An Improved Hierarchical Energy Efficient Clustering Algorithm for Wireless Sensor Networks

FAGBOHUNMI, Griffin Siji[1], Uchegbu Chinenye E.[2]
[1]*Department of Computer Engineering Abia State University, Uturu, Abia State, Nigeria*
[2]*Department of Electrical and Electronics Engineering, Abia State University, Uturu, Abia State, Nigeria*

***Abstract*: A wireless sensor networks consists of spatially distributed nodes randomly deployed in a given environment. It consists of a sensor, for sensing environmental phenomenon, a transceiver for communication purposes, an energy constrained power source and a small amount of memory used in storing programs that controls the operation of the nodes. Sensors in WSN are normally used to gather information from the environment, these data are then aggregated and sent to a base station. Due to the limited energy and memory capacity of the nodes it becomes imperative that the protocol needed to aggregate this data must be light weight so as to increase the network lifetime. It has been observed that clustering which can be defined as dividing the nodes in a network into groups can help to minimize the energy consumption of the WSN. The idea behind a clustering algorithm is that sensors in a particular group sends their data to a cluster head while the cluster heads send the aggregated data in the group to the base station instead of individual sensors sending their individual data directly to the sink. This paper proposes an improved, distributed, randomized clustering algorithm that organizes the sensors in a network into hierarchical groups. The algorithm is an improvement over the hierarchical clustering algorithm employed currently by researchers in literature. The algorithm proved to minimize the energy consumption of the network especially in a networks of thousands of nodes. We used results in stochastic geometry to derive solutions for the values of parameters of our algorithm that minimize the total energy spent in the network when all sensors report data through the cluster heads to the sink. Simulation experiments show an improvement of at least 15% over the hierarchical clustering algorithms used in comparison in reducing total spent in the network thereby prolonging the lifetime of the network.**

***Keywords*: Sensor Networks; Clustering Methods; Voronoi Tessellations; Algorithms. Aggregate, cluster head**

I. INTRODUCTION

Due to advances witnessed in the manufacture of micro electro-mechanical systems, its innovation has led to the development of extremely small, low-cost sensors that possess sensing, signal processing and wireless communication capabilities. The cost of these sensors are much lower than traditional wired sensor systems. Two researches that has taken advantage of this low cost sensors are, the Smart Dust Project at University of California, Berkeley [13, 14, 15] and WINS Project at UCLA [1, 16], The size of the sensors are approximately 1 cubic millimetre. It is possible to build an ad-hoc wireless network consisting of large numbers of such inexpensive but less reliable and accurate sensors to be used in a wide variety of commercial and military applications. These applications include target tracking, security, environmental monitoring, system control, etc. The sensors are equipped with small batteries with capacity at most 1 Joule [11] so as to keep the cost and size of these sensors small. This development limits both the transmission range and the data rate of the sensors. The implication here is that a sensor can therefore communicate directly only with other sensors that are within its transmission range. In order to enable communication between sensors not within each other's communication range, the sensors form a multi-hop communication network. Sensors in these multi-hop networks detect events from their environment and then communicate the collected information to a central location where parameters characterizing these events are estimated. The cost of information computation is less than that of transmitting a bit [1], therefore it is advantageous to organize the sensors into clusters. In the clustered environment, the data gathered by the sensors are communicated to the sink through a hierarchy of cluster heads. The sink is responsible for computing the final estimates of the parameters in question using the information communicated by the cluster heads. The sink can be a specialized device or just one of these sensors itself. In this scenario, a group of sensors communicate their information to a cluster head whose distance is much smaller to the sink. The energy spent in the network will be much lower than the energy spent when every sensor communicate directly to the information processing centre (sink). Researchers has proposed many clustering algorithms in various contexts [2-6, 22-27]. These algorithms are mostly heuristic in nature in that they aim at minimizing the number of nodes in each cluster so that any node in any cluster is at most d hops away from the cluster head. Most of these algorithms have a time complexity of O(n) , where n is the total number of nodes. Many of them also demand time synchronization among the nodes, which makes them suitable only for networks with a small number of sensors. The Max-Min d-Cluster Algorithm [5] generates d-hop clusters with a run-time of O(d) rounds. However this algorithm does not guarantee that the energy used in communicating information to the information centre is minimized. The clustering algorithm proposed by researchers in [6] aim at maximizing the network lifetime, but it assumes that each node is aware of the whole network topology, which is usually impossible for wireless sensor networks which have a large number of nodes. Many of these clustering algorithms [22, 25, 26, 27] are specifically designed with an objective of generating stable clusters in environments with mobile nodes. But in a typical wireless sensor network, the location of sensors are fixed and the instability of clusters due to mobility of sensors is not an issue. In a wireless sensor networks consisting of a large number of energy-constrained sensors typically thousands of sensors, it becomes imperative to design a light weight algorithm that organize sensors in clusters to minimize the energy used to communicate information from all nodes to the sink. This paper proposes a light weight randomized, distributed algorithm for organizing the sensors in a wireless sensor network in a hierarchy of clusters with an objective of minimizing the energy spent in communicating the information to the sink. Results in stochastic geometry was used to derive values of parameters for the algorithm that minimize the energy spent in the network of sensors.

The remainder of the paper is organized as follows: related work is reviewed in Section 2,; Section 3 describes the stochastic analytical geometry and the algorithm for the hierarchical clustering; Section 4 shows the simulation details, result and analysis while in section 5, the conclusion and an overview of future work were provided.

II. RELATED WORK

There are varieties of design goals in wireless sensors networks, this includes but not limited to, (i) design of low-power signal processing architectures, (ii) low-power sensing interfaces, (iii) energy efficient wireless media access control and routing protocols [3, 19], (iv) low-power security protocols and key management architectures [28, 29], (v) localization systems [20, 21], etc. According to Gupta and Kumar [18], they analysed the capacity of wireless ad hoc networks and thereafter derived the critical power at which a node in a wireless ad hoc network should communicate to form a connected network with probability of one [19]. Many clustering algorithms in various contexts have also been proposed in the past [2-6, 22-27], but to the best of our knowledge, none of these algorithms have been able to minimize the energy spent in the system especially in networks consisting of tens of thousands of nodes. Most of these algorithms are heuristic in nature and their aim is to divide the sensors in the network into groups such that each is at most d hops away from the cluster head. In the scenario presented in this paper, generating a cluster of minimal nodes cannot guarantee minimum energy usage. In the Linked Cluster Algorithm proposed by the researchers in [2], a node with the highest identity among all nodes within one hop of itself or among all nodes within one hop of one of its neighbours becomes the cluster head.

