distance between the CH and the sink node is between 45–145m.**

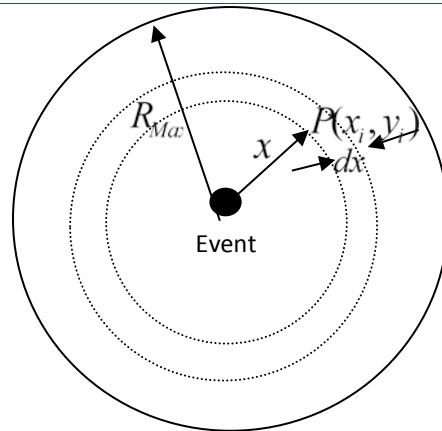
Energy Model	$K_{opt}$	
	$\epsilon_{fs} = 7nJ / bit / m^2$	$\epsilon_{fs} = 10pJ / bit / m^2$
	N=100, M=100	N=100, M=100
Halgamuge et al.[51]	1-6	0-2
Zhu et al.[24]	1-12	0-4
Mille et al.[23]	2-13	0-5
Heinz et al.[31]	2-16	1-7

The difference in the optimal number of clusters ( $K_{opt}$ ) between these energy models is getting closer as the distance between the sink and the CHs increases, This is because, as this distance increases the energy dissipation for communication becomes more and more dominant in the cost function.

- ii. **Sensing Model:** It affects the optimal number of clusters because the distance between the cluster head and base station change with change of sensing model of the nodes. Basically there are two types of sensing models reported in literature. The first model is deterministic, and the second model is probabilistic (Boolean sensing model, shadow-fading sensing model and Elfes sensing model). According to this model [26, 27], the probability that a sensor detects an event to a distance  $x$

Where  $R_1$  defines the starting of uncertainty in sensor detection and the parameters  $\psi$  and  $\nu$  are adjusted according to the physical properties of the sensor.  $R_{max}$  is the maximum sensing range of the node

$$p(x) = \begin{cases} 1 & x \leq R_1 \\ Exp(-\psi(x - R_1)^\nu) & R_1 \leq x \leq R_{Max} \\ 0 & x \geq R_{max} \end{cases} \quad (1)$$



**Figure 6: Probabilistic sensing model**

With the help of above given probability density function one can find the distance between the base station and cluster heads. Figure 6 represents the probabilistic sensing model.

- iii. **Shadowing and Path Loss Exponent Model:** The impact of the shadowing effect on optimal number of clusters is more significant for a larger value of the path loss exponent ( $n$ ) than that for a smaller  $n$ . A larger  $n$  may amplify the shadowing effect on the optimal number of clusters [53, 62].

### 3.2.2 Network Level Parameters

- i. **Node Density:** Experimental results show that for lower densities (275-375 nodes/km<sup>2</sup> approximately) the optimal cluster size is large and for higher densities (400-500 nodes/km<sup>2</sup> approximately) the optimal cluster size is one-hop when sensing field size is fixed. This is due to total intra cluster communication overhead for low density nodes [57].
- ii. **Transmitter and Receiver Circuit Energy:** The energy consumption of the transmitter circuitry  $E_{tx\_elec}$  has no impact on the optimal number of clusters. The energy consumption of the receiver electronics  $E_{rx\_elec}$  can greatly change the optimal number of clusters. It is a very important factor that can decide on whether or not it is worth performing clustering in the sensor network [62].
- iii. **Size of Sensing Field:** The optimal number of clusters can be independent of the sensing field size under the following conditions [7, 62]:
  - ✓  $E_{rx\_elec}$  is small compared to  $E_{tx\_elec}$
  - ✓ The wireless transmissions are governed by free space radio propagation model
  - ✓ The BS is not located outside of the sensing field.

If the above requirements are satisfied, the optimal number of clusters will only depend on the number of sensor nodes across the network  $K_{opt} = f(N)$ . If the sensing area is square-shaped or circular, the optimal number of clusters can be expressed as  $K_{opt} = \tau \sqrt{N}$  where  $\tau$  is a constant with maximum value is one. Table V shows the impact of sensing field, base station position, path loss model on selecting the optimal number of clusters, which support the justification of the above statements.



- iv. **Number of base stations:** Inter cluster communication increases when the number of base stations increases. For large network the energy consumption is nearly the same for single and two hop clusters. However, with three base stations and large networks, the optimal cluster size is 2 hops for reducing energy consumption in the network [57].
- v. **Base Station Positioning:** Optimal number of clusters will be larger when BS is located in the center of the sensing field. As BS moves away from the center of the sensing field area towards boundary to outside of sensing field, optimal number of clusters will be reduced.
- vi. **Single and Multi-hopping:** The impact of selecting the optimal cluster size on total energy consumption is more prominent in single-hop communication than in a multi-hop WSNs. This is because the energy function is proportional to the square of the distance over which data transmission is done and this distance for the single-hop communication is often longer than that for the multi-hop case. When the degree of data correlation is high, then the wider range of cluster sizes will occur. This is also the reason why the optimal number of clusters in multi-hop communication gets larger than that of single hop approach as data correlation degree increases [67]. Single-hop clustering performs best for a large spectrum of different size of sensing field, node densities and the number of base stations. For very high density networks (more than 1000 nodes), multiple base stations (more than three) or very low density network (less than 400 nodes) 2-hop clustering performs better [57].
- vii. **Free Space Fading Energy:** Selecting a number of clusters does not play a key role for reasonable size sensor networks if free space fading energy is low. But for large-scale networks, finding the number of clusters is still important, even if free space fading is low [57].

### 3.2.3 Clustering Level Parameters

- i. **Distance between Base station and CHs:** Selection of optimal number of clusters also depends upon the distance between CHs and BS. The distance between the cluster head and base station depends upon the size of sensing field and type of sensing models.
- ii. **Position of the cluster head:** The Optimal number of clusters will be large when BS is located in the center of the sensing field. As BS moves from the center of the sensing field area towards a boundary to outside of *sensing field*, *optimal number of clusters will be reduced*. *The position of the CH is also important because of two reasons*. First, it affects the load balance, and therefore energy consumption inside the cluster for data routing aggregation. Second, the overhead during routing of data inside the CHs with the head placement [57].
- iii. **In Network Processing:** Consider the possibility of processing data inside the cluster. After processing of the data, data will be transferred to CHs. Energy consumption decreases with increasing cluster sizes because of the fact that data traffic decreases and data aggregation rate grows. However, for very large clusters the performance is rather irrelevant; therefore preference should be given to 3-4 hop clusters since they have simultaneously low energy consumption and low data aggregation rate [57].
- iv. **Data correlation and distortion:** Finding the optimal cluster size depends on the value of the correlation. A large cluster size is optimal for low correlation and a small cluster size performs optimally for high correlation, there exist intermediate cluster sizes that perform near optimally over a wide range of spatial correlations. This near-optimal cluster size depends only on base station position and total number of nodes in the sensing field [36, 67].

## **4 Algorithms for analytically estimating the optimal number of clusters in WSNs:**

There have been several different criteria to initially classify the algorithms for finding optimal number of clusters in WSNs. Two of the most common classifications are the algorithms for homogenous sensor networks the algorithms for heterogeneous sensor networks. Both classifications are based on the characteristics and functionality of sensor nodes in the cluster. Figure 7 shows the taxonomy of various algorithms for an optimal number of clusters. Table 3 represents symbols and their meaning used in different algorithms surveyed. Table 4 represents comparative analysis of algorithms based on network type, analytical value of optimal number of clusters and base station positioning. Table 5 presents comparative analysis of algorithms based on sensing field, radio model, base station position and analytical value of optimal number of clusters particularly for [62, 81, and 82]. Table 6 presents timeline comparative analysis of algorithms based on cluster variability, node distribution, energy model and type of sensing fields and finally table 7 presents comparative analysis of algorithms based on their strengths, weaknesses and applications.

### **4.1 Homogeneous sensor networks**

A homogeneous sensor network consists of identical sensor nodes. It means that the sensor nodes have the same energy and hardware complexity. For static clustering in a homogeneous network, it is shown that the cluster head (CH) nodes will be over-loaded when long range transmission is required to transmit data to remote base station. In a heterogeneous sensor network, the nodes are enabled with extra battery energy and extra hardware. Sensor nodes in a particular cluster, use multi-hopping to send data to the respective cluster head. The sensor nodes that are closest to the cluster head have the highest energy burden due to relaying. When the sensor nodes use single hopping to transfer data to the cluster head, the sensor nodes that are farthest from the cluster heads always spend more energy than the sensor nodes that are closer to the cluster heads. Non-uniform energy consumption in the network takes place. Thus, there are two desirable characteristics of a sensor network: minimum hardware cost, and uniform energy consumption. While heterogeneous networks achieve the minimum hardware cost, the homogeneous networks achieve uniform energy [70, 83]. Most of the known algorithms for finding optimal number of clusters for WSNs can be further distinguished into two main parts: probabilistic (random) and non probabilistic, depending on the cluster formation method and the parameters used for CH selection. Optimal Probabilistic algorithms are further divided into two parts based on node distribution: uniform and non-uniform.

