KNN based Efficient Inter Cluster Routing Protocol for WSN

Vandana Srivastava,
Electronics and Communication,
BBDU, Lucknow

Abstract: In this work, we look at routing protocols, which can have significant impact on the overall energy dissipation of these networks. Based on our findings that the conventional routing protocols of direct transmission, minimum-transmission-energy, multi hop routing, and static clustering may not be optimal for sensor networks, we propose KNN based routing protocol , a inter clustering-based protocol that utilizes randomized rotation of local cluster base stations (cluster-heads) to evenly distribute the energy load among the sensors in the network. It uses a non-localized coordination to enable scalability and robustness for dynamic networks, and incorporates data fusion into the routing protocol to reduce the amount of information that must be transmitted to the base station. Simulations show that it can achieve as much higher factor of reduction in energy dissipation compared with conventional routing protocols. In addition, it is able to distribute energy dissipation evenly throughout the sensors, doubling the useful system lifetime for the networks we simulated.

Keywords: Clustering, KNN, Routing Protocol, WSN

1. Introduction:

In hierarchical routing, sensor nodes are groupped into clusters. Every cluster has a frontrunner, a node dedicated to communication with more computation and power resources. Every sensor node then sends facts simplest to the cluster chief unburdening nodes from routing problems which can be moved to cluster leaders. [2] Many functions of WSN (including time synchronization, duty-cycles) have to be also contemplated in the base station implementation. Therefore, the bottom station developer must recognize the basic ideas of WSN or at least the ones associated with developing the base station. Hierarchical strategies are typically utilized in stressed out network for scalability [11]. For wireless networks, a hierarchical clustering and routing scheme based upon bodily area management was currently proposed in [6, 12]. This scheme, but, creates implementation problems that are probably complicated to solve. First, it does allocate Cluster IDs dynamically. This allocation ought to be specific - not an easy assignment in multi-hop cell environment, in which the hierarchical topology ought to be frequently reconfigured. Second, every cluster can dynamically merge and cut up, primarily based at the wide variety of nodes inside the cluster. Frequent cluster modifications may also degrade the community overall performance appreciably. In a large, cell network the trouble of locating customers and services by means of their names isn't always a trivial one. In a stressed out Internet the DNS affords a mapping among symbolic names and community addresses. The community cope with is then processed by the routing tables and leads immediately to the vacation spot. In stressed out networks with mobile radio extensions, Mobile IP became advanced to deal with the ultimate hop indirection, from Home Agent to mobile user. In multihop wi-fi networks there's no fixed Home Agent. We endorse the Wireless Hierarchical Routing Protocol (WHIRL) to attack this hassle with a ”multihop extension” of the Mobile IP concept. We will distinguish among the ”physical” routing hierarchy, dictated by means of the geographical relationship among nodes, and the ”logical” hierarchy of subnets corresponding to participants within the equal institution (e.G., tanks in the battlefield, or travelling salesman of the equal organisation). We will hold music of logical subnets the usage of a DNS hierarchy geared toward lowering manage traffic O/H. Physical MAC layer clustering [7, 4] offers the primary level of an efficient ”physical” routing hierarchy.