The LCA2 was an improvement on the Linked Cluster algorithm, [8] in the LCA2 algorithm, the node with the lowest id among all nodes that are neither a cluster head nor are within 1-hop of the already chosen cluster heads is chosen as the cluster head. In the algorithm proposed by researchers in [9], a cluster head is chosen as the node with highest degree among its 1–hop neighbours. In [4], the authors proposed a distributed algorithm that is similar to the LCA2 algorithm. In [27], two load balancing heuristics for mobile ad hoc networks were proposed by the authors. When the first heuristic is applied to a node-id based clustering algorithm like LCA or LCA2, it leads to longer, low-variance cluster head duration. The other heuristic is for degree-based clustering algorithms. Degree-based algorithms, in conjunction with the proposed load balancing heuristic produce longer cluster head duration. The parameters for electing a cluster head in the Weighted Clustering Algorithm (WCA) includes: (i) the number of neighbours, (ii) transmission power, (iii) battery life and (iv) mobility rate of the node [26]. The algorithm also restricts the number of nodes in a cluster so that the performance of the MAC protocol was not degraded. In the Distributed Clustering Algorithm (DCA) [24], it uses weights associated with nodes to elect cluster heads. These weights are generic and can be defined based on the application. The node with the highest weight among its 1-hop neighbours is elected as the cluster head. It should be noted that the DCA algorithm is suitable for networks in which nodes are static or moving at a very low speed. The Distributed and Mobility-Adaptive Clustering Algorithm (DMAC) is a modification to the DCA algorithm which incorporate high mobility sensors during or after the cluster set-up phase [25]. The algorithm described in the paper generates 1-hop clusters and require synchronized clocks and have a complexity of O(n) . This makes them suitable only for networks with a small number of nodes. The time complexity of the Max-Min d-cluster algorithm proposed by researchers in [5] is o(d) rounds. This algorithm generates fewer clusters [5], does not require clock synchronization and achieves better load balancing among the cluster heads, than the LCA and LCA2 algorithms. The researchers in [6] proposed a clustering algorithm that aims at maximizing the lifetime of the network by determining optimal cluster size and optimal assignment of nodes to cluster heads. Their proposal assumes a priori knowledge of the number and location of the cluster heads, this is however not possible in all scenarios. The algorithm also requires that each node be aware of the complete topology of the network, which is generally not possible in the context of large sensor networks. McDonald et al [22] proposed a distributed clustering algorithm for mobile ad hoc networks that ensures that the probability of mutual reachability between any two nodes in a cluster is bounded over time. The researchers in [10], designed a 2-level hierarchical telecommunication network in which the nodes at each level are distributed according to two independent homogeneous Poisson point processes and the nodes of one level are connected to the closest node of the next higher level. They later studied the moments and tail of the distributions of characteristics such as the number of lower level nodes connected to a particular higher level node and the total length of segments connecting the lower level nodes to the higher level node in the hierarchy. The results presented in their paper was used to obtain the optimal parameters for the algorithm used in this paper. Baccelli and Zuyev [12] extended the above study to hierarchical telecommunication networks with more than two levels. They designed a network in which a list of subscribers at the lowest level are connected to concentration points at the highest level, directly or indirectly through distribution points. These subscribers, distribution points and the concentrators form the three levels in the hierarchy and are distributed according to independent homogeneous Poisson processes. Now assuming that a node is connected to the closest node of the next higher level, point processes and stochastic geometry was used to determine the average cost of connecting nodes in the network as a function of the intensity of the Poisson processes governing the distribution of nodes at various levels in the network. With this they derived the intensity of the Poisson process of distribution points (as a function of the intensities of the Poisson processes of subscribers and concentration points) that minimizes this cost function. The results obtained from this was then extended for non-purely hierarchical models in order to derive the optimal intensity of Poisson process of distribution points when the numeric values of the intensities of other two processes was given. With this they were able to generalize the cost function for networks with more than three levels. Seema and Edward [26] proposed a hierarchical clustering algorithm (HCA) for micro-sensor networks in which the sensors with certain range of probability elect themselves as cluster heads and broadcast their decisions. The other sensors within the sensor's transmission range join the cluster head that requires minimum communication energy. They employed stochastic geometry to analyse the maximum hop count and maximum number of nodes allowable in a cluster to reduce total energy spent by sensors in the network. They later extended their protocol for hierarchical cluster head formation. Some of the analytical formulations and algorithms used in this paper were similar to those used in the HCA algorithm. HCA led to reduction in total energy expended by sensors in the network, however certain procedure in their algorithm for forming clusters may lead to reducing the network life time. This is the area that this paper intend to address. The shortcoming of the HCA algorithm is that it allows only 1-hop clusters to be formed, which might lead to a large number of clusters. Simulation results were provided showing how the energy spent in the system changes with the number of clusters formed and have observed that, there is a number of clusters that minimizes the energy spent for a given density of nodes. The algorithm is run periodically, and for each period, the probability of becoming a cluster head is chosen to ensure that every node becomes a cluster head at least once within 1/p rounds, where p is the desired percentage of cluster heads. This distribution ensures that none of the sensors are overloaded due to the added responsibility of being a cluster head. In [26], the authors assumed that the sensors are capable of being in different modes: (i) active mode, (ii) idle mode and (iii) sleep mode. This is needed so as to conserve the energy of the sensors. The sensors will only be in active mode when they have data to transmit, the idle or listening mode is activated when they need to sense signals from their environment, while the sleep mode is activated when the sensors are neither transmitting data nor sensing signals. Also when they are in active mode, they have the capability of tuning the power at which they transmit and they communicate with power enough to achieve acceptable signal-to-noise ratio at the receiver. In this paper the sensors are assumed to be simple and homogeneous, i.e. they transmit at a fixed power level. A probabilistic approach was also employed to determine when the sensors transits between the three modes, i.e. idle listening, active and sleep modes. This probabilistic function was based on the history of sensing in the network. Data between two communicating sensors not within each other's radio range is forwarded by other sensors in the network. In the simulation experiments performed by authors, in [26], it was observed that in a network with one level of clustering, there was an optimal number of cluster heads that minimizes the energy used in the network. This paper adopts the results provided in [26] to obtain the optimal number of cluster heads at each level of clustering analytically, this was then extended using our algorithm to generate one or more levels of clustering.

