IOSR Journal of Engineering (IOSRJEN) **www.iosrjen.org**
ISSN (e): 2250-3021, ISSN (p): 2278-8719
PP 11-18

A Unique Cluster for Homogeneous and Heterogeneous Wireless Sensor Networks

Utkarsha S Pacharaney[1], Rajiv Kumar Gupta[2]
[1](Department of Electronics, Datta Meghe College of Engineering, Airoli, India)
[2](Department of Electronics and Telecommunication , Terna Engineering College,Nerul, India)

Abstract: *Internet of Things (IOT) is an amalgamation of both homogenous and heterogeneous Wireless Sensor Network (WSN). Heterogeneous WSNs are a variant of homogenous WSN, where more energy-enabled nodes are deployed along with normal nodes to take the responsibility of energy draining tasks. In IOT, heterogeneity is a combination of various devices carrying out various tasks with the expectation of cooperation to solve these tasks to achieve superior performance. Grouping these devices to perform multi-task could simplify their task accomplishment. Grouping of devices in traditional WSN is clustering. However, the clustering parameters used for homogeneous WSN are not applicable for heterogeneous WSN. In this paper, we propose a novel clustering technique for IOT that works for both homogeneous and heterogeneous WSN. We partition nodes based on their spatial location and nodes are clustered with their closest neighbor if the distance between them is shorter than the sensing range. For homogeneous network, with this clustering technique, a strong correlated cluster is set up. Hence, only the cluster head is required to carry out the sensing and transmitting task on behalf of the entire cluster. Further, we apply Compressive Sensing (CS) on Cluster Head (CH) to prolong the network lifetime. For heterogeneous network, the node with the highest sensing range will take the responsibility of carrying out tasks on behalf of other like devices. We prove a relation exists between distance, correlation and sensing range. We have used MTALAB to simulate and form clusters for the proposed technique and demonstrated that low RMSE is obtained by applying CS at the CH for homogenous network.*

Keywords: *Homoegnous Networks; Heterogenous Networks;Sensing Range; Internet of Things; Clustering*

I. Introduction

Wireless Sensor network (WSN) is a collection of spatially distributed sensor nodes, which work collaboratively to fulfill an assigned task. With the shift of modern world to the new age of Internet of Things (IOT), major evolution are bound to take place in WSN assisted IOT. The foundation of IOT is an integration of heterogeneous devices like smart phones, laptops, PDAs and other futuristic devices along with heterogeneous technologies like Radio Frequency Identification (RFID) tags, sensor/actuators and communication technologies [1]. These heterogeneous devices, performing multiple tasks, are expected to co-operate and solve these tasks to achieve superior performance [2]. However, the topology of these ad hoc heterogeneous devices is usually unknown but of paramount importance, even important than homogeneous network devices. Hence, grouping these devices based on some parameter and then performing multi-task could simplify their task accomplishment. This grouping of devices together in traditional WSN is clustering. Clustering is well studied in WSN, which divides nodes into disjoint sets with respect to some criteria like number of clusters required [3], nodes energy level [4],intercommunication distance [5].These clusters are typically for homogeneous WSN where nodes having similar capabilities and similar goals. The goal is to conserve and balance energy in a battery operated sensor node to prolong network lifetime. However, heterogeneous network, where some nodes are more energy able than others and perform energy draining task, are not constrained by energy. Hence, the aim to prolong network lifetime in heterogeneous network is achieved with different partitioning criteria such as distance between nodes, distance between nodes and base station. Thus, clustering parameters are different for homogeneous and heterogeneous networks as presented in survey [6]. However, IOT is an amalgamation of both homogenous and heterogeneous network. Can we cluster nodes in a network with common parameter that works for both homogenous and heterogeneous WSN? One parameter that stands common between them is their or position or location that is spatial coordinates. Considering spatial coordinates, we build a cluster that works for both type of networks. Clusters based on spatial distribution are already present in literature for homogeneous WSN. These clusters are based on the combination of spatial correlation with clustering technique. Spatial correlation for homogeneous WSN is defined as: Given a set S containing N nodes, spatial correlation refers to the similarity of reading between these N nodes within a geographical proximity for all $N(N-1)$ nodes. However, heterogeneous WSNs with spatial clustering in IOT, does not seek to find similarity of readings between nodes rather a method to enhance their cooperation to achieve superior performance is expected. Hence, the spatial clusters used in homogenous network will not work for heterogeneous one.

