

Efficient Data Gathering using Compressive Sensing with Cluster based Method in WSN

R. Rajalakshmi

Computer Science Department,
Velammal Engineering College,
Chennai, India

T. Subashini

Assistant Professor ,
Computer Science Department,
Velammal Engineering College,
Chennai, India

Abstract - Wireless Sensor Networking is one of the most prominent technologies that are implemented in almost all real time applications. Sensor Network is used for efficient data transmission where the sensor nodes are deployed in hostile environment. In this environment the sensor nodes sense the data and forward the data to the base station. When the data is being forwarded to the base station the data can be compressed at each cluster head and transmit to base station. Network lifetime, throughput and load balancing are the most important requirements in wireless sensor network. Here we have used an energy efficient clustering protocol LEACH and LEACH-C which is a time schedule based protocol. It reviews with the Compressed Sensing techniques that envisioned a useful technique to improve the performance and load balancing in wireless sensor networks. We first do with analytic model by optimizing the cluster size and continue with centralized algorithm for creating cluster based model. Simulation indicates that this method reduces data transmission extensively.

Keywords – Compressed Sensing, Leach-C protocol, Clustering, Data Collection.

I. INTRODUCTION

Wireless sensor networks consists of many spatially distributed sensor nodes which are used to monitor transmission of energy efficient data. The sensor nodes transmit data to the CH, while the nodes aggregate the data and transmit them to the base station either directly or through the intermediate communication with other CH nodes. The BS is the data processing unit for the data received from the sensor nodes. The Base Station is fixed at a place in a stationary manner which is far away from the all the sensor nodes. The function of each CH, is to perform common functions for all the nodes in the cluster, like aggregating the data before sending it to the BS. In some way, the CH is the sink for the cluster nodes, and the BS is the sink for the CH.

The advantage of cluster based environment is: 1) supporting network throughput and decreasing energy consumption through data aggregation 2) It can localize the route setup within the cluster and thus reduce the size of the routing table stored at the node. The main parameters included in clustering are: Number of clusters, Nodes and CH Mobility, Nodes types and roles, Cluster formation methodology, Cluster-head selection. Here hierarchical network routing is used which divides the network into energy

efficient cluster, LEACH - Low Energy Adaptive Clustering Hierarchy is belongs to hierarchical network routing protocol. A cluster head is responsible for conveying the message to base station as it aggregate and compress the data from each node before transmission. Cluster head often drains out from energy consumption which can be rectified by using LEACH protocol. LEACH-C is similar to LEACH as it based on location information of each sensor nodes which is useful for efficient increase in data transmission. As this protocol collects each sensor nodes location and form optimal clustering by computing energy level of each node. The node which is less than this average energy will be left out. The base station sends message to each node with CH id, which node matches this id that node becomes the cluster head. That is BS broadcast the information to all nodes about its position weather cluster head or node with its time schedule. In compressed sensing the aggregation is done by summing the coding vectors and determines the traffic path, finally the sink collects the aggregated vector as few samples, and the decoding algorithm is used to recover the samples. The encoding of each node will be done by distributed method, by performing matrix multiplication and summation whose cost is smaller than the non CS method. Compressed sensing operation requires each node in the individual sensor network to send k packets irrespective of what it has received, compared with nodes more work load for nodes away from sink and less load for nodes close to sink. Hybrid CS is the combination of non CS and Plain CS, it reduces the traffic load while preserving the integrity of the original data set. Plain CS consumes more energy than hybrid CS, the reason is plain CS overacts by forcing every node to be an aggregation, also the aggregation are in minority for optimal configuration. The hybrid CS reduces cost when compared to plain and non CS data aggregation. Here centralized clustering algorithm is used (Centralized- Low Energy Adaptive Cluster Hierarchy) for cluster formation. After CH election each cluster head data will be compressed by encoding method – similar data compression and transmit to base station. The base station will decode the data by collecting similar data's back. The simulation results describes that the method reduces transmission than non CS method. The Compressed method can be matrix multiplication or a Greedy algorithm. Hence the transmission is efficient in leach-c when compared to leach transmission protocol.