In WHIRL, the complete community is divided into logical subnets. Each subnet has one primary Home Agent (HA). It will have several secondary HAs from which, within the case of primary HA failure, a new primary HA may be selected . Each node has a completely unique identifier NODEID. The deal with of the node includes components: logical. This is used to discover the logical subnet to which each node belongs and the is utilized in physical routing. In our look at, we use the Link State (LS) physical routing scheme which is on the top of MAC layer clustering defined in [7, 4] as the bodily routing infrastructure for WHIRL. However, the idea of WHIRL can be built upon any routing scheme the usage of because the bodily routing address. The key obligation of the HA is to maintain the physical clustering statistics of its logical subnet participants. HA additionally desires updates its personal clustering facts to all of the cluster heads. There are two levels within the WHIRL. The packet is routed first from the source to destination HA. Then it is routed from the destination HA to the final vacation spot. The header of the packet carries the of the destination cluster head. Initially, the is about to. The source sends the packet to its cluster head. The cluster head will look up its HA clustering desk. The cluster head will set the to be the of the vacation spot cluster head of the destination HA in the packet header in keeping with the desk. If the vacation spot HA has more than one cluster heads, it'll select the one that has the minimum distance. All the intermediate gateways and cluster heads will direction the packet according to the inside the packet header using the physical routing scheme. Once the vacation spot cluster head receives the packet, it's going to ship the packet to the HA. The HA scans its subnet member clustering table, unearths out the cluster head for the vacation spot node and units the with it. The HA will then ship the packet to its cluster head and the packet may be on the adventure to its very last vacation spot cluster head. The vacation spot cluster head will pass the the packet to destination node.

2. Related Work:

The dynamic nature of wi-fi sensor networks (WSNs) and severa possible cluster configurations make attempting to find an most desirable community shape at the- fly an open undertaking. To cope with this problem, **Xiaohui Yuan, (2017) [1]** proposed a genetic algorithmbased, self-organizing network clustering (GASONeC) technique that offers a framework to dynamically optimize wireless sensor node clusters. In GASONeC, the residual energy, the anticipated energy expenditure, the distance to the base station, and the quantity of nodes in the location are employed in search for an top of the line, dynamic community structure. Balancing those elements is the important thing of organizing nodes into suitable clusters and designating a surrogate node as cluster head. Compared to the cutting-edge methods, GASONeC greatly extends the community lifestyles and the improvement as much as forty three.Forty four %. The node density substantially influences the network toughness. Due to the improved distance between nodes, the community existence is normally shortened. In addition, while the base station is placed far from the sensor field, it is desired that greater clusters are fashioned to preserve strength.

A cluster-based totally version is leading in wi-fi sensor network because of its ability to reduce power consumption. However, handling the nodes inside the cluster in a dynamic environment is an open assignment. Selecting the cluster heads (CHs) is a cumbersome system that significantly impacts the community overall performance. Although there are several research that propose CH selection strategies, maximum of them are not suitable for a dynamic clustering environment. To keep away from this hassle, several strategies had been proposed by way of **Mohamed Elhoseny, (2017), [2]** primarily based on smart algorithms along with fuzzy good judgment, genetic set of rules (GA), and neural networks. However, these algorithms paintings better inside a unmarried-hop clustering version framework, and the network lifetime constitutes a huge trouble in case of multi-hop clustering environments. This paper introduces a new CH selection method based on GA for both unmarried-hop and the multi-hop cluster fashions. The proposed method is designed to satisfy the requirements of dynamic environments by way of electing the CH primarily based on six major functions, namely, (1) the closing energy, (2) the ate up power, (3) the variety of nearby neighbors, (four) the energy aware distance, (5) the node vulnerability, and (6) the diploma of mobility. We shall see how the corresponding consequences display that the proposed set of rules greatly extends the network lifetime.

NitinMittal (2016), [3] worked on nature-inspired algorithms are becoming popular among researchers due to their simplicity and versatility. The nature-stimulated metaheuristic algorithms are analysed in terms in their key capabilities like their diversity and version, exploration and exploitation, and sights and diffusion mechanisms. The fulfillment and challenges regarding these algorithms are based totally on their parameter tuning and parameter manipulate. A relatively new set of rules prompted by using the social hierarchy and hunting conduct of gray wolves is Grey Wolf Optimizer (GWO), that's a completely successful algorithm for fixing actual mechanical and optical engineering troubles. In the authentic GWO, 1/2 of the iterations are committed to exploration and the opposite half of are committed to exploitation, overlooking the impact of proper balance between these to assure an correct approximation of global most desirable. To triumph over this shortcoming, a changed GWO (mGWO) is proposed, which focuses on proper balance between exploration and exploitation that results in an top-rated performance of the set of rules. Simulations based on benchmark troubles and WSN clustering trouble reveal the effectiveness, efficiency, and stability ofmGWO as compared with the basicGWO and some well-known algorithms.