In this paper, we propose a spatial clustering technique that works for both homogenous and heterogeneous WSN in IOT. Different from other spatial clusters discussed in section II, we cluster nodes based on their spatial coordinates if distance between them is shorter than sensing range. For homogeneous network, with this clustering technique, a strong correlated cluster is set up. Hence, only the cluster head is required to carry out the sensing and transmitting task on behalf of the entire cluster. Further, we apply Compressive Sensing (CS) on Cluster Head (CH) to prolong the network lifetime. For heterogeneous case, node with highest sensing range will take the responsibility of grouping nodes and will carry out task on behalf of other devices.

The organization of the paper is as follows: section 2 presents a related work. Derivation of relation between distance, correlation and sensing range is in section 3. The novel node-clustering algorithm is given in section 4. Implementation of the proposed clustering technique is in section 5. Section 6 presents the overview of compressive sensing. We conclude the paper in section 7.

II. Related work

Various clustering algorithms are presented in [7] for homogenous WSNs whereas various clustering algorithm are presented in [8] for heterogeneous WSNs. The heterogeneous WSN in these survey papers are a variant of homogenous WSN in which some more energy-enabled nodes are added along with the homogenous nodes to prolong the network lifetime. A true heterogeneous network, were all nodes have different capabilities and different tasks to perform as given in [2], requires cooperation to carry out tasks together .Hence, clusters used for homogenous WSN will not work for heterogeneous WSN. In addition, the comparative survey given in [6], shows that clustering parameters for both of them are different. One parameter, that makes the same cluster to work for homogenous and heterogeneous network, is position or spatial location of nodes. For homogenous WSN, work on spatial clusters has been done, which combines spatial correlation with clustering, but the volume of the work is very little as presented in this section.

In [9] Vuran *et.al*. revealed that significant energy saving is possible by allowing less number of nodes to send information to sink in dense WSN. These nodes are the cluster heads of a spatially correlated cluster formed during the contention of the media at the MAC layer. The node that capture channel after first contention phase becomes representative node of the area determined by correlation radius r_{corr} . Yuan *et.al*. [10] Integrated clustering technique with spatial-correlation and proposed two algorithms Cluster Construction (CC) and K-hop Cluster Construction (KCC), in which spatial correlation is estimated between nodes by calculating similarity degree over a time period. C Hung *et.al*. [11] proposed to group sensor nodes with similar reading and only representative node reporting to sink. To solve the challenging task of identifying sensor groups and representative nodes, centralized DCglobal and distributed DClocal algorithm were proposed. The nodes with the highest energy level and wide data coverage range are selected as representative node. Zhidan *et.al.* in [12] proposed Distributed Spatial Correlated-based Clustering (DSCC). By exploiting spatial correlation range R_{sc} and application specific parameter namely, error-tolerance ϵ,the algorithm selects node with high residual energy as cluster heads. The Data Density Correlation Degree (DDCD) algorithm proposed in [13] calculates the correlation within the communication range of the sensor nodes and forms correlated clusters. The spatial clusters are constructed with the aim to predict the readings of nearby neighboring nodes, to avoid redundant transmission. Apart from ignoring the cost of learning the correlation, the drawbacks of these correlated clusters is that there is no uniform measure on tolerance error range or distance between sensors. Also, abundant communication overhead for spatial clustering, more number of iterations to select the CH and above all, the construction of cluster takes several rounds of message exchange and computation of correlations. For heterogeneous network, other than using spatial correlation different parameters are used while clustering. Hence, we propose clustering technique, which works for both homogeneous and heterogeneous WSN. For homogeneous network the cluster so formed will be highly correlated hence we can further collect the data compressively to prolong the network lifetime. Various in network compression approaches exist in WSN such as entropy coding, principal component analysis or transform coding [14] [15], [16].However, these include significant computations and control overheads that are often not suitable for sensor networks. CS is a novel paradigm that senses the data in its compressed form and can be used effectively for gathering data in a WSN. By leveraging ideas from CS, we can enable more economic use of sensing resource in an energy-constrained homogenous network. For heterogeneous network, the node with the highest sensing range takes the responsibility to carry out task on behalf of other nodes in the cluster. Based on order of cluster formation that is whether cluster is formed or cluster head is selected first, two approaches are present: Cluster First (CF) approach and Leader First (LF) approach. In this work we follow the cluster first approach similar to [17][18][19] with the cluster head at the center of the cluster. The choice of the cluster head at the center is as given in [20].