I. Overview of Nodes clustering in compressed method.

The sensor nodes are organized into clusters, each cluster has a cluster head, the cluster nodes transmit their data to their cluster head, The sensor nodes are uniformly distributed here. There are two levels of transmission in our clustering method they are, non-inter-cluster and inter cluster transmission. The size of clustering will be same for both. If we reduce the number of transmission automatically the energy consumption can be reduced. As cluster size increases, the number of Non-inter-cluster transmission can be increased. But, when decreasing the cluster size, the number of clusters and inter cluster transmission also can be increased. So optimal cluster size and distributed clustering method can be implemented. The sensor nodes transmission takes place by transmitting data from sensor nodes to CH, where the CS is not implemented in sensor nodes but CH data can be compressed and transmitted to base station. All sensor nodes have fixed transmission power and rate. The nodes know its individual location position that can be sending to base station. The task here is to optimize the cluster size by developing centralized algorithm to minimize the total number of transmission. The clustering algorithm is a LEACH based

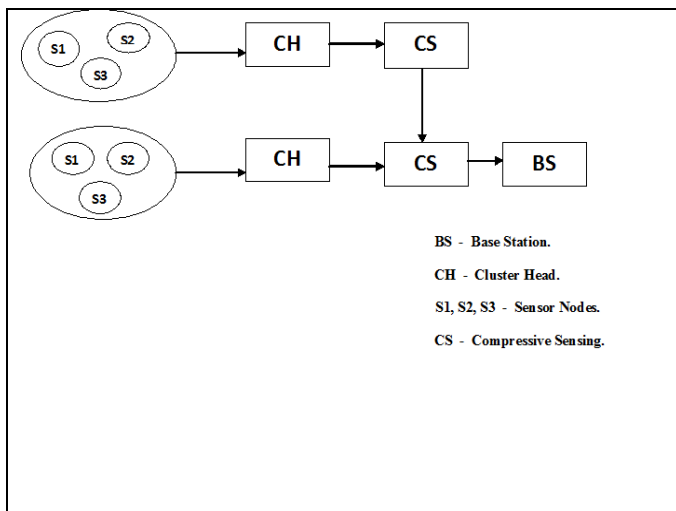


Fig 1: Compressed Sensing based Cluster

II. CENTRALIZED CLUSTERING ALGORITHM:

Protocol which is improved to LEACH-C algorithm called centralized algorithm here cluster head is static but in Leach the cluster head would be dynamic. LEACH is abbreviated as Low energy adaptive clustering hierarchy – it is the first hierarchical cluster-based routing LEACH protocol for cluster head creation. CH is responsible for creating and manipulating TIME (Time division Multiple Access) schedule and sending the data from nodes to BS, while other nodes are cluster members. The protocol is divided into two phases’ setup and steady state phase which is carried out as rounds. The setup phase is further divided into advertisement phase and cluster setup phase. The steady state phase consists of schedule creation and data transmission. In Leach each node decides independent other nodes weather it is going to be CH or not. In advertisement phase the CH inform other nodes that it

become a CH. The sensor nodes accept the advertisement and join the CH as members. After this phase the CH knows its member and its id, based on this message the CH prepares a time schedule and pick the code randomly and broadcast the time schedule table to its members. After this steady state begins, In steady state data transmission begins by nodes sending their data to individual CH based on time schedule when all the data’s send to CH ,the aggregated data can be sent to base station. Leach has no idea of total no cluster and placement of cluster formation. The Leach-C uses a centralized algorithm which is similar to Leach steady state phase. During setup phase the nodes send its information including position and residual energy to base station. The base station determines whether the energy load distributed equally to each node. The base station computes average energy of al nodes and determines which node has energy below this average one. The cluster head and its clusters formation , the base station sends information to all nodes that it obtained cluster head ID for each nodes, If the cluster head ID matches a node ID then that node becomes a cluster head. Otherwise the nodes computes its time schedule slot for data transmission and goes to sleep mode until it’s time to transmit