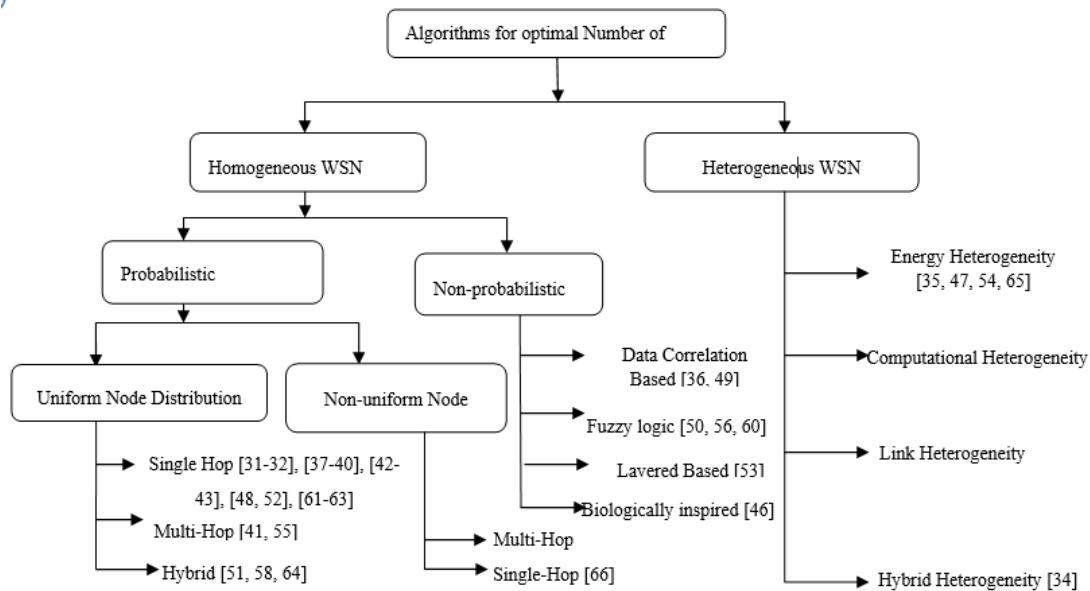


Figure 7: Taxonomy for the optimal number of clusters in WSNs

#### 4.1.1 Probabilistic and Uniform Node Distribution

The main aim of probabilistic algorithm is to reduce the energy consumption and prolong the networks lifetime. Some of the algorithms follow a random approach for CH selection, whereas other follows a hybrid probabilistic model for CH selection.

Heinzelman et al. proposed [31], Low Energy Adaptive Clustering Hierarchy (LEACH) protocol, is the first and the oldest algorithm for the calculation of optimal number of clusters in WSNs. It uses the following techniques to achieve the design goals: randomized, self-configuring and adaptive cluster formation, local control for data transfers and low-energy media access control and application-specific data processing. LEACH protocol has many rounds, and each round has two phases, a setup phase and steady-state phase. In the setup phase; it provides cluster formation in an adaptive manner, and in the steady-state phase transfer of data take place. 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 [80]. Depedri et al. [32] proposed decentralized algorithms for cluster formation in which sensor node only knows about its own position and position of final receiver, and not the position of all sensor nodes. It operates in following phases: Cluster head selection algorithm, Cluster formation and data transmission with multiple accesses. Each sensor node chooses its cluster head by evaluating the energy dissipated in the path between final receiver and itself. It provides better energy efficiency than LEACH. In this, each sensor can determine its own optimum number of clusters. This value depends on the total number of sensors in the network (N), path-loss exponent (n), the dimension of sensing field. It does not depend on the distance between the cluster head, and the base station like instead was in original LEACH [31].

Chen et al. [39] have determined two aspects of the problem: first for given the number of clusters, how does one choose the CHs to cover up the sensor network more efficiently? Second, how does one assess the number of clusters needed to utilize data correlation of sensors for a general sensor network? Comeau et. al. [38] have proposed multi-level clustered network based on single-hop communication.

Authors have analytically determined the optimal number of clusters at each level. Kim et. al. [37] has estimated the optimal number of clusters among random distributed sensors in a bounded sensing field. In this algorithm, optimal number of CHs depends on the distance between the base station and sensor node. Vljajic et al. [40] have proposed that in under some specific condition clustered WSNs provide an advantage over non clustered WSNs. This algorithm considerably more efficient in terms of energy conservation than clustering algorithms previously proposed in the literature.

Wang et al. [43] proposed clustering algorithm, which provides fixed optimum number of clusters in each round during the whole lifetime of the network. In LEACH protocol, optimum number of clusters in each round depends on the total number of nodes in the network, but theoretically, the total number of sensor nodes decreased and the optimal number of clusters should be updated too with the lifetime of the network. So authors have suggested that the dynamic optimum number of cluster base on the real time of sensor nodes in the whole network, and it brings practical optimized effect in WSNs. Yang et al. [42], reduced the energy consumption of the network and tried to avoid the severe synchronization needs of Time Division Multiple Access. This algorithm is first accessible for small networks, under the hypothesis of identical expected distance of all CHs from the BS. Then it is extended for large networks to consider the case when the distances of various sections of the network from the BS may be different. This new architecture is easily implementable, scalable, and reduces the hardware complexity of the sensor nodes.

Chan et al. [48] proposed a Fixed Optimal Cluster (FOC) numbers that is to examine the entire network. There are two different types of the optimal cluster numbers depending on the position of the base station. One is that the base station is set up far farther away from the sensing field, and another is that the base station is set up at the center of the sensing field Results show that network lifetime enhanced very well. Shunjie et al. [52] proposed new model for finding the optimal number of CHs which is based on the LEACH energy dissipation model. Optimal number of clusters depends upon the number of frames in steady-state phase and the distance between CHs and the BS.

Yang et al. [61] have proposed a more reasonable energy consumption model called Optimal Energy Consumption Model (OECM) in a homogeneous network. It shows that the optimal number of cluster heads not only depend on node density, but also on the size of sensing field, circuit energy dissipation, packet length. Navid et al [62], analytically provides the optimal number of clusters that minimizes the total energy expenses in the networks, where all sensor nodes communicate data through their elected CHs to the BS in a distributed fashion. The results show that: 1 Under certain condition, the optimal number of clusters can be independent of the size of sensing field 2: The energy consumption of the transmitter circuitry has no impact on the optimal number of clusters, and 3: The energy consumption of the receiver electronics can substantially change the optimal number of clusters and more importantly; it can decide on whether or not is it worth performing clustering. Li et al. [55], have proposed, an uneven virtual grid-based clustering routing protocol. It works in concentric circular forms through the base station-assisted positioning. It's providing better energy efficiency and balanced network load.

Xia et al. in [41], proposed Local Negotiated Clustering Algorithm (LNCA) that tries to minimize the overall energy cost of the network by using the similarity of nodes. Experimental results of the algorithm

indicate that 2-hop LNCA outperforms existing LEACH [31] algorithms in terms of energy consumption. Optimal number of clusters depends on the sensing field area, node density in the network the length of data, the rate at which data is generated at each node and the length of time for inter cluster communication. Halgamuge et al. in [51] have developed an energy model and used this model to assess energy expenditure and node lifetime for a sensor network with fixed configuration. The authors also have two observations 1: The optimal number of clusters increases with the increase of free space fading energy 2: The number of clusters does not play a significant role for realistic size sensor networks if the free space fading energy is low. For large networks, cluster optimization is still important, even if free space fading energy is low. Li et al. [58] presented an analytical model for finding the optimal number of clusters, for minimizing the communication costs in a clustered sensor network. In this, sensor nodes are placed in the sensing field in a random and distributed manner according to a homogeneous poisson point process.

Tandon in [64], focused on the analysis of extravagant energy consumption within a uniform CH election model and suggested a model to reduce the overall consumption of energy usage amongst the CHs in the WSNs. The optimal number of clusters depends on the distance from the BS. Chen et al. [63] proposed CH optimization based on energy. In this, the authors considered a threshold value and the residual energy of nodes, to optimize the selection of a cluster head. Results show this algorithm can prolong the network lifetime efficiently compared with LEACH protocol. The algorithms based on uniform node distribution have following problems:

- ✓ Cluster head selection is uncontrolled.
- ✓ Uniform sensor node distribution.
- ✓ Cannot be applied to all the practical cases.

#### **4.1.2 Probabilistic and Non-uniform Node Distribution**

Tripathi et al. [66] introduced clustering of non-uniform random distributed nodes, and calculated the optimal number of clusters in WSN. Results show that there is balanced energy expenditure in this non-uniform clustering. Dabirmoghaddam et al. [67], that the general problem of optimal clustering is NP-hard. They try to optimize the algorithm and produce the best possible clustering of the network in terms of energy consumption. It is found that with non-uniform clustering in heterogeneous WSNs, clusters are more energy-efficient in WSNs with spatial data correlation.