Vehicular Ad hoc NETworks (VANETs) are a primary element currently used within the development of Intelligent Transportation Systems (ITSs). VANETs have a highly dynamic and portioned network topology because of the steady and speedy motion of cars. Currently, clustering algorithms are extensively used as the manipulate schemes to make VANET topology much less dynamic for Medium Access Control (MAC), routing and security protocols. An efficient clustering algorithm must recall all of the essential records associated with node mobility. In this work, **Mohamed Hadded,(2015) [4]** proposed an Adaptive Weighted Clustering Protocol (AWCP), mainly designed for vehicular networks, which takes the dual carriageway ID, route of automobiles, function, pace and the wide variety of neighboring motors under consideration a good way to decorate the stableness of the community topology. However, the a couple of manipulate parameters of our AWCP, make parameter tuning a nontrivial problem. In order to optimize the protocol, we define a multi-objective problem whose inputs are the AWCP's parameters and whose targets are: imparting solid cluster structures, maximizing statistics transport price, and lowering the clustering overhead. We cope with this multi-goal hassle with the Nondominated Sorted Genetic Algorithm version 2 (NSGA-II). We evaluate and examine its performance with other multi-objective optimization techniques: Multi-goal Particle Swarm Optimization (MOPSO) and Multi-objective Differential Evolution (MODE). The experiments reveal that NSGA-II improves the effects of MOPSO and MODE in terms of spacing, spread, ratio of non-dominated answers, and inverse generational distance, which are the performance metrics used for comparison. Because of the swiftly changing topology and the dearth of infrastructure, it is very challenging to installation clustering methods in vehicular networks.

Every form of network, be it stressed or wi-fi, could be prompted with the aid of several key elements for its efficient functioning. Routing difficulty, applicable to all styles of networks, is one a number of the several such key factors. Wireless Sensor Networks (WSN) has not been exception to this. Moreover, such issues are very essential due to excessive resource constraints like efficient strength usage, lifetime of community, and drastic environmental situations in WSNs. Neither hop-through-hop or neither direct attain potential is viable in case of WSNs. In this regard, many routing protocols were proposed by way of **Geetha. V.(2012) [5]** to optimize the efficiency of WSNs amidst of above cited intense aid constraints. Out of these, clustering algorithms have gained extra importance, in increasing the life time of the WSN, due to their approach in cluster head selection and information aggregation.

3. Methodology:

The conventional KNN text category set of rules used all training samples for class, so it had a massive wide variety of schooling samples and a high diploma of calculation complexity, and it additionally didn't mirror the exclusive significance of different samples. With the fast improvement of net, a big range of textual content information begin to exist with the form of pc-readable and increase exponentially. The records and useful resource of internet take on the individual of big.

In order to correctly manipulate and utilize this huge amount of report statistics, textual content mining and content-based facts retrieval have regularly come to be the hotspot studies area in the world. Text class is an critical basis for facts retrieval and textual content mining, the main project is assigning textual content file to 1 or extra predefined categories according its content material and the labeled training samples [1]. Text class has been used substantially. For example, a few government departments and businesses made use of textual content classification for e-mail filtering. This type of e-mail classifiers cannot simplest filter out junk emails, but also distribute emails to the corresponding departments in keeping with the content. Text category generation also widely utilized in internet search engines like google, that may filter out the message that customers don't challenge approximately and deliver their fascinated content material. So we are able to take a conclusion that in the manner of information offerings, text category is a essential and vital technique, it may help users to prepare and access to information and it has very important research fee. Study on text classification abroad dated lower back to the past due Nineteen Fifties, H. P. Luhn had executed some ground-breaking studies paintings and proposed the methodology of the usage of word frequency for textual content automated type. In 1960, Maron posted the primary paper on text computerized classification, and then, a big quantity of scholars got fruitful research on this area. So far, the text type technology in foreign united states of america has been carried out in lots of fields, such as e mail filtering, digital conferences and statistics retrieval. There has additionally evolved some of extraordinarily mature software, including Intelligent Miner For Text evolved by using IBM, which can classify, cluster and get summary for the industrial documents; Net Owl Extractor evolved by means of SAR applied the feature of text clustering, category and electronic mail filtering; and additionally Insight Discoverer Categorizer advanced with the aid of TENIS can clear out junk emails and knowledge control for the commercial documents [2].