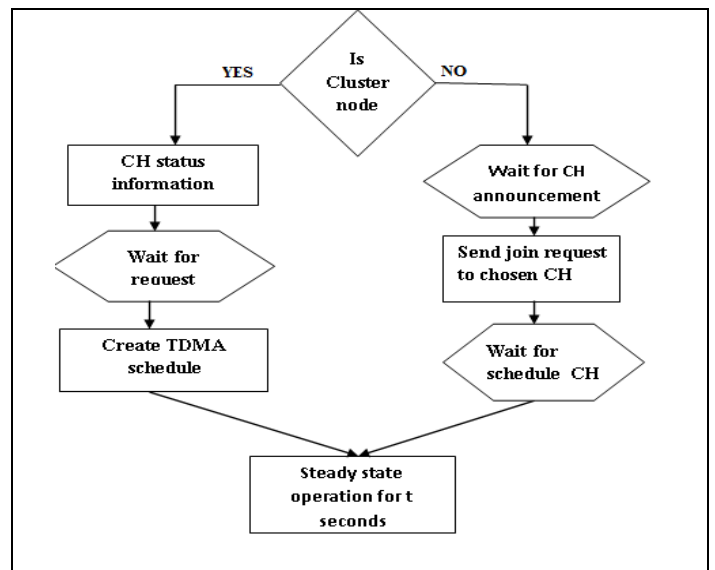


Fig 2: LEACH Protocol based Clustering

data. The tasks performed by leach are random formation of cluster heads and cluster formation. Data communication can be reduced by local compression. Low energy media access control can be performed. The data processing would be application specific. The task performed by setup phase is to organize whole network into non-inter-cluster formation. Advertisement of CH to its members. Transmission of schedule for setup phase. In steady state the process of the transmission within the different clusters of network. Transmission of data via different cluster heads. By using Leach protocol energy consumption can increase energy seven times better than direct communication. The clustering makes optimal cluster region formation so the number of CH can be scalable. The data transmission depends on inter-cluster, intra-cluster transmission. The cluster head becomes static after

computing optimal region of clustering, then transmission starts with reduced energy level..

III. LEACH PROTOCOL CLUSTERING:

Leach protocol provides throughput for aggregation in CH which makes less traffic load in the whole network. Single hop routing makes data transmission easier with less energy consumption. Distributive property is satisfied by using CH to distribute its message to its members.

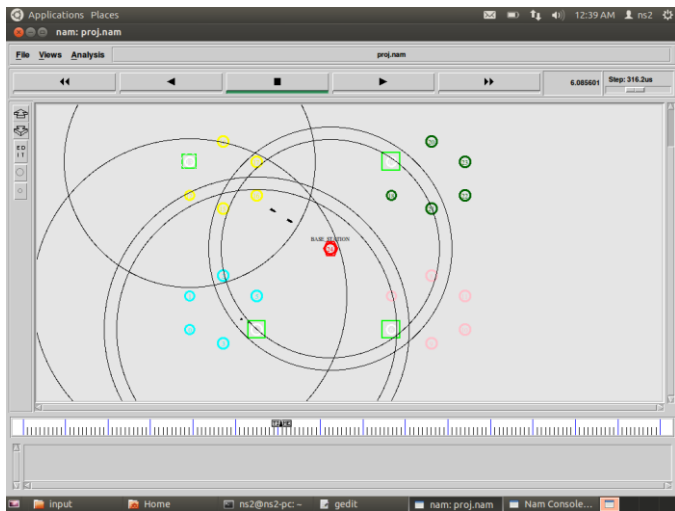


Fig3: LEACH protocol cluster formation

It increases network lifetime by CH distribution, data aggregation by CH and finally time scheduling were CH creates for its members. It is very simple and does not need nodes location for cluster formation. It is dynamic cluster based protocol where periodic data collection approach can be done for constant monitoring applications.