#### **4.1.3 Non-probabilistic Algorithms**

In this type of clustering algorithms, selection of cluster heads and cluster formation is based on deterministic criteria such as connectivity, degree and information received from the other closely located group.

Pattem et al. [36] found the optimal number of clusters, which depend upon the distance from the cluster head to the BS and the degree of correlation. As we know that the small cluster sizes and large cluster sizes perform well at low and high correlations respectively. Nevertheless, it appears that an intermediary cluster size performs correctly across the range of correlation values that is called "near-optimal" static cluster size. The value of near optimal cluster size depends only on BS position and the number of nodes. Chen et al. [49] worked on a multi-cluster sensor network, which is used for source extraction in a sensing field. Performance of source extraction and the complete energy consumption in

the sensor network depends upon the number of clusters. Performance of source extraction and the entire energy consumption in the sensor network can be affected by different factors: unreliable wireless links, sensor failure, etc.

Raghuvanshi et al. [50] have used an objective function based method to divide a data set into a set of clusters. In comparison to standard clustering, fuzzy clustering offer to assign a data point to more than one cluster, so that overlapping clusters can be handled confidently. In this paper Gustafson- Kessel (GK) algorithm is used for partitioning the sensor nodes into clusters and optimal number of clusters are determined using Xie-Beni validation index. This algorithm uses fuzzy clustering for partitioning the data into a pre-defined number of clusters with fuzzy boundaries. G-K clustering algorithms can find clusters of approximately equal areas, thus load balancing is good. In a G-K clustering, number of clusters is a range of numbers rather than a unique number. G-K clustering provides more coverage range comparisons to LEACH clustering algorithm. Selvakennedy et al. [46] tried to find the appropriate number of clusters with well-balanced memberships. The proposed algorithm is speedy with very partial overhead. Due to the robustness of any biologically-inspired algorithm, this protocol could handle unexpected circumstances in the environment and node failures.

Wang et al. [53] proposed the main challenges in deploying a high dense cluster based sensor network. In this paper, the authors have considered a basic Observational Area (OA) in WSNs and determined optimal numbers of clusters in basic OAs from the cross layer approach. The authors find a number of clusters from different layer aspects, and they have shown and concluded that from the physical layer point of view if there are more clusters results in more energy savings. From MAC layer and network layer point of view, fewer cluster results in more energy savings. Hence, authors have considered and developed a cross layer optimized model for physical, MAC, and Network Layer (PHY/MAC/NET). They have shown that optimizing the number of clusters in a sensor network becomes a cross layer tradeoff design issue among the required transmission power in the physical layer, cluster representative in the MAC and hop count in the network layer.

Wang et al. [56] proposed fuzzy based clustering, as we know that Fuzzy c-Means (FCM) and its derivatives suffer from two problems: local minima and cluster validity—which have a straight impact on the formation of the final clustering. The problem of local minima can be short out by optimization and center initialization strategies. This paper proposes a center initialization approach based on a minimum spanning tree to keep FCM from local minima. Raghuvansi et al. [60], describes a method for finding a fuzzy membership matrix that provides cluster membership values for all the objects based on the proximity matrix. Membership matrix related with fuzzy is found by first finding a set of vectors that approximately have the same inter-vector euclidean distances as the proximate that are provided. A dimension of these vectors can be very low (less or equal to 5). FCM algorithm is used for the optimal number of cluster calculation in WSNs. The authors found that this method to be very effective and no more computationally expensive than other relational data clustering methods. The FCM algorithm is more energy efficient compared to a G-K clustering algorithm. There are several validity measures proposed in the literature, given the optimal number of clusters following seven measures are considered. 1: Partition Coefficient (PC) (L), 2: Xie and Beni's Index (XB) (L), 3: Classification Entropy (CE) (M), 4: Partition Index (SC) (L), 5: Separation Index (S) (L), 6: Dunn's Index (DI) (M), 7: Alternative Dunn Index (ADI) (L).

## 4.2 Heterogeneous sensor networks

There are three common types of hardware heterogeneities in WSNs: **Computational heterogeneity** in which the heterogeneous node has a more powerful microprocessor, and more memory, than the normal node. **Link heterogeneity**: In which some nodes have long-distance highly reliable communication links (IEEE 802.11 connectivity) than a normal node. **Energy heterogeneity** where nodes have unlimited energy resources or battery is replaceable. The most important heterogeneity is the energy heterogeneity because both computational heterogeneity and link heterogeneity will consume more energy resources. Heterogeneity provides improvement in terms of response time, reliability and lifetime [68, 69]. If there is no energy heterogeneity, computational and link heterogeneity will impact the whole sensor network negatively, resulting decrement of the network lifetime [28, 30, 47].

Mhatre et al. [34] determined expressions for the required number of CH along with type of communication like single hopping, multi-hopping and hybrid modes. Authors have two types of sensor nodes, which are less robust due to heterogeneity of the nodes. If any cluster head (CH) nodes fail, the system no longer will function. In LEACH protocol, the system is more robust because every node is competent of acting as a CH, and hence the subside of a few nodes does not seriously affect the working of the whole system. Smaragdakis et al. [35] proposed, a heterogeneous-aware protocol; it had applications where the feedback from the sensor network must be reliable.

Kumar et al. [47] have introduced hierarchically clustered network in which sensor nodes have energy heterogeneity, means some of the nodes equipped with additional energy. Kumar et al. [54] have developed an energy-efficient cluster head election (EECHE) protocol for heterogeneous WSN with three types of sensor nodes. Some of the sensor nodes are equipped with the extra energy resources than the other nodes. Results show that EECHE can enhance the lifetime and stability of the system and provide better results than LEACH protocol. Tuah et al. in [65] proposed heterogeneous sensor networks and investigated the optimal number of clusters, which can minimize the energy consumption per round. In this two type of radio energy models are used, in which intra cluster communication using free space model and inter cluster communication using multipath model respectively.

It is seen that the performance of clustered WSNs is not always better than the performance of non-clustered WSNs under certain condition. The ability of clustered WSNs to outperform their non-clustered counterparts depends on number of nodes, the total number of cluster heads, and average factor of in-cluster data reduction [10].

**Table 3: Symbols and their meaning used in different algorithms surveyed**

Symbol	Abbreviation	Symbol	Abbreviation
$K_{opt}$	Optimal number of clusters or CHs	N	Number of sensor nodes
M	Side of the Square	$d_{toBS}$	Distance between the cluster head and base station
$E_{elec} = E_{tx\_elec}$ $= E_{rx\_elec}$	Energy dissipated to run the transmitter or the receiver circuitry to transmit or receive one bit of the data packet	$E_{DA}$	Data Aggregation Energy
$\epsilon_{fs}$	It is energy dissipation of the transmission amplifier to convey one bit of data packet to the receiver node with a distance of $d=1m$ away, for $n=2$ (Free Space)	$\epsilon_{mp} = E_{amp}$	It is energy dissipation of the transmission amplifier to convey one bit of data packet to the receiver node with a distance of $d=1m$ away, for $n=4$ (Multipath)
d	Distance between transmitter and receiver	$d_{toCH}$	Distance between sensor nodes and cluster head
m	Data compression rate	B	Distance from the centre of sensing area to the outside location of the base station.
$i^{nth}$	$i^{th}$ ring	$A_{net}$	Area of circular field
$R_{net}$	Radius of circular fields	$\lambda_i$	Variable density parameter
$j$	No. of annular bands of same radial width	$N_i$	No. of nodes in an annular band of $\lambda_i \times A_i$
L	Signal frame with L-samples	$E\{d_{toCH}^4\}$	Expected Value of distance between cluster heads and base station
$\lambda$	Intensity of homogeneous spatial Poisson process	$l^1 + \mu^1 d^{k^1}$	Energy Spent in transmitting a packet from cluster head to Base Station
$\alpha_i$	Hardware cost of the node	$\beta$	Used to model node battery
T	Data gathering cycle	$\mu$	Propagation Loss Term
n	Propagation loss exponent	$r$	Radio range of node
c	Average of $C_i$	$c_i$	Degree of correlation between a sensor and its neighborhood sensors
$f$	Number of data frames	$l_1$	Bit length of control information
$l_2$	Bit length of data information	T	The time to transmit one byte on the network
$T_{inter}$	time devoted for inter cluster communication	$k_{inter}$	$\frac{T_{inter}}{T}$
$d_{broad\ cost}$	Distance between the CH and the farthest point of observed area	$c_o$	Spatial Correlation
D	No. of hop between CHs and BS	$\rho$	Energy Multiplication factor of the super advanced node
$\gamma$	Data Fusion Rate	$\alpha$	Energy Multiplication factor of the advanced node
$m_0$	Percentage of Advanced Node	$m_1$	Percentage of super Advanced Node
$\eta$	Required received power	$L_o$	Path loss at distance $k_o$
$P_e$	Receiver electronics energy	$d_{oA}$	Minimal distance that results in uncorrelated information
R	Radius of circular Field	k	Propagation loss constant for