III. A NEW, ENERGY-EFFICIENT, SINGLE-LEVEL CLUSTERING ALGORITHM

A. Algorithm

With probability p, each sensor that are at most distance d from the centre of the cluster becomes a cluster head (CH) (d is less than the transmission radius of the sensor) and advertises itself as a cluster head to the sensors within its radio range. These cluster heads are called the volunteer cluster heads. This advertisement is forwarded to all the sensors that are no more than k hops away from the cluster head. Any sensor that receives such advertisements and is not itself a cluster head joins the cluster of the closest cluster head. Any sensor that is neither a cluster head nor has joined any cluster itself is removed from the network. This is in fact the modification made to the previous researcher's work in [26]. The idea behind this modification is that any mode that is not k hops away from a volunteer cluster head will require a lot of energy to transmit data packets to the processing centre and hence may cause a weak link in the network which might lead to shorter lifetime of the network. This can be inferred if a sensor does not receive a CH advertisement within time duration t (where t units is the time required for data to reach the cluster head from any sensor k hops away) it then means that it is not within k hops of any volunteer cluster head and hence excluded from the network. Moreover since data packets in WSN most often represent redundant data, the exclusion of these sensors will have minimal effect on the efficiency of the network as will be deduced from the simulation experiments in the next section. In the previous researcher's algorithm, any sensors not within k hop of a volunteer cluster head becomes a cluster head; these were referred to as the forced cluster heads. Moreover, since all the sensors within a cluster are at most k hops away from the cluster-head, the cluster head can transmit the aggregated information to the processing centre after every t units of time. This places limits on the number of hops allowed from any sensor to its cluster head which enable the cluster heads to schedule their transmissions therefore helping the sensors

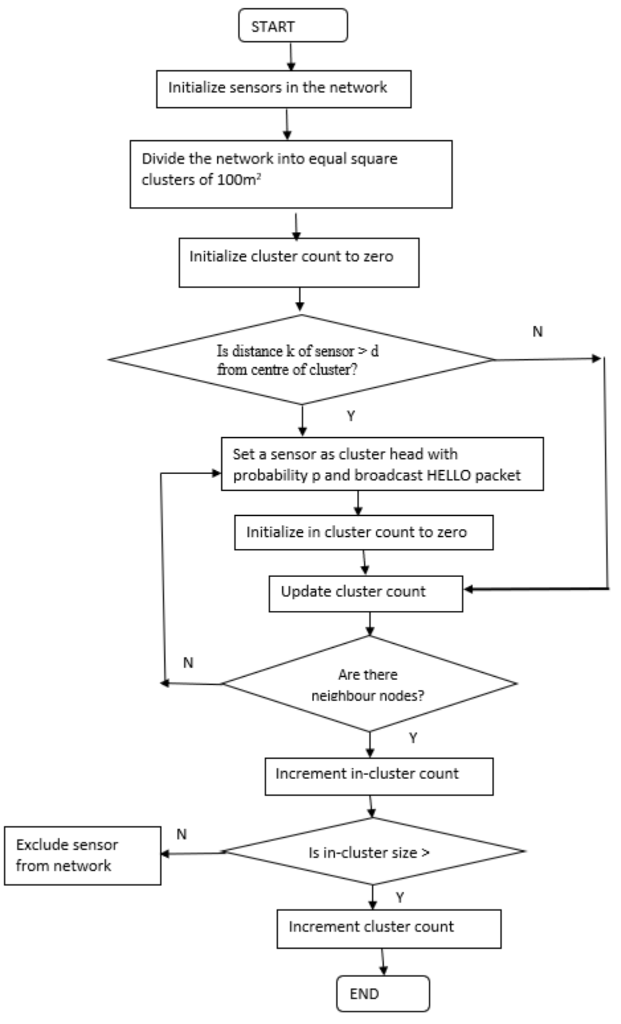

Fig. 1. Flowchart for single level (Hierarchical) cluster formation

B. Formulation of Optimal Parameters for the Algorithm

The following assumptions were made in order to determine the optimal parameters for the algorithm proposed in this paper.

transit between its three modes of sensing as explained earlier. The flowchart for the single level cluster formation is shown in figure 1.

The algorithm described above is a distributed algorithm and does not require clock synchronization between the sensors. It should be noticed that the energy required by the cluster head to transmit the aggregated data from the sensors in their individual clusters to the sink will depend on the parameters p and k of our algorithm. The aim of this paper is to make a WSN easily scalable by organizing its sensors into clusters so as to minimize energy consumption, hence it is imperative to find the values of the parameters p and k in our algorithm that would ensure minimization of energy consumption. The optimal values of p and k are derived in the next subsection.

a) The distribution of the sensors in the WSN follows a spatial Poisson process of intensity α in a 2-dimensional space. The sensors are also homogeneous, i.e. they all have the same amount of energy.
b) When data is exchanged between two communicating sensors outside their communication (radio) range, such data is forwarded by other sensors.
c) If the distance between any sensor and its cluster head is d, this is equivalent to d/r hops.
e) For each sensor, 1 unit of energy is used to transmit or receive 1 unit of data.
e) A routing infrastructure is embedded in the algorithm, hence end to end communication between sensors involves only the sensors on the routing path.
f) Before the clustering algorithm is activated, a security mechanism is performed on the sensors in the network, (this is beyond the scope of this paper) so as to ensure that the communication environment is contention and error-free as such, sensors do not have to retransmit any data.
The need for the derivation of the optimal parameter values is to define a function for the energy used in the network to communicate information to the information-processing centre (sink) and then find the values of parameters (p and k) that would minimize it.

C. *Computation of the Optimal Probability of Becoming a Cluster Head:*

As said earlier, the sensors are distributed according to a homogeneous spatial Poisson process therefore, the number of sensors in a square area of side 2b is a Poisson random variable, N with mean λB, where $B = 4b^2$. If there are m sensors in this area, and the sink is also located at the centre of the square. The probability of becoming a cluster head is p, therefore on average, mp sensors will become cluster heads. Let F be a random variable that denotes the length of the segment from a sensor located at (x_i, y_i), i = 1,2,……m to the processing centre. Without loss of generality, we assume that the processing centre is located at the centre of the square area. Then,

$$E[F_i \mid N = n] = \int_B \sqrt{x_i^2 + y_i^2}\left(\frac{1}{4b^2}\right) dB = 0.765b \qquad (1)$$

From the above explanation, there will be an average of mp cluster heads (CH) and the location of any CH is independent of the locations of other CHs, the total length of the segments from all these CHs to the processing centre is 0.765mpb.

It has already been stated that a sensor becomes a cluster head with probability p, and both the cluster heads and non-cluster heads follows a homogeneous spatial Poisson processes distribution PP_1 and PP_0 of intensity $\lambda_1 = p\lambda$ and $\lambda_0 = (1 - p)\lambda$ respectively. In the mean time it is assumed that there is no limit to the maximum number of hops in the clusters. Each non-cluster head joins the cluster of the closest cluster head to form a Voronoi tessellation [8].