The unique cluster that we propose in this work, take node location i.e. spatial coordinates, calculate the distance between nodes, and then group them within their sensing range to form a cluster i.e. the radius of the cluster is the sensing range. For homogenous network this lead to a correlated cluster. The relationship between distance, correlation and sensing range is given in section III.

III. Relation between Distance, Correlation and Sensing Range

Assume two sensor nodes with sensing range rs_1 and rs_2 located at spatial location$(0,0)$ and $(d, 0)$ respectively as shown in Fig. 1

Using geometry to set up model to find relation between distance, sensing range and correlation with meaning of symbols explained as follows:

R_i and R_j denotes sensing range of node $n_i$and n_j of disk with radius rs_1 and rs_2 respectively .

A_i Area R_i denoting area of R_i

A_j Area R_j denoting area of R_j

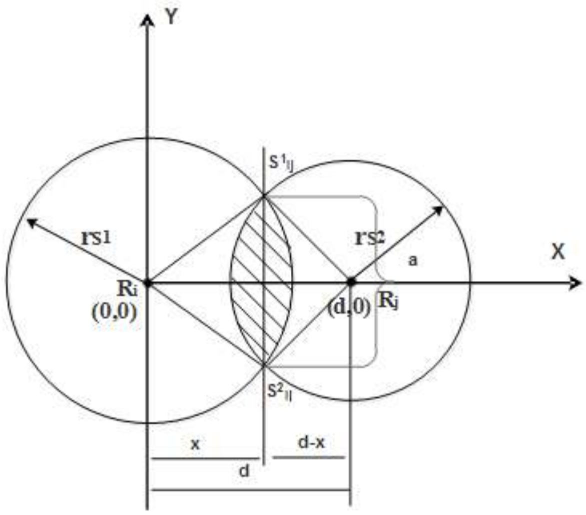

Fig1: Model to find relation between distance, sensing range and correlation.

R_i^j region demarcated by the perpendicular bisector of R_i and R_j and next to belongs to R_i

A_i^j Area R_i^j denoting the area of R_i^j

A_j^i Area R_j^i denoting the area of R_j^i

R denotes the sensing region

A denotes the area of R

The Euclidean distance between R_i and R_j is given by:

$$\sqrt{(0-d)^2 + (0-0)^2} = d$$

S_{ij}^1 and S_{ij}^2 are intersection points of the two nodes

a length of the common chord joining S_{ij}^1 and S_{ij}^2

$$a = \frac{1}{d}\sqrt{4d^2 {rs_1}^2 - (d^2 - {rs_2}^2 + {rs_1}^2)} \quad (1)$$

x is the distance between R_i and chord a

$$x = \frac{d^2 - {rs_2}^2 + {rs_1}^2}{2d} \quad (2)$$

If $d < rs_1 + rs_2$ then R_i overlaps with R_j and correlation is defined as

$$K_\vartheta(d_{ij}) = \frac{A_i^j + A_j^i}{A} \quad (3)$$

$A_i^j + A_j^i = A^{int}$ is the area of the asymmetry lens which intersect the sensing range of two sensor node and is calculated as using the formula of circle segment of radius R' and triangle height d'

$$A^{int}(R', d') = R'^2 \, arccos(\frac{d'}{R'}) - d'\sqrt{R'^2 - d'^2} \quad (4)$$