$\kappa^l$	Propagation loss constant for communication between the cluster heads and Base Station	$\mathbf{Z}$	communication between the cluster head Number of hops between node and Base station
$n_o$	Type "0" Nodes	$d^o$	Discrete time
$\mu A^K$	Amount of Energy Spent in the RF Amplifier to counter propagation Loss	$\delta$	Distance between two rings
$d_{toCh1}$	Average Distance between two CHs	$N_1$	Number of nodes in first ring
$E_{tranCH}$	Energy Dissipation due to operating mode at the CH per Round	$E_{loggCH}$	Energy consumed for logging sensor reading at the CH per round
$E_{SensCH}$	Total Energy dissipation for sensing activity at the CH per round	$L$	Position of BS from the centre of sensing field
$E_c$	Energy Consumed in computation of Data		

**Table 4: Comparative Analysis of algorithms based on network type, Analytical value of optimal number of clusters and Base Station Position (network topology)**

Paper Ref.	Network Type	The analytical value of $K_{opt}$	Base Station Position
[31]	Homogenous	$K_{opt} = \left( \frac{N}{2\pi} \times \frac{\epsilon_{fs}}{\epsilon_{mp}} \right)^{\frac{1}{2}} \frac{M}{d^2_{toBS}}$	Center of Sensing Field
[32]	Homogenous	$K_{opt} = \left[ \frac{N}{2} \times \left( \frac{M^2}{2\pi} \right)^{\frac{n}{2}} \times \frac{1}{(d_{broad\ cost})^n} \right]^{\frac{n+1}{2}}$	Outside of sensing field
[34]	Heterogeneous	$K_{opt} = \left\{ \frac{kn_o\beta T\mu A^k}{2(\alpha_1 + c\beta T(l^1 + \mu^1 d^k))} \right\}^{\frac{2}{k+2}}$	Outside of sensing field (Single Hopping)
		$K_{opt} = \left\{ \frac{n_o\beta T r_n^2(2l + \mu r^k)}{r^2(\alpha_1 + c\beta T(l^1 + \mu^1 d^k))} \right\}^{\frac{1}{2}}$	Outside of sensing field (Multi Hopping)
[35]	Heterogeneous	$K_{opt} = \sqrt{\frac{N}{2\pi}} \times \frac{2}{0.765}$	Center of Sensing Field
[36]	Homogenous	$K_{opt} = \sqrt{2Dc_o}$	Moving
[37]	Homogenous	$K_{opt} = \left\{ \frac{0.5855 \times N \times \epsilon_{fs} \times M^2}{\epsilon_{mp} d^4_{toBS} - E_{elec}} \right\}^{\frac{1}{2}}$	Outside of sensing field
[38]	Homogenous	$K_{opt,1} = M \cdot \left\{ \frac{N\epsilon_{fs}}{\pi(\epsilon_{mp}d^4_{toBS} - E_{elec})} \right\}^{\frac{1}{2}}$	Outside of sensing field
[39]	Homogenous	$K_{opt} = \left\{ \frac{cN}{2(1-c)} \right\}^{\frac{2}{3}}$	Center of Sensing Field
[40]	Homogenous	$K_{opt} = \sqrt{4Z - 1}$	Center of Sensing Field
[41]	Homogenous	$K_{opt} = \frac{1}{36} (1 + P + P^{-1})^2$ $P = (1 + 54N + 6 \times \sqrt{3 \times N + 81N^2})^{\frac{1}{3}}$	Center of Sensing Field
[42]	Homogenous	$K_{opt} = \left\{ \frac{N\epsilon_{amp}M^2}{3E_{elec}k_{inter} + 2\epsilon_{amp}M^2K_{data}} \right\}^{\frac{1}{2}}$	Center of Sensing Field
[43]	Homogenous	$K_{opt} = \left( \frac{N}{2\pi} \times \frac{\epsilon_{fs}}{\epsilon_{mp}} \right)^{\frac{1}{2}} \frac{M}{d^2_{toBS}}$	Inside Sensing field
[46]	Homogenous	$P_{opt} = \frac{1}{2} \times \sqrt{\frac{\epsilon_{fs}}{\lambda l \frac{\sqrt{6M}}{r} (2E_{elec} + \epsilon_{fs}r^2) - 2E_{elec}}}$	Center of Sensing Field

[47]	Heterogeneous	$K_{opt} = \left( \frac{N}{2\pi} \times \frac{\epsilon_{fs}}{\epsilon_{mp}} \right)^{\frac{1}{2}} \frac{M}{d^2_{toBS}}$	Center of Sensing Field
[48]	Homogeneous	$K_{opt} = \sqrt{N}$	Center of Sensing Field
		$K_{opt} = \sqrt{6N} \times \sqrt{\frac{\epsilon_{fs}}{\epsilon_{mp}}} \times \frac{M^2}{M^2 + 6B^2} \quad B \geq \sqrt{\frac{\epsilon_{fs}}{\epsilon_{mp}} - \frac{M^2}{6}}$	<b>When</b> $d_{toBS} \geq \sqrt{\frac{\epsilon_{fs}}{\epsilon_{mp}}}$ For multipath fading model
		$K_{opt} = \sqrt{\frac{N\epsilon_{fs}}{6\epsilon_{mp}}} \times \frac{M}{d^2_{toBS}}$	$d_{toBS} \leq \sqrt{\frac{\epsilon_{fs}}{\epsilon_{mp}}}$ <b>For free Space Model fading model</b>
[49]	Heterogeneous	$K_{opt} = \left\{ \frac{\epsilon_{fs} d^0 l N M^2 + 4\pi d^0 N^2 E_c}{2\pi[(8-2l)d^0 E_{elec} + 8\epsilon_{mp} d^0 E_{toCH}] + (2l+7)d^0 E_c} \right\}^{\frac{1}{2}}$	Above the center of field
[50]	Homogeneous	Numerical results is given	Center of Sensing Field
[51]	Homogeneous	$K_{opt} = \sqrt{\frac{N_s}{6}} \times \frac{M}{d^2_{toBS}} \times \sqrt{\frac{E_{fs}}{D_\alpha}}$ $D_\alpha = (E_{amp} + E_{sensCH} + E_{tranCH} + E_{log gCH})$	Out-side of field
[52]	Homogeneous	$K_{opt} = \left\{ \frac{(l_1 + l_2 f) N \epsilon_{fs} M^2}{2\pi[\epsilon_{mp}(l_1 d^4 + l_2 f d_{toBS}^4) - (2l_1 + l_2 f) E_{elec}]} \right\}^{\frac{1}{2}}$	Center of Sensing Field
[53]	Homogeneous	$K_{opt} = \begin{cases} \left( \frac{\left( \frac{n-1}{2} \right) \xi}{2P_e} \right)^{\frac{n}{2}} & \text{for } n > 2 \\ 1 & \text{for } n = 2 \end{cases}$ $\xi = \eta L_o \left( \frac{\sqrt{5} \times d_{OA}}{k_o} \right)^n$	Center of Sensing Field (Physical and MAC)
		$K_{opt} = \begin{cases} \left( \frac{z}{2P_e} \right)^{\frac{2}{n}} & \text{for } n > 2 \\ 1 & \text{for } n = 2 \end{cases} \quad z = \left( \frac{n-1}{2} \right) \times \xi \times 10^{Q^{-1}(\theta) \frac{\sigma}{10}}$	Center of Sensing Field (Shadowing effects)
[54]	Heterogeneous	$K_{opt} = \left( \frac{N}{2\pi} \times \frac{\epsilon_{fs}}{\epsilon_{mp}} \right)^{\frac{1}{2}} \frac{M}{d^2_{toBS}}$	Center of Sensing Field
[55]	Homogeneous	$K_{opt} = \sqrt{\frac{N_1}{2}} \times \frac{\delta}{\sqrt{E[d_{toCH1}^2]}}$	Center of Sensing Field
[56]	Homogeneous	Referred [56]	N/A
[57]	Homogeneous	based on Experimentally Analysis	N/A
[58]	Homogeneous	$K_{opt} = \left\{ \frac{3N^2 \epsilon_{fs}}{\pi \lambda (2M^2 \epsilon_{fs} - 3E_{elec})} \right\}^{\frac{1}{2}}$	The BS is located in one vertex of an square area
[60]	Homogeneous	Numerical Results given	Center of Sensing Field
[61]	Homogeneous	Numerical Results given	Center of Sensing Field