The plane is thus divided into zones called the Voronoi cells, each cell corresponding to a PP_1 process point, called its nucleus. If N_i is the random variable denoting the number of PP_0 process points in each Voronoi cell and L_i is the total length of all segments connecting the PP_0 process points to the nucleus in a Voronoi cell, then according to results in [9],

$$E[N_v \mid N = m] \approx E[N_v] = \frac{\beta_0}{\beta_1} \qquad (2)$$

$$E[L_v \mid N = m] \approx E[L_v] = \frac{\beta_0}{2\beta 1^{3/2}} \qquad (3)$$

If C_1 is defined to be the total energy used by the sensors in a Voronoi cell to communicate one unit of data to the cluster head. Then,

$$E[C_1 \mid N = m] = \frac{E[L_v \| N=m]}{r} \qquad (4)$$

If C_2 is defined to be the total energy spent by all the sensors in communicating 1 unit of data to their respective cluster heads, as said earlier, there are mp cells, therefore the expected value of C_2 conditioned on N, is given by

$$E[C_2 \mid N = m] = mpE[C_1 \mid N = m]. \qquad (5)$$

If the total energy spent by the cluster heads to communicate the aggregated information to the processing centre is denoted by C_3, then,

$$E[C_3 \mid N = m] = \frac{0.7647\,mpb}{r} \qquad (6)$$

If C is defined to be the total energy spent in the system, then,

$$.E[C \mid N = m] = E[C_2 \mid N = m] + E[C_3 \mid N = m]$$

$$= \frac{mp}{r}\frac{(1-p)}{2p^{3/2}\sqrt{\beta}} + \frac{0.7647\,mpb}{r} \qquad (7)$$

Removing the conditioning on N yields:

$$E[C] = E[E[C \mid N = m]]$$

$$= E[N]\left[\frac{1-p}{2r\sqrt{p\beta}} + \frac{0{,}7647\,pb}{r}\right]$$

$$= \beta B\left[\frac{1-p}{2r\sqrt{p\beta}} + \frac{0{,}7647\,pb}{r}\right] \qquad (8)$$

From this equation, E[C] is the minimized value of p which is a solution of

$$cp^{3/2} - p - 1 = 0 \qquad (9)$$

Equation 9 has three roots, two of them are imaginary. The value of the second derivative of the above function is positive for the only real root of equation 9 and therefore will minimize the energy spent. The only real root of equation 9 is given by:

$$P = \left[\frac{1}{3c} + \frac{\sqrt[3]{2}}{3c\left(2+27c^2+3\sqrt{3c}\sqrt{27c^2+4)}^{\frac{1}{3}}\right)} + \frac{2+27c^2+3\sqrt{3c}\sqrt{27c^2+4)^{1/3}}}{3c} * \right.$$
$$132 \qquad (10)$$

D. Computation of the range of hops allowed within a cluster

It should be noticed that in the model described above, a limit or range was not placed on the number of hops (k) allowable within a cluster. However this will be needed for two reasons: (i) to enable a synchronous communication between the cluster members and the cluster head and (ii) to prolong the lifetime of the WSN. This second reason is the improvement proffered over the previous researcher's work because, we argue here that having less than three sensors to make a cluster will result in high energy consumption for communication between such sensors and the sinks, as frequent communication exchange will be done between such sensors and the sinks (processing centre) thereby leading to quick energy drain and lower network lifetime. Allowing isolated nodes to form cluster heads as propounded by the previous researcher will only serve to lower the network lifetime as validated from the simulation results in the next section. As a result of this, the minimum number of sensors allowed in a cluster is three.

Now we will define a formula to find the maximum possible distance (call it R_{max}) at which a PP_0 process point can be from its nucleus in a Voronoi cell, we can find the maximum value of k by assuming that a distance R_{max} from the nucleus is equivalent to R_{max}/r hops. Similarly by setting $k = R_{max}/r$ will also ensure that the number of sensors in the cluster almost will easily converge to a central value. However getting a fixed value of R_{max} is not possible, and therefore one cannot say with certainty that any point of PP_0 process will be at the most R_{max} distance away from its nucleus in the Voronoi

Tessellation. A probabilistic approach will now be adopted. R_{max} will be set to a value such that the probability that any point more than distance R_{max} from a PP_0 process to all points of PP_1 process is very small. Using this value of R_{max}, will ensure that the probability of any sensor being more than k hops away from all cluster heads will be very small.

Assume that the radius the radius of the minimal ball centred at the nucleus of a Voronoi cell is Ωm, which contains the Voronoi cell. Ωp is defined to be the probability that Ωm is greater than a certain value R , i.e. $\Omega_R = P(\Omega_M > R)$ It can be proved that

$\Omega_R \leq 7\exp(-1.09\lambda_1 R^2)$ [10], If R_α is the value of Rsuch Ωp is less than α , then

$$R_\alpha \leq \sqrt{\frac{-0.9172\ln(\alpha/7)}{p_1\beta}} \qquad (11)$$

From equation 11 it can then be shown that the expected number of sensors that will not join any cluster is αn if we set

$$k_1 = \frac{1}{r}\sqrt{\frac{-0.9172\ln(\frac{\alpha}{7})}{p_1\,\beta}} \qquad (12)$$

In order to ensure minimum energy consumption, a very small value for α will be used , this means that the probability that all sensors will be within k hops from at least one cluster head will be very high.

When α = 0.001 and the values of p and k are computed according to equations (10) and (12), Assuming we have a network of 1000 sensors, on an average 1sensor will not join any cluster heads, this again validates the idea or the improvement proposed in this paper that any sensor not within radius R_{max} from any cluster head will be excluded from the network. In the previous researchers work [26] the isolated sensors is made to be a forced cluster head. This is the weakness in the protocol and it leads to a reduction in the network lifetime as will be shown in the simulation experiments in the next section. Also considering the previous researcher's work, the optimal value of p for a network with 1000 nodes in an area of 100 sq. units is 0.008; this means 8 nodes will become forced cluster heads on the average. These isolated sensors was excluded in this work.

IV. SIMULATION EXPERIMENTS AND RESULTS

The algorithm described in section IIIA was simulated using MATLAB for networks with varying sensor density d (i.e. number of sensors per cluster) and with different values of the parameters p and k.

The communication range for each sensor in all simulations was assumed to be 1 unit. The output of the first set of these simulations is shown in Fig 2. with parameters p and k set to between 0.1 and 2 on a network of 500 sensors distributed uniformly in a square area of 100 square units. In order to validate the fact that the optimal values of the parameters p and k used in the algorithms computed according to equations 10 and 12 actually minimize the energy spent in the system, the clustering algorithm was simulated on sensor networks with 500, 1000 and 2000 sensors in a square area of 100 sq. units wherein the sensors were distributed uniformly. Without loss of generality, it is assumed that a cost of 1 unit of energy is required for transmitting 1 unit of data. Also the processing centre (sink) is assumed to be located at the centre of the square area. The first set of simulation experiments used a range of values as the probability (p) of a sensor within a cluster to become a cluster head using the algorithm proposed in Section III. The maximum number of hops (k) in each cluster required for each set of probability values was also computed using equation 12. These values were used for the maximum number of hops allowed in a cluster in the simulations. Fig. 3 shows the results of these simulations. Each of the data point in Fig. 3 corresponds to the average energy consumption over 500 experiments. It can be deduced from Fig. 3. that the energy spent in the network is lowest at the optimal values of the parameter p computed theoretically using equation 10 (This optimal value is referred to as opt p), These values are given in Table I for 500, 1000, 2000 and 3000 sensors in the network.