The disadvantage about Leach protocol is that it depends on CH rather than to its members in order to reach base station which can lead to cluster heads failure. It avoids overhead to the process for cluster head changes in each iteration for communication of information. It also avoids overhead in calculations for energy efficiency dynamic clustering in large scale networks. It has no inter - cluster transmission since CH communicates directly to the base station. The process requires high range of power in the network, for this only leach is not for large scale network. In Leach CH are not uniformly distributed that is CH can be located at the edge of the cluster. Here CH election is random processed which does not take energy consumption of the different nodes and this lead to re-electing of cluster CH as the same node is used for data iteration. It does not work well with the application that needs large area coverage for hopped inter-cluster communication. The animated file for this protocol is based on cluster head election method. The scenario is filled up with sensor nodes were the nodes are deployed with position of starting as well as ending location value. Here cluster node as well as member nodes generated on the basis of random number generation. Then the node which is energy efficient that node will become as cluster head on the basis of random number generation method. After the cluster head formation

the cluster head informs other node that is the member node to join in to the cluster. The cluster head starts time schedule for data transmission. After few seconds cluster head would be drained up. Then reelection would be done on next round, the protocol has higher transmission power for CH to base station transmission. So data transmits fo large area with reduced energy consumption.

IV. LEACH-C PROTOCOL CLUSTERING:

It uses a centralized clustering algorithm and it have same steady state protocol like leach. In this algorithm each node sends its energy level and position to base station. Based on this information the base station will determine different clusters along with CH and non CH nodes of each and every cluster. By this method base station would be able to produce better cluster with less energy transmission. Here the number of cluster head would be equal to predominant value where the Leach is dynamic and number of CH would vary. But in Leach-C the cluster head is static. The scenario file describes that the protocol is based on base station election for cluster formation. Here setup phase is different were the base station receives each node location based on which cluster election would be done. In advertisement phase nodes send location value to base station it elects by computing the average energy of all nodes. The node which has energy less than this will be dropped out. The base station sends id to all nodes which node matches its id it will become cluster head. The cluster formation is done for optimizing the cluster size.

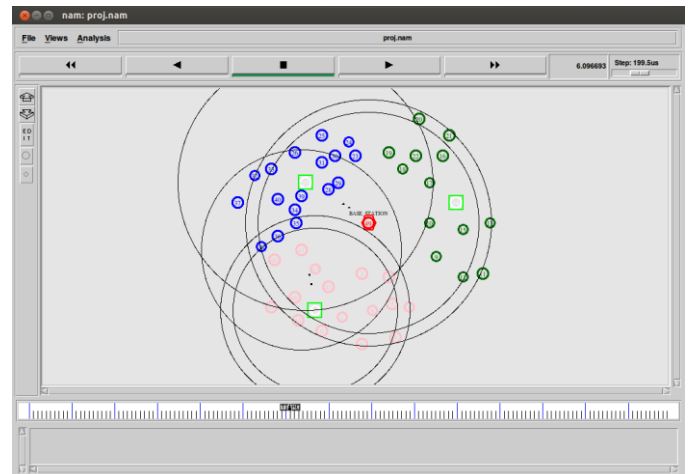


Fig 4: LEACH-C protocol cluster formation

If the cluster size increases the as cluster head total will be decreased. If the cluster size is decreased then cluster head would be increased hence the optimization of inter-cluster, non-inter-cluster should be done. The main advantage is that it is location dependent so optimization of cluster would be easier and also here the cluster head would be static. It is efficient with leach-c using compressed sensing when compared to leach for more than 90%. The received data is compressed by 30% so in leach-c based simulation will be taken on the basis of minimum and maximum value of compressed data.

V. COMPRESSED SENSING BASED DATA AGGREGATION:

The routing protocol cannot able to maintain transmission in huge traffic so compression technique is considered to be the best one before data transmission proceeds. The compressed sensing technique can reduce huge traffic and consume power consumption than other routing protocols. There are two types of data compression approach that is distributed and local compression techniques. The distributed spatial network has high spatial correlation in sensor nodes with wide networks. Here the coding can be jointly encoded at each nodes and decoded at the base station. So the signal can be compressed and sampled at each sensor nodes and decoded at the base station with high probability. The purpose of transmitting vast number of nodes is difficult task so a well-developed signal processed communication system is required for huge scale network. No data assumption is required in sampling hence it is a decentralized method. Different hopping method can reduce data transmission by projecting the sensor nodes by amplitude the data in cluster projection. The signal can be taken as $s = (s_1, s_2, \dots, s_{n-1}, s_n)$ be n number dimensional arrays and z is an array of vector co-efficient. The $s = z m$ but $v = d l$ where d is a $k \times k$ matrices. The value of s is compressed which includes normal value and location. The original data can be recovered at the base station by using gossip algorithm. The decoded signal is determined to same way as it is encoded and it is possible with projection of CS algorithm. The task of gathering compressed data is to read the signals and gather it in order to reduce it with huge network traffic and and prolong network lifetime, the sensed data can be relayed by computing the measurement and transmit the data

```

ns2@ns2-pc ~
File Edit View Search Terminal Help
-----
ch_(27)      2000    1241    37
-----
ch_(43)      1010     680    32
-----
ch_(43)       900     669    25
-----
ch_(14)      1560    1010    35
-----
ch_(7)       1010     680    32
-----
ch_(9)       790     504    36
-----
ch_(23)      1780    1120    37
-----
ch_(46)      1890    1296    31
-----
Starting Simulation...
-----
The Transmission saved -- 516
The Transmission efficiency increased by -- 34 %
-----
Compression Ratio available