Study on textual content class in China had a past due start. Professor Hou Han-Qing had executed a good deal studies on text mining and delivered the idea of foreign laptop control tables, pc statistics retrieval, and computer textual content automatic classification in 1981. Afterwards, many researchers and institutions have all started to look at the text class. Currently, the research on textual content class has been made a whole lot of improvement, and the commonplace algorithms

for text class consist of K nearest neighbor set of rules (KNN), Bayes algorithm, Support Vector Machine algorithm (SVM), selection tree algorithms, neural community algorithm (Nnet), Boosting algorithm, etc [2]. KNN is one of the maximum popular and good sized amongst those, but it nonetheless has many defects, together with wonderful calculation complexity, no distinction among function phrases, does not keep in mind the associations among the key phrases and so forth. In order to keep away from those defects, many researchers had proposed a few enhancements. On account of the truth that the conventional method lacked of attention of institutions among the keywords, literature [3] proposed an stepped forward KNN approach which carried out vector-combination technology to extract the related discriminating phrases consistent with the CHI statistic distribution. The era of vector combination can reduce the size of the textual content function vector and improve the accuracy effectively, but it can't spotlight the important thing phrases that have more contribution to type. Literature [4] proposed a fast KNN set of rules named FKNN directed to the shortcomings of incredible calculation, but it could't improve the accuracy. In order to reduce the high calculation complexity, this paper used clustering technique and selected the cluster facilities because the representative factors which made the education sets grow to be smaller, and for overcoming the defect of no distinction among feature phrases.

3.1 KNN Text Classification Algorithm

KNN is one of the most vital non-parameter algorithms in pattern recognition subject [11] and it's a supervised studying predictable type set of rules. The classification guidelines of KNN are generated by using the schooling samples themselves with none extra data. KNN category set of rules predicts the take a look at pattern's class in line with the K schooling samples which are the closest neighbors to the test pattern, and choose it to that class which has the largest category probability. The technique of KNN set of rules to categorise file X is [12]:

Suppose that there are j training categories as C1 ,C2 ,C3.... Cj , and the sum of the training samples is N . After pre-processing for every document, all of them become m-size characteristic vector.

1. Make record X to be the same text feature vector shape (X1 ,X2 ,X3.... Xm) as all training samples.

2. Calculate the similarities among all education samples and record X . Taking the ith file (di1, di2,.....Dim) as an example, the similarity SIM (X, di) is as following

$$SIM(X,d_i) = \frac{\sum_{j=1}^{m} x_j.d_{ij}}{\sqrt{(\sum_{j=1}^{m} x_j)^2}.\sqrt{\sum_{j=1}^{m} d_{ij})^2}}$$

3) Choose k samples which are larger from N similarities of SIM(X ,di),(i= 1,2,.... ,N), and treat them as a KNN collection of X . Then, calculate the probability of X belong to each category respectively with the following formula.