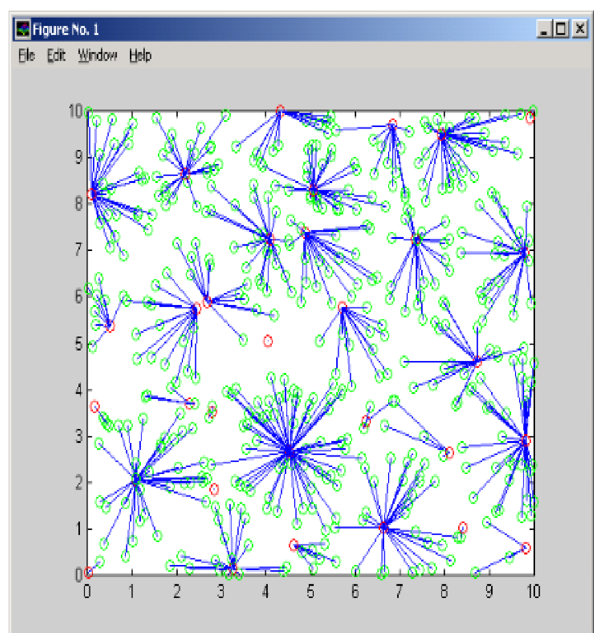

Fig. 2. Simulation output of a single level clustering algorithm

The time complexity for most of the clustering algorithms in the literature (LCA [2], LCA2 [8] and the Highest Degree [9, 23] algorithms) are O(n) , hence they are not realistic for a constrained sensor networks that have large number of sensors. The time complexity for The Max- Min d-Cluster Algorithm [5] is O(d), this may be acceptable for large networks. The time complexity for the Hierarchy clustering algorithm (HCA) herein referred to as the previous researcher is log.O(d). Based on this facts the Improved Hierarchy clustering algorithm proposed in this paper (IHCA) is compared with the Max-Min d-Cluster Algorithm (for d = 2,3,4,)and the Hierarchy clustering algorithm. The performance measure used in the simulation was total energy spent in the system. The experiments were conducted for networks of different densities. The probability of becoming cluster head computed from the optimal value (opt p) using equation 10 was used for each network density. The theoretical computation was also used to compute the size of each cluster. For the simulation analysis, the probability of the sensors in each cluster to assume cluster head status was set to the optimal value using equation 10, similarly the maximum number of hops (k) allowed between any sensor and its cluster head was equal to the value calculated using opt p in equation 12.

Table 1. Parameters for Minimizing Energy in the Algorithm

Number of Sensors (n)	Spatial Density (d)	Probability (opt P)	Maximum number of Hops (k)
500	5	0.10123	5
1000	10	0.07921	4
1500	15	0.06883	3
2000	20	0.06224	3
2500	25	0.05761	3
3000	30	0.05412	3

Table 1 shows the computed values of opt p and the corresponding values of maximum number of hops (k) in a cluster for networks of various densities. Fig 4 shows the results of the simulation experiments based on the comparison of the Max-Min d-algorithm with the Hierarchical clustering algorithm (HCA) and the Improved Hierarchical clustering algorithm (IHCA) proposed in this paper. From Fig 4, it can be deduced that the IHCA on the average, reduced the total energy spent by 12% compared to HCA and by 42% and 48% compared to the Max-Min d-clustering algorithm with d = 2 and 3 respectively. This reduction can be attributed to the setting the values for allowable number of sensors in a cluster to between 3 and k as opposed to the partially opened cluster size used in HCA. It can also be observed that the energy savings increases as the density of sensors in the network increases. This can also be attributed to the fact that an increase in the density of sensors will lead to more isolated sensors in the networks.

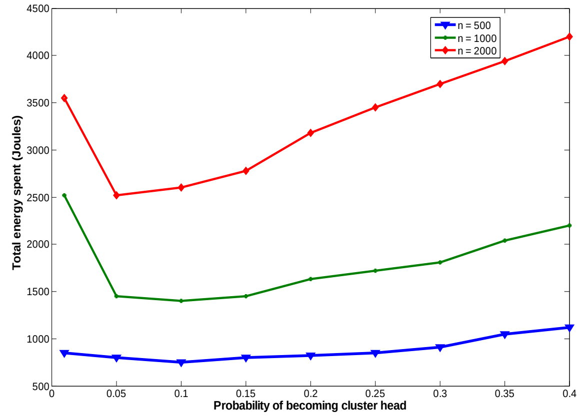

Fig.3 . Probability of becoming a cluster head and Total Energy Spent

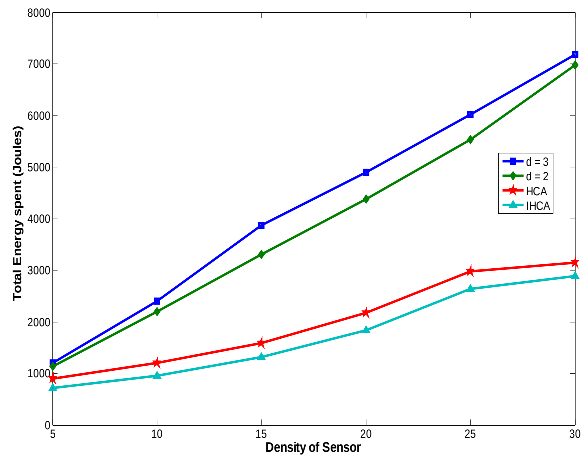

Fig. 4. Comparison of IHCA with HCA and the Max-Min D-Cluster Algorithms .

Isolated sensors here are defined as sensors without neighbour nodes within their transmission range. Such isolated sensors further reduce the network lifespan, hence the progressive increase in energy savings in IHCA as compared to HCA as network density increases.

A. *An Energy-Efficient, Hierarchical Clustering Algorithm*

In the previous section we designed an analytical model for one level of clustering, this model will be extended for the design of different levels of clustering herein defined as hierarchical clustering algorithm. It will be assumed that there are m levels in the clustering hierarchy with level 1 being the lowest level and level m being the highest. In this scenario, the sensors in the first level communicate the gathered data to level-1 cluster heads (CHs). This level-1 CHs now aggregate the data and communicate the compressed version of the aggregated data to level-2 CHs and so on (the compression algorithm is beyond the scope of this paper, it is however assumed that it is such as would not be cumbersome for the memory constrained sensors). Finally, the level-m CHs communicate the aggregated data or compressed version of the aggregated data to the processing centre. The cost of communicating the information from the sensors to the processing centre (sink) is the sum of the energy spent by the sensors to communicate the information to level-1 cluster heads (CHs), and the energy spent by the level-1 CHs to communicate the aggregated information to level-2 CHs, …, up to and including the energy spent by the level-m CHs to communicate the aggregated information to the information processing centre.