```

Fig: 5 compressed data aggregation

VI. CLUSTER HEAD ELECTION:

In a clustering scheme the nodes are divided into groups, mostly based on geographical properties. Each group contains a single leader and several ordinary nodes. A cluster head normally serves as a local coordinator for its cluster, performing non inter-cluster transmission arrangement, data aggregation, forwarding, and so on. In the network model that

we are dealing with, we assume multiple-hop network, where gateway nodes connect between different clusters if there is not direct communication between the cluster heads. The clustering objectives are to minimize the total transmission power aggregated over the nodes in the selected path, and to balance the load among the nodes for prolonging the network lifetime. Cluster head selection is based on energy consumption of each node. The radius of cluster size is determined based on non-inter-cluster formation. The nodes close to the middle of the cluster becomes CH (cluster head) of cluster. Clustering is triggered at every c and no of seconds to form new CH. Probability of CH formation is based on $\text{Probability} = \text{Probability} * (\text{estimated energy} / \text{maximum energy})$. If there is no sensor node within the hops from central point then there is no CH for it. And this node joins other near cluster. After CH election CH sends an advertisement message to other sensor nodes to join in clusters. Clustering algorithm for data transmission will select x CH from set of nodes and perform x clusters. Distance from nodes to CH should be minimized. Repeat these steps until we get optimal cluster size. A back bone tree will be considered to connect all CH to sink by CS or non CS by broadcasting messages. Smaller the radius larger the cluster head and larger the radius larger the cluster head. If no sensor node falls in this area there would be no CH for that area. Finally we use a group of clusters to perform clustering. Considering the Poisson distribution of sensor nodes in the cluster formations makes the cluster sub divided and makes analytic method of data transmission for optimal cluster formation. It minimizes total transmission in aggregated data overall selected path. Balance the load to prolong network life time. It increases network scaled and reduces energy consumption. Optimal cluster selection is challenge thing. Grouping nodes into a hierarchical model is a popular method to achieve network scaled. Clustering usually localizes the routing setup within the cluster and therefore it reduces the routing overhead by each node and the topology maintenance overhead. Using clustering, the network appears smaller and more stable. The information, generated from nearer sensor nodes, is often redundant and highly correlated, so data aggregation by each cluster head conserves communication bandwidth as well. Moreover, the ability to use different power levels in inter-cluster and non inter cluster Communication reduces the inter and the collisions in the network resulting in a better throughput. Clustering is a challenging mission because optimal clusters selection is equivalent to the dominating set problem which is an NP-complete problem. Additionally, not only cluster selection is required, but also maintenance is essential. Because CH often lose more energy compared to regular nodes, it is necessary to perform re-clustering periodically in order to select energy-abundant nodes to serve as CH, thus distributing the load uniformly on all the nodes. Transmit aggregated data to the data sink is reducing number of nodes taking part in transmission. Useful energy consumption for transmission can be done by optimizing the cluster size. Scaled for large number of nodes is also possible by optimization. It reduces communication overhead for both single and multiple hops. In clustering algorithm sink will update the data for which the nodes are divided into cluster form. So a cluster head will be constructed for back bone tree. Here base station will