Hence

$$A^{int} = A(rs_1, x) + A(rs_2, d - x) \quad (5)$$

$$A^{int} = rs_1{}^2\, arccos\left(\frac{d^2 + rs_1{}^2 - rs_2{}^2}{2drs_2}\right) + \quad rs_2{}^2\, arccos\left(\frac{d^2 + rs_1{}^2 - rs_2{}^2}{2drs_1}\right) - \frac{1}{2}\sqrt{4d^2 rs_1 - (d^2 - rs_2 + rs_1)^2} \quad (6)$$

Hence, from 2 we obtain

$$K_\vartheta(d_{ij}) = \frac{A^{int}}{A} \quad (7)$$

For the case when the sensing range of the two nodes is same i.e. $rs_1 = rs_2 = r$

$$K_\vartheta(d_{ij}) = \frac{2}{\pi}\arccos\left(\frac{d}{2r}\right) - \frac{d}{\pi r^2}\sqrt{r^2 - \frac{d^2}{4}} \quad (8)$$

We observe that the correlation is a function of Euclidean distance and the sensing range of the sensors. When $d = rs_1 + rs_2$ then correlation is zero while when $d = 0$ then maximum correlation exist between nodes. For any value, $1 < d < rs_1 + rs_2$ correlation exists between nodes. In addition, if any nodes lies within the sensing range of a node, it will be correlated with the other. Considering this, we construct a cluster with the radius of the cluster equal to the sensing range. For homogeneous network, the nodes within this correlated cluster can be put to sleep with only CH sensing and sensing compressed data to the BS. While for heterogeneous network, the CH is the node with the highest sensing range and it will carry out the processing task for all the nodes within its sensing area. The algorithm at the BS is given in the next section

IV. Node Clustering based on Spatial location

In this section, we present a novel clustering algorithm with the following assumptions

1. Sensor nodes knows their geographic location via GPS or other localization algorithms
2. Base Station (BS) knows the entire region were the nodes are randomly deployed via UAV.
3. The sensing range of the node is half the communication range [21].

In this paper, Gaussian probability distribution function is used to model the spatial distribution of N sensor nodes in aregion $A \in \mathbb{R}^2$, where $\mathbb{R}^2$ denotes the 2-D Euclidean domain. Let $N_1, N_2, \dots \dots \dots N_n$ denote nodes with spatial location $(x_1, y_1), (x_2, y_2), \dots \dots \dots .(x_n, y_n)$ .The sensing range of the nodes is given as $r_{s_1}, r_{s_2}, \dots \dots r_{s_n}$ The BS broadcasts network wide tuple, requesting location information, energy and sensing range i.e. (W^L, W^E, W^S) from each node. Nodes reply by sending these three values in the RREP packet along with their ID. After receiving the data from all the nodes, the BS runs the algorithm 1 to form clusters.

Algorithm 1

1. Note the(x_i, y_i) coordinates for each node, where $i, \forall\ 1, n$
2. Compute pairwise distance between all the nodes d_{ij}

$$d_{ij} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}$$

3. Sort the distance between nodes $N_i \& N_j$ in ascending order $'D_s'$ along with their node IDs
4. Select highest sensing range among $r_{s_1}, r_{s_2}, \dots \dots r_{s_n}$
5. Start clustering with $N_i \& N_j$

a. If $D_s \leq$ sensing range
   Begin new cluster
b. Check whether node $N_i \& N_j$ already exist in some other cluster
c. If already in different cluster and distance between them less than sensing range merge clusters.

6. Thus clusters, $C_1, C_2, \dots \dots \dots C_k$ where $k \ll n$, are formed
7. Compute the centroid for each cluster
8. The node with the value nearest to the centroid is elected as a cluster head.

Thus, clusters are formed and cluster heads selected by the BS are announced in the networks.

V. Implementation

In this section, we demonstrate our clustering algorithm with the sensors nodes distributed in Berkeley Research Lab [22]. The deployment of nodes in Berkeley Lab is as shown in Fig 2.