B Algorithm

This algorithm works in a bottom-up fashion. The level-1 cluster heads are first elected, the level-2 cluster heads, then level-3 cluster heads, and so on. The processes for choosing the level-1 cluster heads are as follows. As opposed to the previous researcher's method, where each sensor in level-1 becomes a CH with probability p_1, after which it advertises itself as a cluster head to the sensors within its radio range, in this paper, only sensors closer to the centre of the cluster can be elected as cluster heads. This has the advantage of reducing the overhead in initially choosing a cluster head for a given cluster hierarchy. It is also assumed that making the cluster head to be at the centre will reduce transmission energy consumption within the cluster. The few qualified sensors within the cluster then advertise themselves by forwarding the control packet to all the sensors within k hops of the advertising CH. Each sensor that receives an advertisement joins the cluster of the closest level-1 CH; any sensor without any neighbour within its transmission range is excluded from the network. The next step involves the election of level-2 CHs from Level-1 CHs. The level-2 CH is elected from the pool of leval-1 CH with probability p_2. and thereafter broadcast their decision of becoming a level-2 CH. This decision is forwarded to all the sensors within k_2 hops. The level-1 CHs that receive the advertisements from level-2 CHs joins the cluster of the closest level-2 CH. All other level-1 CHs become forced level-2 CHs, this is a deviation from our earlier assertion of excluding such sensors because in this case these particular set of sensors contain compressed aggregated data from other sensors, and it may be catastrophic o exclude such from the network. Cluster heads at level 3,4,…..,m are chosen in similar fashion, with probabilities $p_3, p_4, \ldots p_m$ respectively, to generate a hierarchy of CHs, wherein any level-i CH is also a CH of level (i-1), (i-2),…, 1. The flowchart for the multiple level cluster head formation is shown in Fig. 5.

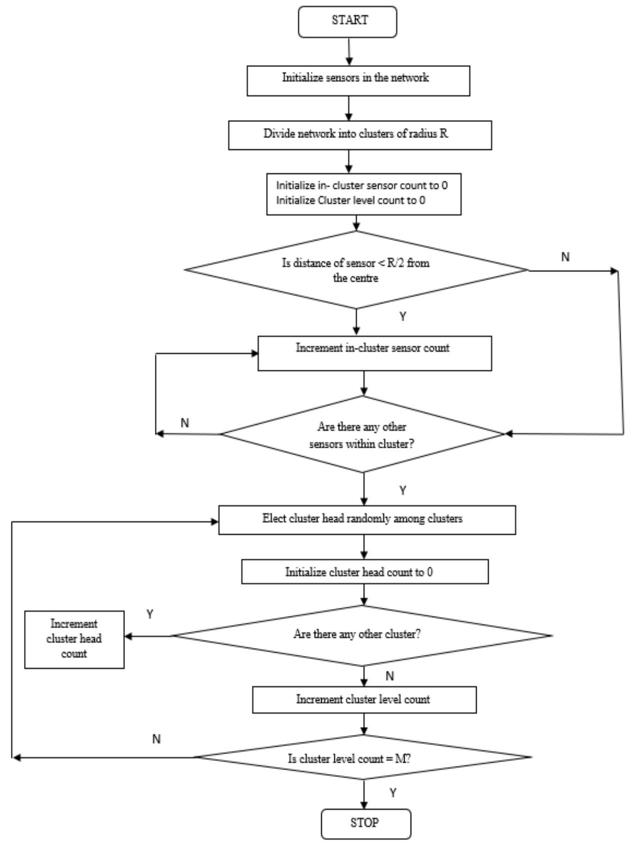

Fig. 5. Flowchart for multiple level (Hierarchical) cluster formation

C Determining Optimal parameters for hierarchical clustering algorithm

The energy required to transmit the aggregated data by the sensors to the information processing centre through the hierarchy of cluster heads will depend on the maximum number of hops within a cluster and probabilities of becoming a cluster head at each level in the hierarchy. In this subsection, the optimal values for the parameters of the algorithm described in Section IV-A will be obtained, such as would minimize this energy consumption. In order to achieve this, the same assumptions that was made as in Section 3B will be used. As pointed out in section 3B, the sensors are points of a homogeneous Poisson process of intensity λ , the number of sensors in a square area of side 2b is a Poisson random variable (assume this to be N) with mean λA , where $A = 4b^2$is the area of the square. For a particular realization of the process, it was assumed that there are n sensors in this area. The following definitions are also given:

Ni: Number of sensors in a level-i cluster,

Li: Sum of distances between the members of a level-i cluster and their level-i CH, Hi: Number of hops from a member to its CH in a typical level-i cluster,

TCHi: Total number of level-i CHs,

Ci: Total cost of communicating information from all level-i CHs to the level-(i+1) CHs, and

C: Total cost of communicating information from the sensors to the data processing centre through the hierarchy of cluster heads generated by the clustering algorithms.

As previously stated in the proposed algorithm, the sensors elect themselves as level-1 CH with probabilities p_1 and the level-i CHs elect themselves as level-(i+1) CHs with probability p_{i+1}, I = 1,2,,…(m-1).

Therefore, using the properties of the Poisson process, level-i CHs, I = 1,2,…,m are governed by homogeneous Poisson processes of intensities, $\lambda_{i1} = \lambda\prod_{i-1}^{i} p_j$. Using similar arguments to those in Section 3A, the sum total of distance of level-(i-1) CHs from a level-i CH, I =2,3,….,m in a typical level-i cluster is given by

$$E[L_i \mid N = n] = \frac{(1-p_i)\lambda \prod_{j=1}^{i-1} p_j}{2\left\{\lambda \prod_{j=1}^{i} p_j\right\}^{3/2}} \quad (13)$$

The expected number of level-(i-1) CHs in a typical level-i cluster is given by

$$E[N_i \mid N = n] = \frac{1-p_i}{p_i} \quad (14)$$

The expected number of hops between a level-(i- 1) CH and its level-i CH in a typical level-i cluster is therefore given by

$$E[H_i \mid N = m] = \left\lceil \frac{1}{r}\frac{E[L_i|N=m}{E[N_i|N=m} \right\rceil$$

$$= \left\lceil \frac{1}{2r\sqrt{\beta \prod_{j=1}^{i} p_j}} \right\rceil \quad (15)$$