broadcast information to all sensor nodes and data aggregation will be performed respectively. The optimal cluster size will be constructed on the basis of n sensor nodes with clustering. Data will be transmitted by shortest path method via cluster head to base station. Non inter cluster transmission will be based on the sum of the nodes within CH. It can be calculated depending on k -median problem. Also the distance between two nodes is calculated by hopping method. Data aggregation compressed by CS at cluster head, which is send to base station. The sink will calculate the cluster by heuristic method. The clustering algorithm connects sensor nodes to CH. For each cluster chooses a new CH and repeats the iteration. The iteration makes the cluster optimal. The clustering makes use of optimal cluster size in order to make the data transfer which will be based on clustering, here used is hybrid Compressed sensing which makes reduces data transmission higher than other methods.

VII. SIMULATION RESULTS:

The network has been simulated in network simulator with data gathered spatially to perform the analyse. The simulation is done with randomly deployed nodes with size $100 * 100m$, and the nodes deployed varied from 50 to 100 with base station location x, y .

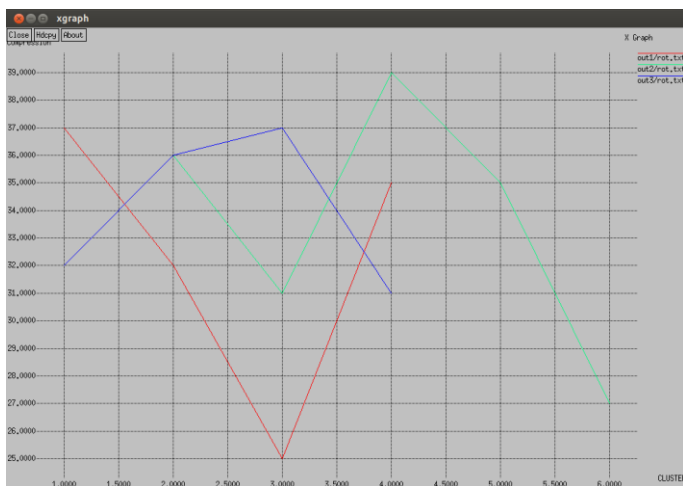


Fig 6: Graphical representation of CS

To measure energy consumption we depend on the parameters of energy and radio model. Energy consumption is divides into receiving and transmitting. The transmission message requires more energy than receiving the message. The ratio of transmission computes that the number of transmission in LEACH-C is transmits efficiently when compared to non CS. The performance of both the protocol is based on energy consumption and network lifetime. The average energy consumed is less than the leach based compressed with number of rounds. The network life time can be measured from the start of operation to the death of the first alive node. So our approach LEACH –C with Compressed Sensing has less consumption of energy than LEACH. Our method reduces the number transmission by about 60 percent than with clustering without CS method. It demonstrates a improvement than to worst case. The graphical representation

gives information about how the data is transmitted with the compressed method. As CS improves throughput with clustering method, the clustering is also developed with different protocol like leach-v,c-leach, MT-leach . hence the enhanced version is determined with compressed data of leach-c with leach.

VIII. CONCLUSION:

Here hybrid CS method is used for data transmission with LEACH –C based clustering. It is the low energy adaptive clustering hierarchy method that makes cluster formation with random CH election. The task performed by compressive sensing is to aggregate the similar data and transmit it. Our feature work is to implement compressed sensing in sensor nodes of sensor network which reduce energy higher than previous one. LEACH is an important protocol used for developing energy efficient clustering for numerous data transmission network, followed to this LEACH-C is implemented. From the simulation results we can draw conclusion that messages created by LEACH-C is base station based static CH is improved than LEACH. The compressed sensing method works well with LEACH-C protocol that gives increase throughput than non CS method. Our future enhancement is to perform compression to all sensor nodes by Leach-C method.