$$P(X,C_j) = \sum_{d_j \in KNN} SIM(X,d_i).y(d_i,C_j)$$

Where y(di ,Cj) is a category attribute function, which satisfied

$$y(d_i, C_j) = \begin{cases} 1, d_i \in C_j \\ 0, d_j \notin C_j \end{cases}$$

4. Judge document X to be the category which has the largest P(X, Cj,).

The traditional KNN text classification has 3 defects [13]: 1) Great calculation complexity. When using traditional KNN type, if you want to locate the K nearest neighbor samples for the given take a look at pattern, it have to be calculated with all of the similarities between the training samples, as the size of the text vector is typically very high, so its has exquisite calculation complexity on this method which made the efficiency of textual content type very low. Generally speaking, there are 3 techniques to lessen the complexity of KNN set of rules: reducing the size of vector text [4]; the usage of smaller records units; the usage of stepped forward algorithm that could accelerate to find out the K nearest neighbor samples [5]; 2) Depending on training set. KNN algorithm does not use extra data to explain the classification rules, but the classifier are generated by way of the self education samples, this made the set of rules depend on training set excessively, for instance, it want to re-calculated when there may be a small exchange on education set; 3) No weight distinction among samples. As system (2), the category attribute characteristic has mentioned that the conventional KNN algorithm dealt with all education samples similarly, and there is no distinction among the samples, so it don't suit the actual phenomenon which samples have choppy distribution usually.

4. Result and Discussion:

In this paper for the algorithm development we have used MATLAB 2015 software and developed an algorithm using programming. We have considered a simulated network of 100 randomly distributed nodes in a sensor field of 100* 100 area. The location of base station is fixed and base station ID is 101. The simulation parameters used are shown in Table 1

Table 1: Simulation Parameters

No of nodes	100
Network area	100 * 100
Channel Type	Wireless channel
Source node	100 sensor nodes
Antenna Model	Antenna/Omni Antenna
Interface Queue Type	Queue/Drop Tail/PriQueue
Initial Energy of sensor nodes	.5j
Time for each round	5 sec
Transmission power	50 pj
Receiving power	10 pj
Transmission Range	40m

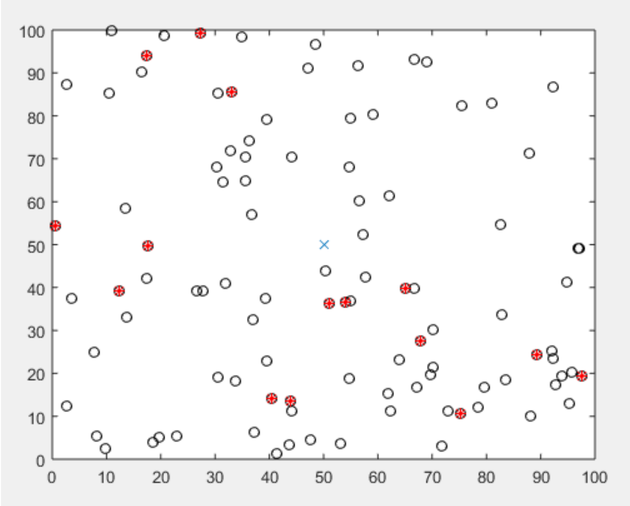

Fig 1: Network layout for developed WSN model

Figure 1 shows the network layout of developed WSN model having each node is distributed randomly over the complete area. Base station is placed at (x, y)=(50 , 50) position. Each cluster head is selected in every round. They are shown by '*' mark in above figure 1.These cluster heads are considered under K nearest neighborhood algorithm. Where k is the number of nearest neighbors here k is taken to be 4.

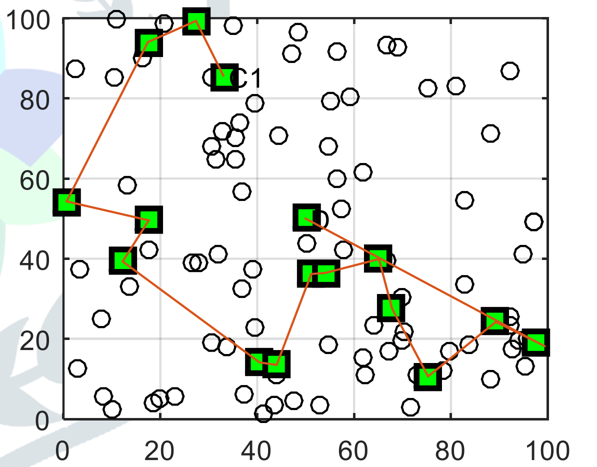

Fig 2: Routing from cluster head C1 to base station.

