Use of Entropy to Control Coverage and Energy Dissipation in Wireless Sensor Networks

R.Zilan, Y.M.Erten, Member, IEEE

Abstract— **In this study we proposed a scheme to achieve adequate coverage while maximizing network lifetime in a sensor network. The proposed approach is based on controlling the number of active sensors using information theory. It assumes that the area of coverage is divided into grids and uses a threshold probability to decide whether the sensors should be ON or OFF. Based on the number of ON sensors in each grid we define an entropy value which we use to adjust the threshold probability for the next epoch. The control mechanism is devised to maximize entropy while keeping the number of ON sensors at a minimum. This approach improved the coverage and network lifetime.**

I. INTRODUCTION

There are many definitions of quality of service (QoS) in sensor networks and Iyer and Kleinrock [1] define it as the optimum number of sensors sending information towards the information-collecting sinks, typically base stations. In the light of this, we have to have enough sensors sending packets to have adequate coverage of the area of interest. On the other hand, it is also desirable to maximize network lifetime.

This study examines these two aspects by focusing on a method to obtain optimum number of sensors transmitting in each epoch. As pointed out in [2], coverage is a measure of the QoS of the sensing function. The goal is to have each location in the physical space of interest within the sensing range of at least one sensor. The effective range of the sensors attached to a sensor node defines the coverage area of a sensor node. Coverage may be sparse or dense but ideally we should have adequate number of sensors transmitting for a robust system while receiving as little redundant data as possible.

Having redundant information from sensors which need not be ON does not improve the information quality but affects the network life adversely and causes the sensors to die early. This problem may be remedied if we can make sure that only adequate number of sensors is turned ON during each epoch. To adjust the number of ON sensors without considering coverage does not really offer improved QoS as the active sensors may be clustered in one part of the network hence the information received from the area of interest may not be complete. We therefore used entropy function as a measure of homogeneity of the received information. We divided the area to be covered into grids and used the ratio of the number of ON sensors in each grid to the total number of ON sensors in all the grids as the probability value to be used in the entropy calculations. As this probability is closer to 1/N; where N is the number of grids, we would obtain the maximum entropy value and the most desired coverage for efficient energy usage. We introduced a threshold probability of being ON transmitted by the base station and we adjusted this probability to make sure that we have the required coverage, i.e. at least one sensor active in each grid. We used MATLAB in our simulations and presented the results of our proposed approach in the proceeding sections. In our simulations we assumed that each sensor in a grid is capable of covering the whole grid area and that each sensor node can communicate with the base station.

The paper is organized as follows: section 2 describes the previous work on the subject and simulation setup used in this study is given in section 3 followed by the results in section 4. Conclusions are included in section 5.

II. PREVIOUS STUDIES

As mentioned before, Iyer and Kleinrock [1], defined QoS as the number of sensors sending information and presented an algorithm using the Gur Game paradigm that allowed the base station to specify the optimal number of sensors which is ON at a particular epoch. As it uses broadcast mechanisms, however, the battery life optimization is not very efficient. Frolik [3] presented two new techniques that maintain QoS in a wireless sensor network and Kay and Frolik [4] investigated spatial resolution in wireless sensor networks as a measure of quality of service. This work uses the ACK messages to specify the probability of sensors turning ON hence offers better network lifetime.

Wang et al. [5] also proposed an entropy-based sensor selection heuristic for selecting the desired sensor. They tested the heuristic using simulations and it has been shown that the approach selects the desired sensor with nearly the maximal mutual information.

Johansson and Sternad [6] used information theory in their work where they formulated and solved the problem of resource allocation over a given time with uncertain demands and uncertain capacities of the available resources. They presented a method for allocating the channels to maximize the expected system throughput. The framework includes quality of- service requirements, e.g., minimum-rate constraints, as well as priorities represented by a user-specific cost per transmitted bit. Optimal solutions are found by using the maximum entropy principle and elementary

---

Manuscript received June 9, 2007. M. Erten is with the Department of Computer Engineering, TOBB Economics and Technology University, Turkey (phone: 90-312 292-4071; fax: 90-312-292- 40 91; e-mail: erten@ etu.edu.tr).

R Zilan, was with TOBB Economics and Technology University, Turkey. She is now with the Department of Computer Architectures, UPC, Barcelona, Spain (e-mail: rzilan@etu.edu.tr).

1-4244-1364-8/07/$25.00 ©2007 IEEE

probability theory.

Ammari and Das [7] proposed an energy aware protocol called WEDAS (Weighted Entropy Data Dissemination) for sending data to the mobile sink in WSNs using an information theoretic approach. The protocol is based on quantifying the uncertainty of the position of a mobile sink and the remaining energy of static sensors and aims to select the most appropriate sensors.