[62]	Homogeneous	$K_{opt} = \sqrt{\frac{N\epsilon_{fs}A}{2\pi(\epsilon_{amp}d_{toBS}^n - E_{rx\_elec})}}$	Dynamic base Station
[63]	Homogeneous	$K_{opt} = \left\{ \frac{NM^2\epsilon_{fs}}{0.342M^4\epsilon_{amp} - E_{elec}} \right\}^{\frac{1}{2}}$	Center of Sensing Field
[64]	Homogeneous	$K_{opt} = \left\{ \frac{3\epsilon_{fs}N(2i+1)^3}{2\epsilon_{mp}R_{net}^2((i+1)^6 - i^6)} \right\}^{\frac{1}{2}}$	Center of Sensing Field
[65]	Heterogeneous	$K_{opt} = \frac{M\sqrt{3N}}{\sqrt{2\pi}} \left\{ \frac{\epsilon_{fs}((1-(m_0+m_1))+\alpha m_0 + \rho m_1)}{E_{elec} \left( \frac{3}{\gamma} - N(1+\alpha + \rho) + \frac{3\epsilon_{mp}d_{toBS}^4}{\gamma} \right)} \right\}$	Center of Sensing Field
[66]	Homogeneous	$K_{opt} = \frac{\frac{3}{N^2} \times R}{\sqrt{2 \times \sum_{i=1}^i \lambda_i A_i R_i}}$	Center of Sensing Field
[67]	Homogeneous	Numerical Value is given	Center of Sensing Field

**Table 5: This table is well taken from [62] for comparative Analysis of Algorithm based on sensing field, radio model, base station position and Analytical value of optimal number of clusters**

Reference [62]				
Sensing field	Radio model	Location of BS	Optimal number of clusters	Optimal number of clusters (Small $E_{Rx}$ )
Square ( $M \times M$ )	Free space	Center	$K_{opt} = \sqrt{\frac{N\epsilon_{fs}M^2}{2\pi(\epsilon_{fs}\frac{M^2}{\gamma} - E_{rx\_elec})}}$	$K_{opt} = \sqrt{\frac{3N}{\pi}} = 0.977\sqrt{N}$
		Corner	$K_{opt} = \sqrt{\frac{N\epsilon_{fs}M^2}{2\pi(\epsilon_{fs}\frac{2M^2}{\gamma} - E_{rx\_elec})}}$	$K_{opt} = \sqrt{\frac{3N}{4\pi}} = 0.489\sqrt{N}$
		Side midpoint	$K_{opt} = \sqrt{\frac{N\epsilon_{fs}M^2}{2\pi(\epsilon_{fs}\frac{5M^2}{12} - E_{rx\_elec})}}$	$K_{opt} = \sqrt{\frac{6N}{5\pi}} = 0.618\sqrt{N}$
		Outside on the (Axis of symmetry)	$K_{opt} = \sqrt{\frac{N\epsilon_{fs}M^2}{2\pi(\epsilon_{fs}(\frac{M^2}{6} + L^2) - E_{rx\_elec})}}$	$K_{opt} = \sqrt{\frac{NM^2}{2\pi(\frac{M^2}{6} + L^2)}}$
	Two ray	Center	$K_{opt} = \sqrt{\frac{N\epsilon_{fs}M^2}{2\pi(\epsilon_{mp}\frac{7M^2}{180} - E_{rx\_elec})}}$	$K_{opt} = \sqrt{\frac{90N\epsilon_{fs}}{7\pi\epsilon_{mp}M^2}}$
		Corner	$K_{opt} = \sqrt{\frac{N\epsilon_{fs}M^2}{2\pi(\epsilon_{mp}\frac{28M^4}{45} - E_{rx\_elec})}}$	$K_{opt} = \sqrt{\frac{45N\epsilon_{fs}}{56\pi\epsilon_{mp}M^2}}$
		Side mid point	$K_{opt} = \sqrt{\frac{N\epsilon_{fs}M^2}{2\pi(\epsilon_{mp}\frac{193M^4}{720} - E_{rx\_elec})}}$	$K_{opt} = \sqrt{\frac{360N\epsilon_{fs}}{193\pi\epsilon_{mp}M^2}}$
		Out (on the axis of symmetry)	$K_{opt} = \sqrt{\frac{N\epsilon_{fs}M^2}{2\pi(\epsilon_{mp}(\frac{7M^4}{180} + \frac{2}{3}M^2L^2 + L^4) - E_{rx\_elec})}}$	$K_{opt} = \sqrt{\frac{N\epsilon_{fs}M^2}{2\pi\epsilon_{mp}(\frac{7M^4}{180} + \frac{2}{3}M^2L^2 + L^4)}}$
Circle (R)	Free space	Center	$K_{opt} = \sqrt{\frac{N\epsilon_{fs}R^2}{R^2\epsilon_{fs} - 2E_{rx\_elec}}}$	$K_{opt} = \sqrt{N}$
		Circumference	$K_{opt} = \sqrt{\frac{N\epsilon_{fs}R^2}{3R^2\epsilon_{fs} - 2E_{rx\_elec}}}$	$K_{opt} = 0.577\sqrt{N}$
		Outside	$K_{opt} = \sqrt{\frac{N\epsilon_{fs}R^2}{2(\epsilon_{fs}(\frac{R^2}{2} + L^2) - E_{rx\_elec})}}$	$K_{opt} = \sqrt{\frac{NR^2}{2(\frac{R^2}{2} + L^2)}}$

<b>Two ray</b>	<b>Center</b>	$K_{opt} = \sqrt{\frac{N\epsilon_{fs}R^2}{2(\epsilon_{mp}\frac{R^4}{3} - E_{rx\_elec})}}$	$K_{opt} = \sqrt{\frac{3N\epsilon_{fs}}{2\epsilon_{mp}R^2}}$
	<b>Circumference</b>	$K_{opt} = \sqrt{\frac{N\epsilon_{fs}R^2}{2(\epsilon_{mp}\frac{10R^4}{3} - E_{rx\_elec})}}$	$K_{opt} = \sqrt{\frac{3N\epsilon_{fs}}{20\epsilon_{mp}R^2}}$
	<b>Outside</b>	$K_{opt} = \sqrt{\frac{N\epsilon_{fs}R^2}{2(\epsilon_{mp}(\frac{R^4}{2} + L^4 + 2R^2L^2) - E_{rx\_elec})}}$	$K_{opt} = \sqrt{\frac{N\epsilon_{fs}R^2}{2\epsilon_{mp}(\frac{R^4}{3} + L^4 + 2R^2L^2)}}$

**Table 6: Time-line comparative analysis of algorithms based on cluster variability, year of publication, node distribution, energy model and type of sensing fields**

Reference	Year of Publication	Cluster Variability	Node Distribution	Energy Model Used	Sensing Fields
[31]	2002	Fixed	Uniform and random	[31]	Square
[32]	2003	Adaptive	Uniform and random	[63]	Square
[34]	2003	Adaptive	Uniform and random	[31]	Circular
[35]	2004	Adaptive	Uniform and random	[31]	Square
[36]	2004	Fixed	Uniform and random	N/A	Square
[37]	2005	Fixed	Uniform and random	[31]	Square
[38]	2006	Adaptive	Spatial Poisson	[31]	Square
[39]	2006	Adaptive	Uniform and random	[31]	N/A
[40]	2006	Adaptive	Linearly placed nodes	N/A	N/A
[41]	2007	Adaptive	Uniform and random	N/A	Square
[42]	2007	Fixed	Uniform and random	[31]	Square
[43]	2007	Adaptive	Uniform and random	[31]	Square
[46]	2007	Adaptive	Spatial Poisson	N/A	Circular
[47]	2008	Adaptive	Uniform and random	[31]	Square
[48]	2008	Fixed	Uniform and random	[31]	Square
[49]	2009	Adaptive	Uniform and random	[31]	Square
[50]	2009	Adaptive	Uniform and random	N/A	Square
[51]	2009	Adaptive	Uniform and random	[51]	Square
[52]	2009	Adaptive	Uniform and random	[31]	Square
[53]	2009	Adaptive	Uniform Distribution[75]	[31] & [71]	Square
[54]	2009	Adaptive	Uniform and random	[31]	Square
[55]	2009	Adaptive	Uniform and random	[31]	Concentric Ring
[56]	2009	Adaptive	.....	N/A	N/A
[57]	2010	Adaptive	Uniform and random	No Specific Energy Model	Square
[58]	2010	Adaptive	Spatial Poisson	[58]	Square
[60]	2010	Adaptive	Uniform and random	N/A	Square
[61]	2010	Adaptive	Spatial Poisson	[35]	Square
[62]	2011	Adaptive	Uniform and random	[31]	Circular and square
[63]	2011	Adaptive	Spatial Poisson	[74]	Square
[64]	2012	Adaptive	Uniform and random	[31]	Circular
[65]	2012	Adaptive	Uniform and random	[31]	Square
[66]	2013	Adaptive	Spatial Poisson	[31]	Circular
[67]	2014	Adaptive and static	Uniform and random	[31]	Circular