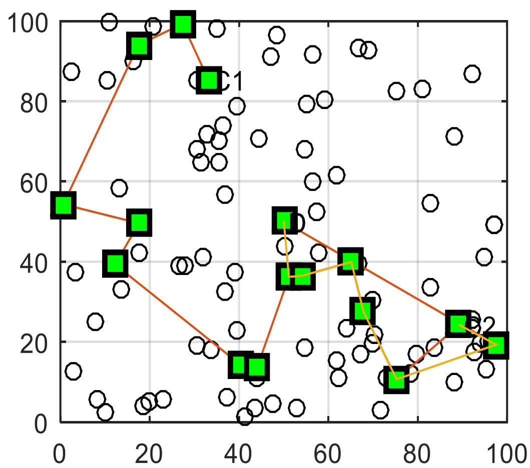

Fig 3: Routing from cluster head C2 to base station.

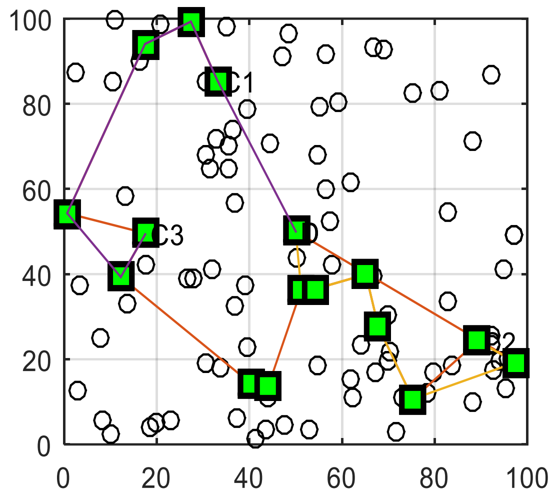

Fig 4: Routing from cluster head C3 to base station.

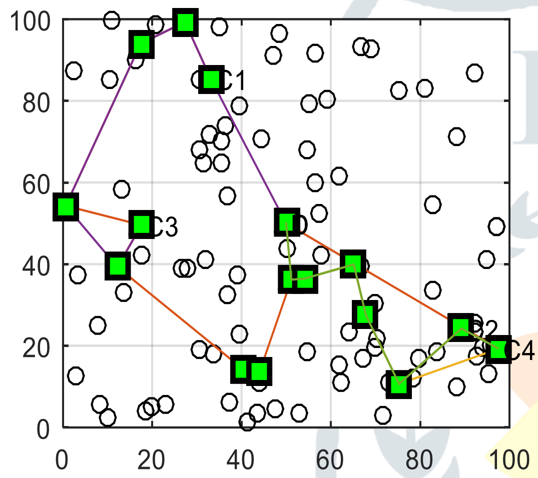

Fig 5: Routing from cluster head C4 to base station.

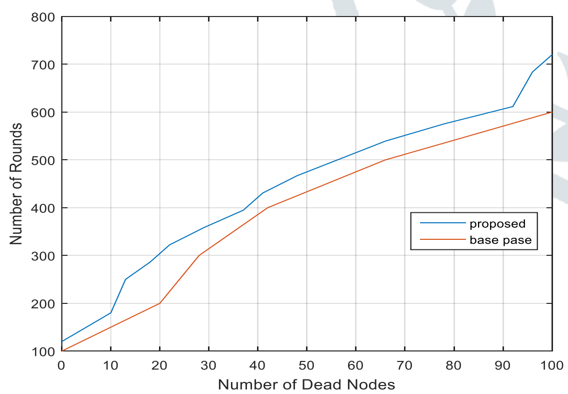

Fig 6: Number of dead nodes w.r.t. number of rounds.