IX. REFERENCES:

- [1] R.Szewczyk, A. Mainwaring, J.Polastre, J.Anderson, and D.Culler, "An Analysis of a Large Scale Habitat Monitoring Application," Proc. ACM Second Int'l Conf. Embedded Networked Sensor Systems (SenSys'04), pp. 214-226, Nov. 2004.
- [2] E. Candes and M. Wakin, "An Introduction to Compressed Sampling," IEEE Signal Processing Magazine, vol.25, no.2, pp. 21-10, Mar.2008.
- [3] R.Baraniuk, "Compressed Sensing [Lecture Notes]," IEEE Signal Processing Magazine, vol.24, no.4, pp. 118-121, July 2007.
- [4] D. Donahue, "Compressed Sensing," IEEE Trans. Information Theory, vol.52, no.4, pp. 1289-1306, Apr. 2006.
- [5] J.Haupt, W. Bajwa, M. Rabbat and R. Nowak, "Compressed Sensing or Networked Data," IEEE Signal Processing Magazine, vol.25, no.2, pp. 92-101, Mar.2008
- [6] C. Luo, F. Wu, J. Sun, and C.W. Chen, "Compressed Data Gathering for Large-Scale Wireless Sensor Network," Proc. ACM MobiCom, pp. 145-156, Sept.2009.
- [7] S.Lee, S.Pattem, M.Sathiamoorthy, B.Krishnamachari, and A. Ortega, "Spatially-Localized Compressed Sensing and Routing in Multiple-Hop Sensor networks," Proc. Third Int'l Conf. GeoSensor networks (GSN'09), pp. 11-20, 2009.
- [8] C.Luo, F.Wu, J.Sun, and C.W. Chen, "Efficient Measurement Generation and Pervasive Sparsity for Compressed Data Gathering," IEEE Trans. Wireless Comm., vol.9, no.12, pp.3728-3738, Dec 2010.
- [9] J.Luo, L.Xiang, and C.Rosenberg, "Does Compressed Sensing Improve the Throughput of Sensor networks?" Proc. IEEE Int'l Conf. Comm. (ICC), pp. 1-6, May 2010
- [10] L.Xiang, J.Luo, and A.Vasilios, "Compressed Data Aggregations for Energy Efficient Sensor networks," Proc. IEEE Sensor, Mesh, and Ad Hoc Comm. and Network (SECON'11), pp.46- 54, June 2011.

- [11]. Fazel, M. Fazel, and M. Stojanovic, "Random Access Compressed Sensing for Energy-Efficient Underwater Sensor Networks," *IEEE J. Selected Areas Comm.*, vol.29, no.8, pp.1660-1670, Sept.2011.
- [12]. J.Wang, S.Tang, B.Yin, and X.-Y.Li, "Data Gathering in Sensor networks through Intelligent Compressed Sensing," *Proc. IEEE INFOCOM*, pp. 603-611, Mar.2012.
- [13]. B.Zhang, X.Cheng, N.Zhang, Y.Cui, Y.Li, and Q.Liang, "Sparse Target Counting and Localization in Sensor networks Based on Compressed Sensing," *Proc. IEEE INFOCOM*, pp.2255-2263, Apr 2011.
- [14]. E. Candes, J. Romberg, and T. Tao, "Robust Uncertainty Principles: Exact Signal Reconstruction from Highly Incomplete Frequency Information," *IEEE Trans. Information Theory*, vol. 52, no.2, pp. 489-509, Feb.2006.
- [15]. E. Candes and T. Tao, "Near-Optimal Signal Recovery from Random Projections: Universal Encoding Strategies?" *IEEE Trans. Information Theory*, vol.52, no.12, pp. 5406-5425, Dec.2006.
- [16]. J.Tropp and A.Gilbert, "Signal Recovery from Random Measurements via Orthogonal Matching Pursuit," *IEEE Trans. Information Theory*, vol.53, no.12, pp. 4655-4666, Dec.2007.
- [17]. M.Youssef, M.Younis "Overlapping Multiplehopping Clustering in sensor networks", *IEEE Trans Parallel and Distributed Systems*, Dec 2009.
- [18]. V.Arya, A.Meyerson, K.Mungala, "Local Search Heuristic for K-Median and Facility Location Problems", *Theory of computing* Pg 21-29 2009.
- [19]. C.perkin and E.Royer, "Adhoc On demand Distance Vector Routing", *IEEE Workshop on mobile computings*, pp 90-10 Feb 1999.
- [20]. D.B.Johnson and D.S Maltz, "Dynamic source network in Adhoc Sensor networks", *A survey on computer network* -2002.