**Table 7: Comparative Analysis of algorithms based on their strengths, weaknesses and applications**

Refer.	Strengths	Weaknesses	Applications
[31]	It reduces data collisions. It is more robust. The random mechanism of cluster head selection has many advantages such as easy realization, easy operation, and better scalability.	The Energy model used is too ideal to some extent, especially when the scale of sensing field is large. Overhead generated is very high.	This algorithm provides the high performance needed under the tight constraints of the wireless channel.
[32]	This algorithm outperform LEACH-A in a large class of situations and applications and especially when the final receiver is closer to the sensors and the deterministic attenuation due to the path-loss has a value of <i>no</i> larger than 2.5.	Aggregation of data is not used in this algorithm	It can used for the monitoring of a big car-park in which sensors, distributed in the area, interact to communicate with an external receiver (mounted over a car) the best way to reach the closest free place.
[34]	It's required less hardware and software complexity	Less robust	It is applicable to the overall design problem through a data aggregation model
[35]	SEP yields longer stability region for higher values of extra energy brought by more powerful nodes.	Due to practical/cost constraints, it is not always possible to satisfy the constraints for optimal distribution between different types of nodes.	One of the applications could be the re-energization of sensor networks. There are also applications where the spatial density of sensors is a constraint.
[36]	It reduces the overhead generated	This algorithm has ignored lossy compression	This algorithm concentrates on applications which involve continuous data gathering for large scale and distributed physical phenomena using a dense WSNs.
[37]	Its minimizing the energy consumption of the system	This method cannot be applied in real time based model	This can be applied to the fact that the optimal number of clusters-heads determines as the different density of sensors in a bounded area
[38]	Energy consumption of the network decreases as the number of levels is increased	Increased overhead and delay relative to single level clustering	It is expected to be most suited for delay insensitive applications where data aggregation is required
[39]	It provides the efficient utilization of data correlation.	Type of Routing used, allows for maximum possible aggregation at each hop, which may not always be practically possible to implement.	It is applicable for the place where energy consumption required with respect to data correlation of the nodes
[40]	Provides Energy conservation in the networks	It is focusing on the optimal cluster size, is based on the assumption that each issue cluster area spans over one, or more, square shaped clusters, which may not always be true in real world WSN environments.	We can take a decision whether clustering will be useful or not.
[41]	Intra-cluster communication can be completely avoided by this algorithm which is not true in the case of LEACH protocol.	It does not include the development of cost-effective and rapid re-clustering mechanisms according to the environmental changes.	It is highly effective in minimizing in-network data-reporting traffic and, accordingly, in reducing the energy usage of individual sensor nodes
[42]	It reduces the overhead of the system	Complexity of the algorithms is high	It can be used for dynamic network

[43]	Network lifetime is enhanced	System complexity is increased	It is more acceptable for monitoring and measuring larger scale of water environment than before [44-45].
[46]	This clustering protocol exhibits fault-tolerance in the case of node failures compared deterministic clustering approaches in the literature.	Identification of the optimal number of ants is a little bit tough process.	Due to the robustness of any biologically-inspired algorithm, this protocol could handle unexpected circumstances in the environment and node failures. This approach may also be useful in applications that require an in-network actuation, to assist in the sensor-actuator coordination.
[47]	It has extended the lifetime of the network by 10% as compared with LEACH in the presence of the same setting of powerful nodes in a network	This algorithm is valid only when intra-cluster communication phase is long enough	This algorithm can be used for monitoring the application
[48]	Its reduces data collisions It is more robust	Overhead generated is very high	It provides best performance needed under the rigid constraints of the wireless channel.
[49]	Performance is greatly improved by adopting a multi-cluster structure	The multi-hop setup was not explored in this algorithm.	This paper studies a multi-cluster sensor network, which is applied for source extraction in a sensing field
[51]	It has a new, realistic and comprehensive energy model for wireless sensor networks	.....	With the help this algorithm we can take decision whether we should do optimal clustering or not the basis of free space fading energy
[52]	It has better performance than LEACH	This algorithm is not applicable to the real time scenario. It is not applicable for asymmetrical channel	It is applicable for dynamic topology
[53]	This algorithm provides an optimal cluster number that can still effectively function, regardless of the different densities of sensors	The complexity of this algorithm is high	The proposed cross-layer analytical model can facilitate the design of the optimal number of clusters for a sensor network in different radio environments
[54]	It has better performance than LEACH	Overhead generated is very high Fault tolerance is also a major issue	This algorithm can equally apply to small sized wireless networks.
[55]	Network lifetime increase.	It's not providing good connectivity and coverage	This algorithm has application where we location based application
[56]	It is better than or comparable with CCIA and Kd-tree [72,73] in terms of pattern recognition rate.	This algorithm is computationally expensive in some cases	It can apply even though the structure of the dataset is complex
[57]	It shows that 1-hop clustering performs best for a large spectrum of different network sizes, node densities and the number of base stations.	.....	It can be used for real transmission
[58]	Less overhead	Algorithms only considered error free communication and do not account for the channel contention	It is applicable for determining the optimal number of clusters based on other MAC and routing algorithms.
[60]	More Stable region.	.....	It can be used for approximation problems and recognition of geometrical shapes in image processing
[61]	Provide Scalable Network	This algorithm does not use a real scenario energy model	Applicable for different network environment and parameters

[62]	Enhanced Network Lifetime	Energy wastage due to variability of the number of clusters	This is applicable for providing the analytical model for all types of sensing fields and the position of the base station in the scenario
[63]	This algorithm can prolong the network lifetime efficiently compared with LEACH protocol	It cannot ensure uniform distributed of cluster heads in the network for it does not consider the location of nodes.	It is useful for homogeneous type of applications
[64]	It provide on an average a 28 % reduction in total energy usage over other existing algorithm	It can not apply for location aware applications	This scheme would be very suitable in applications where either the required accuracy in data is low or the data has very high redundancy
[65]	Its provide the perfect selection of optimal number of clusters	This algorithm has less fault tolerance	The application of this technique in WSNs is possible due to the redundancy in sensor readings and the large number of nodes deployed
[66]	Energy Efficient	Complexity of the system increases	It can be applied to practical scenario like a battlefield where we are placing nodes in an uncontrolled manner
[67]	High Network lifetime	.....	It is applicable for non-uniform distributed nodes in WSNs

## 5 Open Issues and Challenges

The open issues and challenges in WSNs, which can serve as research topics for future work are summarized below

- ✓ Finding optimal number of clusters with moving nodes that provides good coverage and energy-efficient system.
- ✓ Finding optimal number of clusters for non-uniform node distribution (Gaussian distribution) in application for remote places, where sensor nodes are dropped by helicopter.
- ✓ How to select optimal number of clusters without knowing the location of the nodes?
- ✓ Optimal number of clusters using a cross layer approach along with random traffic and data aggregation model.
- ✓ Finding optimal number of clusters using more practical energy model other than an existing one [31]
- ✓ Calculating optimal number of CHs by using probabilistic sensing model of the nodes (Elfes Sensing model)

## 6 Conclusion

WSNs are an emerging technology that has been attracting large pool of researchers in recent years. Optimal number of clusters plays a crucial role in the performance of WSNs, in terms of system scalability, energy efficiency, collision reduction, network lifetime, latency, and efficient routing backbone in the network. We have surveyed the state-of-art of different algorithms for finding the optimal number of clusters that have been analytically reported in the literature of WSNs, and have presented the methodologies utilized by different authors to calculate the optimal number of clusters in WSNs from an energy-efficiency point of view with the help of taxonomy. This paper discusses the fundamental concepts of clustering, need of clustering, optimal clustering and need of optimal clustering in WSNs. We have compared the optimal clustering algorithms based on cluster variability, heterogeneity, analytical formula for optimal number of clusters, type of energy model used, type of node distribution, type of sensing field, the position of the base station, strengths, weaknesses and

applications of each and every algorithm for calculating an optimal number of clusters. We have also discussed the impact of different levels at which WSNs is modeled viz. Radio Energy Model Level, Network Model Level and Clustering Level for selecting optimal number of clusters along with open issues and challenges in WSNs.

## REFERENCES

- [1]. Pottie, G.; Kaiser, W. Wireless Integrated Network Sensors. *ACM Communications*, 2000, 43, 5,p. 51–58.
- [2]. Kushwaha, S.; Kumar, V.; Jain, S. Node Architectures and Its Deployment in Wireless Sensor Networks: A Survey. In *High Performance Architecture and Grid Computing*, published Springer Berlin Heidelberg, 2011, p. 515-526.
- [3]. Yu, J. Y.; Chong, P. H. J. A Survey of Clustering Schemes for mobile ad hoc networks. *IEEE Communications Surveys Tutorials*, 2005, 7,p. 32–48.
- [4]. Kumar, V.; Jain, S.; Tiwari, S. Energy Efficient Clustering Algorithms in Wireless Sensor Networks: A survey. *IJCSI International Journal of Computer Science Issues*, 2011, 8(5),p. 1694-0814.
- [5]. Kumar, V.; Tiwari, S. Energy Efficient Mechanisms in Wireless Sensor Networks: A survey. *International Journal of Advanced Research in Computer Science*, 2011, 2(5),p. 595-604.
- [6]. Abbasi, A. A.; Younis, M. A Survey on Clustering Algorithms for Wireless Sensor Networks. *Elsevier Science direct Computer Communications*, 2007, 30,p. 2826–2841.
- [7]. Deosarkar, B. P.; Yadav, N. S.; Yadav, R. Cluster head selection in clustering algorithms for wireless sensor networks: A Survey. *Proceedings of International Conference on Computing, Communication and Networking*, 2008, p. 1–8.
- [8]. Akyildiz, I. F.; Vuran, M. C. Wireless Sensor Networks. *A John Wiley and Sons, Ltd, Publication*, 2010.
- [9]. Zhang, Y.; Yang, L.T.; Chen J. RFID and Sensor networks: Architectures, Protocols, Security, and Integrations. *CRC press Taylor and Francis* 2010.
- [10]. Heinzelman, W.R.; Chandrakasan, A.; Balakrishnan, H. Energy-efficient communication protocol for wireless microsensor networks. *Proceeding 33<sup>rd</sup> Hawaii International Conference on System Sciences*, 2000, p. 3005–3014.
- [11]. Dali, W.; Chan, H. A survey on cluster schemes in ad hoc wireless networks. *Proceedings of 2<sup>nd</sup> International Conference on Mobile Technology, application and systems*, 2005, p.1-8.
- [12]. Kozat, U.C.; Kondylis, G.; Ryu, B.; Marina, M. Virtual dynamic backbone for mobile ad hoc networks. *IEEE International Conference on Communication*, 2002, p.250-255.
- [13]. Chen, B.; Jamieson, K.; Balakrishnan, H.; Morris, R. SPAN: An energy-efficient coordination algorithm for topology maintenance in ad hoc wireless networks. In *Wireless Networks*, 2002, 8,p. 481–494.