5. Conclusion:

In this paper, we develop an energy efficient routing algorithm based on minimum energy route development for wireless sensor networks. We have used multi-hop architecture in between CH to base station for data transmission. The first step of routing is clustering, where the whole network is divided into some disjoint clusters having some member nodes and a special node known as cluster head. Each cluster head for a particular cluster performs data aggregation and necessary computation. The cluster-heads are responsible for routing received data to the base station. We evaluated the neighboring clusters ids using KNN technique and we calculate the energy required to transmit a message to all the neighbors from source cluster to the destination cluster. Comparing all possible potential routes we select the minimum energy route to transmit the data towards base station. Simulation result shows significant improvement of performance of our scheme.

References:

[1] Xiaohui Yuan, "A Genetic Algorithm-Based, Dynamic Clustering Method Towards Improved WSN Longevity," 9 September 2015 / Revised: 25 March 2016 / Accepted: 7 April 2016 _ Springer Science+Business Media New York 2017

[2] Mohamed Elhoseny, "Dynamic Multi-hop Clustering in a Wireless Sensor Network: Performance Improvement," Springer Science+Business Media New York 2017

[3] NitinMittal, "Modified Grey Wolf Optimizer for Global Engineering Optimization," Journal Applied Computational Intelligence and Soft Computing archive Volume 2016, March 2016 Article No. 8 Hindawi Publishing Corp. New York, NY, United States

[4] Mohamed Hadded, "A Multi-Objectif Genetic Algorithm-Based Adaptive Weighted Clustering Protocol in VANET," 2015 IEEE Congress on Evolutionary Computation (CEC)

[5] Geetha, "Clustering in Wireless Sensor Networks: Performance Comparison of LEACH & LEACH-C Protocols Using NS2," Geetha. V. et al. / Procedia Technology 4 (2012) 163 – 170

[6] Fuad Bajaber, "Adaptive decentralized re-clustering protocol for wireless sensor Networks," Journal of Computer and System Sciences 77 (2011) 282–292

[7] Jorge Tavares, "Application of Wireless Sensor Networks to Automobiles," MEASUREMENT SCIENCE REVIEW, Volume 8, Section 3, No. 3, 2008

[8] Ramesh Rajagopalan, "Data aggregation techniques in sensor networks: A Survey," (2006). Electrical Engineering and Computer Science. 22.

[9] Goldberg, "Genetic Algorithms and Machine Learning," Machine Learning 3: 95-99, 1988 © 1988 Kluwer Academic Publishers - Manufactured in The Netherlands

[10] Wendi B. Heinzelman, "An Application-Specific Protocol Architecture for Wireless Microsensor Networks," IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, VOL. 1, NO. 4, OCTOBER 2002

[11] I.F. Akyildiz, "Wireless sensor networks: a survey," Computer Networks 38 (2002) 393–422

[12] Wendi Beth Heinzelman, "Application- Specific Protocol Architectures for Wireless Networks," Massachusetts Institute of Technology Date Issued: 2000

[13] S. Bandyopadhyay, E.J. Coyle, An energy efficient hierarchical clustering algorithm for wireless sensor networks, in: Proceeding of INFOCOM 2003, April 2003.

[14] V. Mhatre, C. Rosenberg, Design guidelines for wireless sensor networks: communication, clustering and aggregation, Ad Hoc Network Journal 2 (1) (2004) 45–63.

[15] M. Ye, C. Li, G. Chen, J. Wu, EECS: an energy efficient cluster scheme in wireless sensor networks, in: IEEE International Workshop on Strategies for Energy Efficiency in Ad Hoc and Sensor Networks (IEEE IWSEEASN2005), Phoenix, Arizona, April 7–9, 2005.