The expected number of level-i CHs is given by

$$E[TCH_i \mid N = m] = n\prod_{j=1}^{i} p_j \quad (16)$$

Therefore the total expected cost of communicating information from all the level-(i-1) CHs to their respective level-i CHs, I = 2,……., (h – 1), h is given by

$$E[C_{i-1} \mid N = m] = E[TCH_i \mid N = m]E[N_i \mid N$$

$$= m[E[H_i \mid N = m] \quad (17)$$

The total cost of the expected value of communicating information from all the sensors in the network to their level-1 CHs is given by

$$E[C_0 \mid N = m] = E[TCH_i \mid N = m]E[N_i \mid N$$

$$= n]E[H_i \mid N = m] \quad (18)$$

Therefore the total expected cost of communicating information from all the sensors in the network to the processing centre (sink) in the clustered environment is given by:

$$E[C \mid N = m] = n\prod_{i=1}^{m} p_i \left\lceil\frac{0.7647\,a}{r}\right\rceil + \sum_{i=0}^{m-1} E[C_i \mid N = m]$$

$$= n\prod_{i=1}^{m} p_i \left\lceil\frac{0.7647a}{r}\right\rceil +$$

$$n\sum_{i=1}^{m}(1-p_i)\prod_{j=1}^{i-1}(p_j)\left\lceil \frac{1}{2r\sqrt{\beta \prod_{j=1}^{i} p_j}} \right\rceil \quad (19)$$

By un-conditioning on N , we find:

$$E[C] = E[E[C \mid N = m] = \beta A\prod_{i=1}^{m} p_i \left\lceil\frac{0.7647\,a}{r}\right\rceil$$

$$+ \beta A \sum_{i=1}^{m}(1-p_i)\prod_{j=1}^{i-1}(p_j)\left\lceil \frac{1}{2r\sqrt{\beta \prod_{j=1}^{i} p_j}} \right\rceil \quad (20)$$

The function depicted in 20 has a very complex form, A MATLAB plot of the function shows that it has many minima points. Hence the analytical solution of the expression in equation 20 cannot be obtained. We thus switch to a numerical solution. This can be obtained by substituting the optimal probability obtained from equation 12 into the equation shown below:

$$K_1 = \left\lceil \frac{1}{r}\sqrt{\frac{-0.9172\,in\left(\frac{\alpha}{7}\right)}{\lambda i \prod_{j=1} p_j}} \right\rceil \quad (21)$$

From (21) α denotes the probability that the number of hops between a member and its cluster head in a level-i cluster is more than k_i i = 1,2,……,m.

D. Numerical Results and Simulations

In this section, an analysis is made as regards the simulation experiments performed with respect to the algorithm described in section IVA. The sensors were distributed uniformly with various spatial densities. It was assumed that in order to communicate 1 unit of data, 1 unit of energy will be spent. This algorithm was used to generate a clustering hierarchy with different number of levels. A comparison of Energy consumption involving various levels of clustering was done with the HCA and the IHCA clustering algorithm proposed in this paper. From figure 6 it can be deduced that IHCA outperform HCA in terms of reduction in total energy spent by 24%. It can also be seen that in both algorithms, the total energy spent decreases as the number of levels in the hierarchy increase. This means that the Hierarchical clustering algorithm is preferred to single level clustering algorithm especially when the size of the network is large which also means the number of sensors will be high especially in thousands.

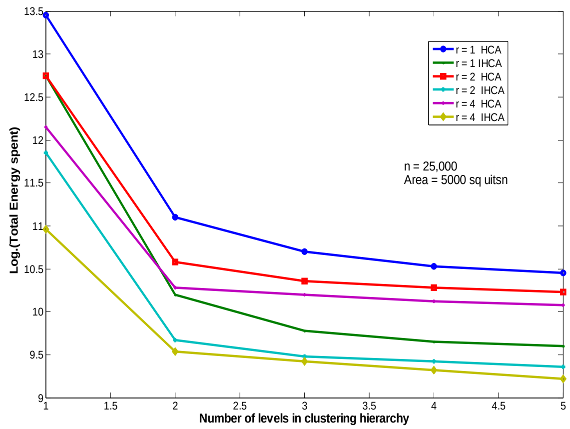

Fig. 6. Total Energy Spent vs. number of levels in the clustering hierarchy in a network of 25000 sensors with communication radiir distributed in a square area of 5000 sq. units.

It can also be deduced from Fig 6 that there was reduction in total energy spent for network of sensors with higher communication radius. The reason for this is because, the higher the communication radius, the lower the number of hops required for data to reach individual cluster heads .This also translates to lower number of hops to get to the processing center (sink). In Fig 7, an increase in the number of hierarchies also led to a more significant reduction in the energy savings.

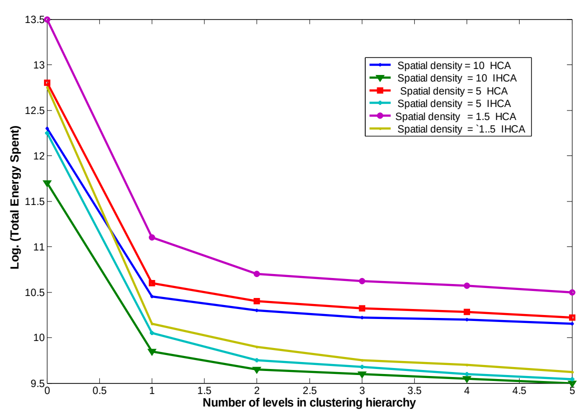

Fig. 7. Total Energy Spent vs. number of levels in the clustering hierarchy in a network of 25000 sensors of communication radius 2 distributed with spatial density λ.

The reason for this is because, the clustering algorithm with higher levels of clustering signifies that transmission of data is distributed among many labels as opposed to a single level.. Hence there is bound to be a reduction in transmission distance because many clusters will be closer to the processing centre as opposed to a single level where the single cluster may be far and therefore incur a lot of energy dissipation.. In terms of energy savings the hierarchical clustering, IHCA outperforms HCA by 26%.

Finally from Fig 6 and Fig 7, the energy consumption is minimized when the number of hierarchical levels is 5 irrespective of the density of sensors and their communication radius. Therefore if one chooses to store the numerically computed values of optimal probability in the sensor memory, only a small amount of memory would be needed.

It must be stated here that the sensors that are elected cluster head in the two algorithms proposed in this paper spends more energy than the other sensors within their respective clusters. In this vein it will be necessary to run the clustering algorithm periodically as proposed in [5], so as to enable load balancing. Another possibility is that cluster heads trigger the clustering algorithm when their energy levels fall below a certain threshold.