- [14]. Krunz, M.; Siam, M. Z.; Nguyen, D. N. Clustering and power management for virtual MIMO communications in wireless sensor networks. *Journal of Ad-hoc network*, 2013, 11(5), p. 1571–1587.
- [15]. Naeimi, S.; Ghafghazi, H.; Chow, C.; Ishii, H. Survey on the Taxonomy for Cluster-based Routing Protocols for Homogeneous Wireless Sensor Networks. *Sensors. Journal of Sensor, MDPI*, 2012, 12(6),p. 7350–7409.
- [16]. Liu X. A Survey on Clustering Routing Protocols in Wireless Sensor Networks. *Journal of Sensors*, 2012, 12,p. 11113-11153.
- [17]. Li, J.; Mohapatra, P. Analytical modeling and mitigation techniques for the energy hole problem in sensor networks. *Pervasive Mobile Computation*, 2007, 3, p.233–254.
- [18]. Tran-Quang, V.; Miyoshi, T. A Transmission Range Adjustment Algorithm to Avoid Energy Holes in Wireless Sensor Networks. *Proceedings of 8<sup>th</sup> Asia-Pacific Symposium on Information and Telecommunication Technologies, Kuching, Malaysia*, 2010, p. 15–18.
- [19]. Ishmanov, F.; Malik, A.S.; Kim, A.S. Energy consumption balancing (ECB) issues and mechanisms in Wireless Sensor Networks (WSNs): A comprehensive overview. *Eurosip Transaction on Telecommunication*, 2011,22, p.151–167.
- [20]. Karl, H.; Willig, A. Protocols and architectures for wireless sensor network. *A John Wiley and Sons, Ltd, Publication*, 2005.
- [21]. Wang, P.; Dui, R.; Akyildiz, I. Collaborative data compression using clustered source coding for wireless multimedia sensor networks. *Proceeding of IEEE Conference on Computer Communications (INFOCOM), San Diego, CA, USA*, 2010, p. 1713–1723.
- [22]. Tripathi, R. K.; Singh, Y.N.; Verma, N.K. Two-tiered wireless sensor networks-base station optimal positioning case study. *IET Wireless Sensor Systems*, 2012, 2(4),p. 351–360.
- [23]. Mille, M. J.; Vaidya, N. H. A MAC protocol to reduce sensor network energy consumption using a wakeup radio. *IEEE Transaction on Mobile Computing*, 2005, 4(3), p.228–242.
- [24]. Zhu, J.; Papavassiliou, S. On the energy-efficient organization and the lifetime of multi-hop sensor networks,” *IEEE Communication Letter*. 2003, 7(110), p. 537–539.
- [25]. Medagliani, P.; Martalò, M.; Ferrari, G. Clustered Zigbee networks with data fusion: Characterization and performance analysis. *Journal of Ad Hoc Networks*, 2011, 9(7),p. 1083-1103.
- [26]. Elfes, A. Occupancy grids: a stochastic spatial representation the active robot perception. *Autonomous Mobile Robots: Perception, Mapping and Navigation*, 1991, p. 60-70.
- [27]. Hossain, A.; Chakrabarti, S.; Biswas, P.K. Impact of Sensing Model on Wireless Sensor Network Coverage. *IET wireless sensor system*, 2012, 2(3),p. 272-28.
- [28]. Gu,Y.; Wu, Q.; Rao, N. S. V. Optimizing Cluster Heads for Energy Efficiency in Large-Scale Heterogeneous Wireless Sensor Networks. *International Journal of Distributed Sensor Networks*, 2010.

- [29]. Dabirmoghaddam, A.; Ghaderi, M.; Williamson, C. Cluster-based correlated data gathering in wireless sensor network. *Proceeding of IEEE International Symposium on Modeling, Analysis and Simulation of Computer and Telecommunication Systems , Miami Beach, USA, 2010, p. 163–171.*
- [30]. Tuah,N.; Ismail, M.; Jumari, K. Energy efficient algorithm for heterogeneous wireless sensor network. *IEEE International Conference on Control System, Computing and Engineering, 2011, p. 92-96.*
- [31]. Heinzelman, W. R.; Chandrakasan, A.; Balakrishnan, H. an application-specific protocol architecture for wireless microsensor networks. *IEEE Transaction on Wireless Communication, 2002, 1(4),p. 660– 670.*
- [32]. Depedri, A.; Zanella, R.; Verdone, R. An energy efficient protocol for wireless sensor networks. *Autonomous Intelligent Networks and System, Menlo Park, 2003.*
- [33]. Chen, P.; Dea, B. O.; Callaway, E. Energy Efficient System Design with Optimum Transmission Range for Wireless Ad Hoc Networks. *IEEE International Conference on Communication, 2002, 2,p.945-952.*
- [34]. Mhatre, V; Rosenberg, C. Design guidelines for wireless sensor networks: communication, clustering and aggregation. *Journal of Ad Hoc Network Journal, 2004, 2(1),p. 45–63.*
- [35]. Smaragdakis, G.; Matta, I.; Bestavros, A. SEP: A stable election protocol for clustered heterogeneous wireless sensor networks. *Proceeding of the International Workshop on SANPA, 2004.*
- [36]. Patten, S.; Krishnamachari, B.; Govindan, R. The Impact of Spatial Correlation on Routing with Compression in Wireless Sensor Networks. *ACM/IEEE IPSN, Berkeley, US, 2004, p.28-35.*
- [37]. Kim, H.; Kim, S.W.; Lee, S.B.; Son, B. Estimation of the optimal number of cluster-heads in sensor network. *Proceeding of KES, Melbourne, Australia, 2005, 3,p. 87–94.*
- [38]. Comeau, F.; Sivakumar, S.C.; Robertson, W.; Phillips W.J. Energy conserving architectures and algorithms for wireless sensor networks. *39<sup>th</sup> Hawaii International Conference on System Sciences, 2006, p.236c.*
- [39]. Chen, H.; Megerian, S. Cluster Sizing and Head Selection for Efficient Data Aggregation and Routing in Sensor Networks. *Proceeding of IEEE WCNC, 2006, p. 2318–2323.*
- [40]. Vljajic, N.; Xia, D. Wireless sensor networks: To cluster or not to cluster. *Proceedings of International Symposium on a World of Wireless, Mobile and Multimedia, 2006, 9, p. – 268.*
- [41]. Xia, D.; Vlagie, X. Near optimal node clustering in wireless sensor networks for environment monitoring. *21<sup>st</sup> international conference and applications, 2007,p. 632-641.*
- [42]. Yang, H.; Sikdar, B. Optimal Cluster Head Selection in the LEACH Architecture. *Proceeding of IPCCC, 2007, p. 93-100.*
- [43]. Wang, H.; Yu, X.; Kong, D.; Yan, X.; Ma, X. Route Protocol of Wireless Sensor Networks Based on Dynamic Setting Cluster. *Proceedings of the International Conference on Information Acquisition, Jeju, Korea, 2007, p. 112–117.*