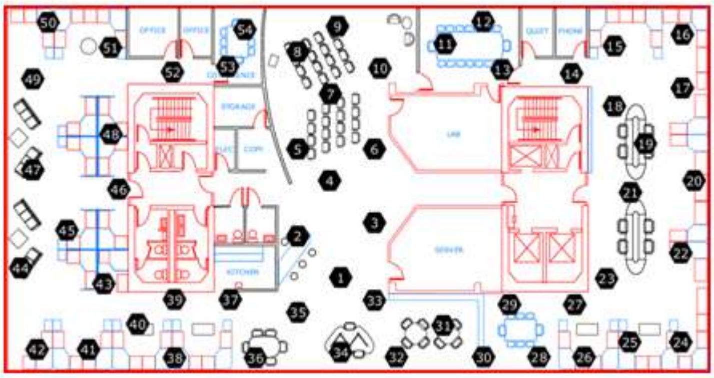

Fig. 2: Intel Berkeley Lab

Table I: Clusters with their spatial location

[26.5,2;24.5,4;21.5,2;19.5,5;16.5,3;13.5,1;12.5,5;8.5,6;22.5,8]
[4.5,30;1.5,30]
[10.5,31;7.5,31;13.5,31]
[39,37,35,40,43]
[30.5,26;27.5,26]
[39.5,30;36.5,30]
[28.5,5;31.5,6]
[8.5,26;6,24;1.5,23]
[19.5,26;21.5,23;24.5,20;19.5,19]
[24.5,12;22.5,15;19.5,12]
[3.5,13;5.5,10;1.5,8]
[15.5,28;12.5,26;17.5,31;21.5,30]
[30.5,31;26.5,31]
[1.5,2;5.5,3]
[4.5,18;0.5,17]
[37.5,19;40.5,22;34.5,16;39.5,14;35.5,10]
[35.5,4;38.5,1;39.5,6]

According to their spatial location, nodes are clustered together and sixteen clusters as shown in Fig 3 a. are formed. The spatial location of nodes in each cluster are tabulated in table 1 Once clusters are formed according to steps 1 to 6 of the algorithm , the centroid of the cluster is computed. The centroid of each cluster is given in table 2.

Table II: Centroid of the Clusters

18.38889,4
3,30
10.5,31
30.3,26.2
38,30
30,5.5
5.333333,24.33333
21.25,22
22.16667,13
3.5,10.33333
16.75,28.75
28.5,31
3.5,2.5
2.5,17.5
37.5,16.2
37.83333,3.666667

The location of the CH in the cluster is the node whose spatial location matches the centroid of the cluster or near to that value. The BS finds the node whose value lies near to that of the centroid and it is assigned as the CH as shown in Fig. 3b. The CH of each cluster is given in table 3.

Table III: Selecting Cluster Head of Clusters

19.5,5
4.5,30
10.5,31
30.5,26
39.5,30
28.5,5
6,24
21.5,23
22.5,15
5.5,10
15.5,28
30.5,31
1.5,2
4.5,18
37.5,19
35.5,4

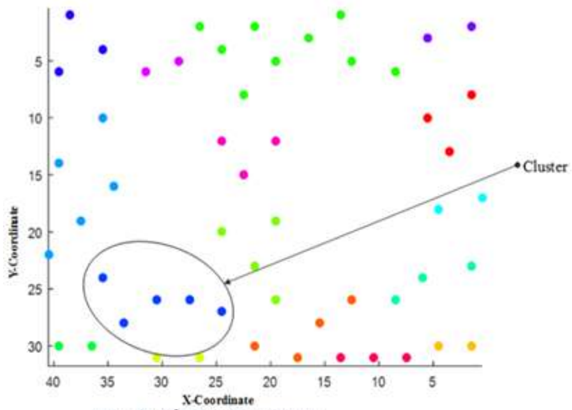

Fig. 3a Cluster Formation

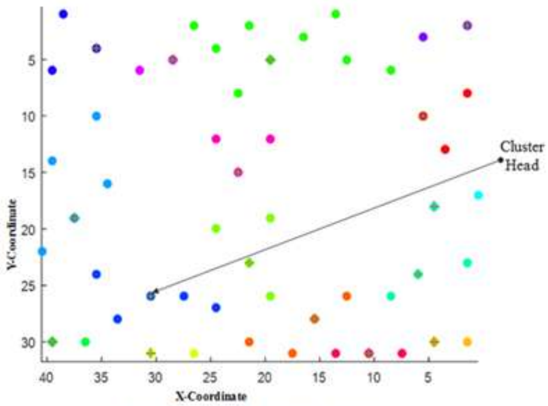

Fig. 3b Cluster Head Selection

Once the CH is selected, For heterogenous network the ch will carry the multi task on behalf of other nodes, while for homogenous network we apply CS on the CH to further compressively send the data to the BS. The overview of which is given in the next section.