V. CONCLUSIONS AND FUTURE WORK

This paper proposed a distributed algorithm for organizing sensors into a hierarchy of clusters with an objective of minimizing the total energy spent in the system to communicate the information gathered by these sensors to the information-processing center. The optimal parameter values for these algorithms that minimize the energy

spent in the network have been analytically determined. In a contention-free environment, the IHCA algorithm as a time complexity of $O(\log(K_1 + K_2 +,,,,,K_n))$ an improvement of $O(K_1 + K_2 +,,,,,K_n)$ in th HCA algorithm and a significant improvement over the time complexity of O(n) in the algorithms proposed in [2,3,4,8,9]. This makes the new algorithm suitable for networks of very large number of nodes. In this paper, it was assumed that the communication environment is contention and error free; in future an underlying medium access protocol will be considered and an investigation on how that would affect the optimal probabilities of becoming a cluster head and the run-time of the algorithm will be performed.

REFERENCES

[1] Pottie G. J and Kaiser W. J "Wireless Integrated Network Sensors", Communications of the ACM, Vol. 43, No. 5, pp 51-58, May 2015.

[2] Baker D. J. and Ephremides A. "The Architectural Organization of a Mobile Radio Network via a Distributed Algorithm", IEEE Transactions on Communications, Vol. 29, No. 11, pp. 1694-1701, November 2000.

[3] Das B. and Bharghavan V. "Routing in Ad-Hoc Networks Using Minimum Connected Dominating Sets", in Proceedings of ICC, 2014.

[4] Lin C. R. and Gerla M. "Adaptive Clustering for Mobile Wireless Networks", Journal on Selected Areas in Communication, Vol. 15 pp. 1265-1275, September 2016.

[5] Amis A. D. Prakash R, Vuong T. H. P. and Huynh D. T., (2017) " Max-Min D-Cluster Formation in Wireless Ad Hoc Networks", in Proceedings of IEEE INFOCOM, March 2017.
[6] Chiasserini C. F. Chlamtac I. Monti P and Nucci A., "Energy Efficient design of Wireless Ad Hoc Networks", in Proceedings of European Wireless, February 2018.
[7] Ephremides A. Wieselthier J. E. and Baker D. J "A Design concept for Reliable Mobile Radio Networks with Frequency Hopping Signaling", Proceeding of IEEE, Vol. 7, No. 1, pp. 56-73, 2016.
[8] A. K. Parekh A. K., "Selecting Routers in Ad-Hoc Wireless Networks", in Proceedings of ITS, 2014.
[9] Okabe A, Boots B, Sugihara K. and Chiu S. N. Spatial Tessellations: Concepts and Applications of Voronoi Diagrams, 2nd edition, John Wiley and Sons Ltd. 2016
[10] Foss S, G. and Zuyev S, A "On a Voronoi Aggregative Process Related to a Bivariate Poisson Process", Advances in Applied Probability, Vol. 28, no. 4, pp. 965-981, 2014.
[11] Kahn J. Katz M. R. H and Pister K. S. J, "Next Century Challenges: Mobile Networking for Smart Dust", in the Proceedings of 5th Annual ACM/IEEE International Conference on Mobile Computing, and Networking (MobiCom 2014), pp. 271-278, Aug, 2014,
[12] Baccelli F. and Zuyev S, "Poisson Voronoi Spanning Trees with Applications to the Optimization of Communication Networks", Operations Research, vol. 47, no. 4, pp. 619-631, 2017.
[13] Warneke B., Last M,, Liebowitz B,, Kristofer and Pister S. J, "Smart Dust: Communicating with a Cubic-Millimeter Computer", Computer Magazine, Vol. 34, No. 1, pp 44-51, Jan. 2018.
[14] Kahn J M., Katz R. H. and Pister K. S. J, "Next Century Challenges: Mobile Networking for Smart Dust", in the 5th Annual ACM/IEEE International Conference on Mobile Computing and Networking (MobiCom 2013), pp. 271-278, Aug 2013
[15] Hsu V., J. Kahn J.M, and Pister K. S. J., "Wireless Communications for Smart Dust", Electronics Research Laboratory Technical Memorandum/ 2, Feb. 2014. http://www.janet.ucla.edu/WINS/wins_intro.htm.
[16] Gupta P. and Kumar P. R., "The Capacity of Wireless Networks,", IEEE Transactions on Information Theory, vol. IT-46, no. 2, pp. 388404, March 2016.
[17] Gupta P.and Kumar P. R. "Critical Power for Asymptotic Connectivity in Wireless Networks", pp. 547-566, in Stochastic Analysis, Control, Optimization and Applications: A Volume in Honour of Fleming W. H.. Edited by W.M. McEneany W.M , G. Yin G , and Q. Zhang Q., Birkhauser, Boston, 2015. ISBN 0-8176-4078-9.
[18] Ye W., Heidemann J., and Estrin D. "An Energy-Efficient MAC Protocol for Wireless Sensor Networks", In Proceedings of the 21st International Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM 2017), New York, NY, USA, June, 2017.
[19] Bulusu N., Estrin D., Girod L., and Heidemann J. "Scalable Coordination for Wireless Sensor Networks: Self-Configuring Localization Systems", In Proceedings of the Sixth International Symposium on Communication Theory and Applications (ISCTA 2016), Ambleside, Lake District, UK, July 2016.
[20] Bulusu N. , Heidemann J., and Estrin D., "Adaptive beacon Placement", Proceedings of the Twenty First International Conference on Distributed Computing Systems (ICDCS-21), Phoenix, Arizona, April 2017.
[21] McDonald A. B., and Znati T., "A Mobility Based Framework for Adaptive Clustering in Wireless Ad-Hoc Networks", IEEE Journal on Selected Areas in Communications, Vol. 17, No. 8, pp. 1466-1487, Aug. 2015.
[22] Gerla M., and Tsai J. T. C., "Multi-cluster, Mobile, Multimedia Radio Networks", Wireless Networks, Vol. 1, No. 3, pp. 255-265, 2014.
[23] Basagni S., "Distributed Clustering for Ad Hoc Networks", in Proceedings of International Symposium on Parallel Architectures, Algorithms and Networks, pp. 310-315, June 2016.
[24] Basagni S. "Distributed and Mobility-Adaptive Clustering for Multimedia Support in Multi-Hop Wireless Networks", in Proceedings of Vehicular Technology Conference, Vol. 2, pp. 889-893, 2016.
[25] Seema B. and Edward J. C. ' An energy efficient Hierarchical Clustering Algorithm for Wireless Sensor Networks in IEEE INFOCOM 2019 pp 362-378.. Sept 2019
[26] Amis A. D., and Prakash R. , "Load-Balancing Clusters in Wireless AdHoc Networks", in Proceedings of ASSET 2016 , Richardson, Texas, March 2016.
[27] Perrig A., Szewczyk R., Wen V. and Tygar J. D., "SPINS: Security protocols for Sensor Networks", in 7th Annual International Conference on Mobile computing and Networking, , pp. 189-199, 2017
[28] Carman D. W., Kruus P. S., and Matt B. J., "Constraints and approaches for distributed sensor network security", NAI Labs Technical Report 00-010, September 2016.