- [44]. Yu, X.; Xu, L.; Wang, H. A Protocol Design of Water in a Remote Measure Based On Mobile Communication Networks and Wireless Sensor Networks. *Automation in Water Resources and Hydrology*, 2006, 1, p. 6-10.
- [45]. Wang, Y.; Ma, X.; Xu, L. The designing model of effectively, wireless sensor network data link layer based on information fusion strategy. *Computer Engineering*, 2005, 31 23, p. 6-10.
- [46]. Selvakennedy, S.; Sinnappan, S.; Shang, Y. A biologically-inspired clustering protocol for wireless sensor networks. *Journal of Computer Communications*, 2007, 30, 14–15, p. 2786–2801.
- [47]. Kumar, D.; Aseri, T.C.; Patel, R.B. EEHC: Energy efficient heterogeneous clustered scheme for wireless sensor networks. *Journal of Computer Communications*, 2009, 32, 4, p.662-667.
- [48]. Chan, T.J.; Chen, M.C.; Huang, Y.F.; Lin, J.Y.; Chen, T.R. Optimal Cluster Number Selection in Ad-hoc Wireless Sensor Networks. *WSEAS Transaction on Communication*, 2008, 7, 8, 837-846.
- [49]. Chen, H.; Tse, C.K.; Feng, J. Minimizing effective energy consumption in multi-cluster sensor networks for source extraction. *IEEE Transactions on Wireless Communications*, 2009, 8(3), p. 1480–1489.
- [50]. Raghuvanshi, S.; Tiwari, S.; Tripathi, R.; Kishor, N. G K Clustering Approach to Determine Optimal Number of Clusters for Wireless Sensor Networks. *Fifth IEEE Conference on Wireless Communication and Sensor network*, 2009, p.1-5.
- [51]. Halgamuge, M. N.; Zukerman, M.; Ramamohanarao, K; Vu, H. L. An estimation of sensor energy consumption. *Progress In Electromagnetics Research B*, 2009, p.12259-295.
- [52]. Li, H.; Shunjie, X.; Shurong, L.; Weixia, Z.; Zheng, Z. Novel Method for Optimal Number of Cluster Heads in LEACH. *WASE International Conference on Information Engineering*, 2009, p. 302-309.
- [53]. Wang, L. C.; Wang, C. W.; Liu, C. M. Optimal number of Clusters in Dense Wireless Sensor Networks: A Cross-Layer Approach. *IEEE Transactions on Vehicular Technology*, 2009, 58(2), p.966-976.
- [54]. Kumar, D.; Aseri, T. S.; Patel, R. B. EECH: Energy efficient cluster head election protocol for heterogeneous Wireless Sensor Networks. *Proceedings of ACM International Conference on Computing, Communication and Control, Bandra, Mumbai, India*, 2009, p. 75-80.
- [55]. Li, H.; Shunjie, X.; Guoqiang, W.; Zhe, J. Uneven Virtual Grid-Based Clustering Routing Protocol for Wireless Sensor Networks. *Proceedings of IEEE International Conference on Information and Automation, Zhuhai/Macau, China*, 2009, p.397-402.
- [56]. Wang, Y.; Li, C.; Zuo, Y. A selection model for optimal fuzzy clustering algorithm and number of clusters based on competitive comprehensive fuzzy evaluation. *IEEE Transaction on Fuzzy System*, 2009, 17(3),568–577.
- [57]. Förster; Förster, A.; Murphy, A. L. 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*, 2010, 28, p. 49–63.
- [58]. Li, W.; Martins, P.; Shen, L. Determination method of optimal number of clusters for clustered wireless sensor networks. *Journal Wireless Communications and Mobile Computing*, 2012, 12(2), p. 158 – 168.

- [59]. Ahmed, A.A.; Shi, H.; Shang, Y. A survey on network protocols for wireless sensor networks. *International Conference on Information Technology: Research and Education, 2003*, p. 301-305.
- [60]. Raghuvanshi, A. S.; Tiwari, S.; Tripathi, R.; Kishore, N. Optimal number of clusters in wireless sensor networks: a FCM approach. *International Journal of Sensor Networks, 2012, 12(1),p.16-24*.
- [61]. Yang, K.; Wu, Y.; Zhou, H. Research of optimal energy consumption model in wireless sensor network. *2<sup>nd</sup> International Conference on Computer Engineering and Technology Chengdu, China, 2010*, p. V7-421 - V7-424.
- [62]. Navid, A.; Alireza, V.; Wenyao, X.; Mario, G.; Majid, S. Cluster size optimization in sensor networks with decentralized cluster-based protocols. *Journal of Computer Communications, 2012, 35(2),p. 207–220*.
- [63]. Chen, B.; Zhang, Y.; Li, Y.; Hao, X.; Fang, Y. A Clustering Algorithm of Cluster-head Optimization for Wireless Sensor Networks Based on Energy. *Journal of Information & Computational Science, 2011,8(11),p. 2129–2136*.
- [64]. Tandon, R. Determination of Optimal Number of Clusters in Wireless Sensor Networks. *International Journal of Computer Networks & Communications (IJCNC), 2012, 4(4),p. 235-249*.
- [65]. Tuah, N.; Ismail, M.; Jumari, K. Evaluation of Optimal Cluster Size in Heterogeneous Energy Wireless Sensor Network. *1<sup>st</sup> IEEE International Symposium on Telecommunication Technologies, 2012*, p.124 – 130.
- [66]. Tripathi, R.K.; Singh, Y.N.; Verma, N.K. Clustering algorithms for non-uniformly distributed nodes in WSNs. *Electronics Letters, 2013, 49, 4*.
- [67]. Dabirmoghaddam, A.; Ghaderi M.; Williamson, C. On the optimal randomized clustering in distributed sensor networks. *Journal of computer network, 2014, 59, 11,p.17–32*.
- [68]. Katiyar, V.; Chand, N.; Soni, S. A Survey on clustering algorithms for Heterogeneous Wireless Sensor Networks. *International Journal of Advanced Networking and Applications, 2011, 2(4),p. 745-754*.
- [69]. Sheikhpour, R.; Jabbehdari, S.; Khadem-Zadeh, A. Comparison of Energy Efficient Clustering Protocols in Heterogeneous Wireless Sensor Networks. *International Journal of Advanced Science and Technology, 2011, Vol. 36, p.27-40*.
- [70]. Mhatre, V.; Rosenberg, C. Homogeneous Vs Heterogeneous Clustered Networks: A Comparative Study . *Proceedings of IEEE ICC, 2004*, p.1-6.
- [71]. Shih, E.; Cho, S.; Lee, F. S. B.; Calhoun, H.; Chandrakasan, A. Design Considerations for Energy-Efficient Radios in Wireless Micro-sensor Networks. *Journal of VLSI Signal Process, 2004, 37(1),p. 77–94*.
- [72]. Khan, S.; Ahmad, A. Cluster centre initialization algorithm for Kmeans clustering. *Pattern Recognition Letter, 2004, 25(1),p. 1293–1302*.
- [73]. Redmond, S.; Heneghan, C. A method for initializing the K-means clustering algorithm using kd-trees. *Pattern Recognit. Lett., 2007, 28(1), p.965–973*

- [74]. Huang, H.; Yao, D.; Shen, J.; Ma, K.; Liu, H. Multi-weight based clustering algorithm for wireless sensor networks. *Journal of Electronics & Information Technology*, 2008,p. 1489-1492.
- [75]. Stein, E. M.; Shakarchi, R. *Real Analysis: Measure Theory, Integration, and Hilbert Spaces*. NJ: Princeton Univ. Press, 2005.
- [76]. Kumar, V.; Dhok, S.B.; Tripathi,R.; Tiwari, S. "Cluster Size Optimization in Gaussian Distributed Wireless Sensor Networks" *International Journal of Engineering and Technology (IJET)*,2014, 6 (3),p. 1581-1592.
- [77]. Kumar, V.; and Tiwari, S. "Routing in IPv6 over low-power wireless personal networks (6LowPAN): A survey" *Journal of computer Networks and Communication*, 2012.
- [78]. Kumar, V.; Raghuvansi, A.S.; and Tiwari, S. "Performance study of beacon- enabled IEEE 802.15.4 standards in WSNs with Clustering" *International conference on power control and embedded system*,2010, p. 1-5.
- [79]. Kumar, V.; and Tiwari, S. "Performance of Routing Protocols for Beacon-Enabled IEEE 802.15.4 WSNs with Different Duty Cycle", *International Conference on Devices and Communications*, 2011, p.1-5.
- [80]. Kumar, V.; Raghuvansi, A.S.; and Tiwari, S. "LEACH and Its Derivatives in WSN: A Survey" *International Conference on Communication and computational intelligence*, 2010/11, p.81-86.
- [81]. Kumar, V.; Jain, S.; and Tiwari, S. "Impact of Node Density and Mobility on Scalable Routing Protocols in Mobile AdHoc Networks" *Special Issue of International Journal of Computer Applications (0975 8887) on Communication Security*, No.5 Mar.2012
- [82]. Kumar, V.; Jain, S.; and Tiwari, S. " Performance of Routing Protocols in MANETs with Node Density and Mobility using Omni and Directional Antenna" *Special Issue of International Journal of Computer Applications (0975 -8887) on Wireless Communication and Mobile Networks*, No.11., [www.ijcaonline.org](http://www.ijcaonline.org), Jan.2012,p.51-55.
- [83]. Kumar, V.; Dhok, S.B.; Tripathi,R.; Tiwari, S. "A Review Study of Hierarchical Clustering Algorithms for Wireless Sensor Networks" *IJCSI International Journal of Computer Science Issues*,2014, 11(3), p.92-101.