VI. Overview of compression sensing

A brief idea of CS is as follows: Consider a signal $f = [f_1, f_2, \dots \dots \dots \dots . f_N]$ denoting a set of sensor readings from N nodes. If data f is k-sparse then we can obtain $Y = [Y_1, Y_2, \dots \dots \dots \dots . . Y_M]$,where $M \ll N$ but includes most of the information of f, by multiplying it with compression matrix $\emptyset$ .Also f can be recovered from Y by solving l_1 optimization problem i.e. $\underset{z \in \mathbb{R}^n}{min} \|Z\|_{l_1}$ subject to $Y = \emptyset \Psi Z$ where Ψ represents a proper basis such that $f = \Psi Z$ [23].

The advantage of using CS is twofold:

Firstly, CS compresses the data before collection as against traditional compression process that discarding the data after collection. Secondly, the expensive reconstruction of the sparse signal is done at the powerful BS, not burdening the battery-operated node.

We apply CS on the CH of cluster shown in fig 3a. Since the nodes are correlated, the other nodes in the network can be put to sleep and only the CH sends the data to BS. The original and reconstructed signal at the BS is shown in fig 4. To measure the accuracy of reconstruction we calculate the RMSE value, which is found to be 0.052. Taking the advantage of correlation we save on the number of transmissions and hence prolong the network lifetime.

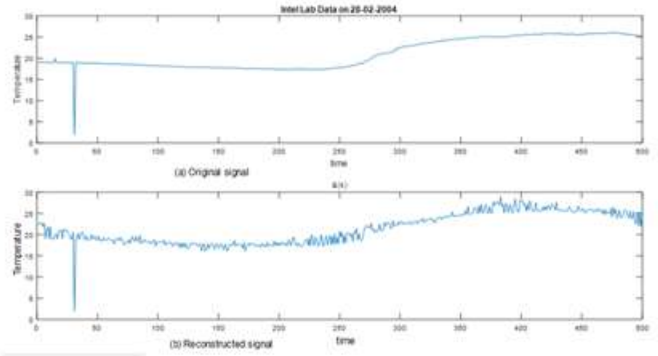

Fig4: Original and Reconstruction signal

VII. Conclusion

WSN is the key in IOT and IOT is the combination of both homogenous and heterogeneous devices. A unique cluster that works for both networks is proposed. The cluster is constructed based on the distance between the nodes, which are within the sensing range. Such a cluster is highly correlated for a homogenous network and hence only the CH can sense and transmit compressed data to the BS. The compression technique that is used for collecting data by the CH is compressing sensing. MATLAB simulation is used to demonstrate cluster formation and applying CS on CH. The original signal is reconstructed with very low value of RMSE. For heterogeneous network which aims to carry on multi-task the highest sensing range node with carry out the task on behalf of other nodes. Thus same cluster can work for both homogenous and heterogeneous networks.

Refrences

[1] Da Xu, Li, Wu He, and Shancang Li. "Internet of things in industries: A survey." *IEEE Transactions on industrial informatics* 10.4 (2014): 2233-2243.

[2] Plata-Chaves, Jorge, et al. "Heterogeneous and Multitask Wireless Sensor Networks—Algorithms, Applications, and Challenges." *IEEE Journal of Selected Topics in Signal Processing* 11.3 (2017): 450-465.

[3] Tandon, Ravi. "Determination of Optimal Number of Clusters in Wireless Sensor Networks." *arXiv preprint arXiv:1208.1982* (2012).

[4] Yu, Jiguo, et al. "A local energy consumption prediction-based clustering protocol for wireless sensor networks." *Sensors* 14.12 (2014): 23017-23040.

[5] Chen, Guihai, et al. "An unequal cluster-based routing protocol in wireless sensor networks." *Wireless Networks* 15.2 (2009): 193-207.

[6] Rostami, Ali Shokouhi, et al. "Survey on clustering in heterogeneous and homogeneous wireless sensor networks." *The Journal of Supercomputing* 74.1 (2018): 277-323.

[7] Liu, Xuxun. "A survey on clustering routing protocols in wireless sensor networks." *sensors* 12.8 (2012): 11113-11153.

[8] G. Eason, B. Noble, and I.N. Sneddon, "On certain integrals of Lipschitz-Hankel type involving products of Bessel functions," Phil. Trans. Roy. Soc. London, vol. A247, pp. 529-551, April 1955. (*references*)

[9] Vuran, Mehmet C., Özgür B. Akan, and Ian F. Akyildiz. "Spatio-temporal correlation: theory and applications for wireless sensor networks." *Computer Networks* 45.3 (2004): 245-259.

[10] Yuan, Jinhui, and Hong Chen. "The optimized clustering technique based on spatial-correlation in wireless sensor networks." *Information, Computing and Telecommunication, 2009. YC-ICT'09. IEEE Youth Conference on*. IEEE, 2009.

[11] Peng, Chih-Chieh Hung Wen-Chih, York Shang-Hua Tsai, and Wang-Chien Lee. "Exploiting spatial and data correlations for approximate data collection in wireless sensor networks."*Knowledge Discovery from Sensor Data (Sensor-KDD 2008)*(2008): 111..

[12] Liu, Zhidan, et al. "Distributed spatial correlation-based clustering for approximate data collection in WSNs."*Advanced Information Networking and Applications (AINA), 2013 IEEE 27th International Conference on*. IEEE, 2013.

[13] Yuan, Fei, Yiju Zhan, and Yonghua Wang. "Data density correlation degree clustering method for data aggregation in WSN." *IEEE Sensors Journal* 14.4 (2014): 1089-1098.

[14] Dabirmoghaddam, Ali, Majid Ghaderi, and Carey Williamson. "Cluster-based correlated data gathering in wireless sensor networks." *Modeling, Analysis & Simulation of Computer and Telecommunication Systems (MASCOTS), 2010 IEEE International Symposium on*. IEEE, 2010.

[15] Rooshenas, Amirmohammad, et al. "Reducing the data transmission in wireless sensor networks using the principal component analysis." *Intelligent Sensors, Sensor Networks and Information Processing (ISSNIP), 2010 Sixth International Conference on*. IEEE, 2010

[16] Gastpar, Michael, Pier LuigiDragotti, and Martin Vetterli. "The distributed karhunen–loeve transform." *IEEE Transactions on Information Theory* 52.12 (2006): 5177-5196.

[17] T. Predrag and G.Agha. Maximal clique based distribution group formation for autonomous agent coalitions. In *Coalitions and Teams Workshop (W10), 3rd Int'l Joint Conf. on Agents and Mutli Agent Systems, 2004*

[18] Y. Xu, J. Heidemann and D. Estrin. Geography-informed energy conservation for ad hoc routing . In *MobiCom*, 2001.

[19] P. Krishna, N. H. Vaidya, M. Chatterjee, and D. K. Pradhan. A cluster-based approach for routing in dynamic networks. *SIGCOMM Computer Communication Review*, 27(2), 1997.

[20] Khalid, Zubair, and Salman Durrani. "Distance distributions in regular polygons." *IEEE Transactions on Vehicular Technology* 62.5 (2013): 2363-2368.

[21] Zhang, Honghai, and Jennifer C. Hou. "Maintaining sensing coverage and connectivity in large sensor networks." *Ad Hoc & Sensor Wireless Networks* 1.1-2 (2005): 89-124.

[22] Institute for Nuclear Theory, Seattle, WA,USA.(2004). Intel Lab Data [Online]. Available: http://db.csail.mit.edu/labdata/labdata.html
[23] Quer, Giorgio, et al. "Sensing, compression, and recovery for WSNs: Sparse signal modeling and monitoring framework." *IEEE Transactions on Wireless Communications* 11.10 (2012): 3447-3461.