The grid based approach we adopted in our study has also been used in the literature by different authors. Nama and Mandayam [8] proposed an information field model that partitions the observation space into a grid, with independent information being generated at each point in the grid. Given this model, they found the optimal node distribution over the field that maximizes the network information capacity or the total information gathered over the lifetime of the network.

Chen et al. [9] proposed a grid-based working node (WN) selection approach for wireless sensor networks. They aimed at identifying minimum number of sensors among the ones deployed which they called WNs. The remaining sensors are turned off to save power. The coverage of the sensors is represented by a number of sample points which are the intersection points of the established grid.

Akl and Sawant [10] designed a grid-based coordinated routing protocol for sensor networks. They described how the coordinators are elected from the given set of nodes and also determined the upper bound on the grid size to ensure connectivity between the coordinators. Finally, they incorporated load balancing in their protocol to distribute routing load over all the nodes and analyzed the effects of varying transmit power, receiver sensitivity, and grid size on the network lifetime.

Tan [11] proposed a hexagon-grid algorithm to schedule the activities of sensor nodes in a WSN. After that the respective analysis is carried out.

In our study we assumed that the base station has a means of knowing the location i.e. the grid of the sensor node. Location information is one of the most important subjects for wireless sensor networks since the accuracy of the collected data can noticeably be affected by the position of the event of interest. Stoleru and Stankovic [12] proposed a location estimation scheme that uses a probabilistic approach for estimating the location of a node in a sensor network.

The Global Positioning System (GPS) is very common in military, civil and industrial applications. Since providing an accurate localization is a remaining problem in real applications, Stoleru et al. [13] proposed a solution which is described as .Walking GPS.. The study proposes a design implementation and real world estimation for the localization of sensor nodes.

To detect the existence and position of objects, Radar systems [14] which is based on transmitted and received radio waves, were developed. It records and uses measured received signals to evaluate the distance or the position of the objects. Bahl and Padmanabhan proposed a radar system which is a location support system for in building applications [14]. Another in building application is called the Cricket [15]. Instead of tracking locations of the users, this method helps devices to learn their places. And then devices send the information to a particular user. This process does not depend on any centralized management. The results of this study show that several location-dependent applications can be improved with little effort.

Bulusu et al. propose an RF-based and receiver based localization technique for a small number of reference points which is called .Centroid. [16]. They studied classification of the design space and proposed a method for coarse-grained localization based on an idealized radio model, and demonstrates its validity and applicability in outdoor unconstrained environments. Even though it has a simple implementation, application of this study is not feasible since the model assumes that some of the nodes are more powerful than the others. APIT [17] which is a novel, range-free, localization scheme, on the other hand, is also based on the same principle, but does not require more powerful nodes. He et al. [17] showed that APIT scheme runs best when an irregular radio pattern and random node position are considered, and low communication overhead is preferred.

Another range-free, robust and adaptive localization algorithm is presented by Nagpal et al [18]. This study does not require more powerful nodes either. The algorithm is used for creating a coordination system without the uses of global control, globally accessible beacon signals, or accurate approximation of inter sensor distances. Niculescu et al. at Rutgers [19] proposed a localized, distributed, hop by hop positioning algorithm called APS (Ad hoc Positioning System). This method extends the capabilities of GPS to non-GPS enabled nodes in a hop by hop fashion in an ad hoc network.

Zhang et.al. [20] presented the Anchor Location Service (ALS) protocol, a grid-based approach to assist sources for the localization of sinks in large-scale wireless sensor networks. This study is functional for a source to find the place of the sink or sinks.

We adopted many of the ideas from the above defined studies. Our entropy approach, however, is novel to the best of our knowledge.

III. SIMULATIONS

In this study simulations are carried out assuming that there are randomly distributed sensors in a terrain. The number of sensors is unknown after the initial deployment as they may die randomly. They could be either “ON” or “OFF” during each epoch. In our simulations we took the lifetime of each sensor as 20 epochs. That is, each sensor can be ON 20 times. The remaining life of each sensor decreases after each “ON” state. All the sensor nodes in the network have a probability of being ON which is denoted as p_s. The base station sends a threshold probability pt to the sensor nodes. If $p_s>p_t$, then the sensor node transmits.

The area of coverage is divided into grids. Each grid should contain at least one ON sensor during each epoch in order to cover the whole network area.

The algorithm tries to minimize the total number of ON sensors and maximize entropy. The former helps us to achieve maximum network lifetime while the latter is used to provide adequate coverage. For this study we used the following relationships to calculate different parameters.

We defined x_i as the number of sensors which are ON in a grid and defined a probability p_i related to the number of ON sensors in a grid as given in equation (1)

$$p_i = \frac{x_i}{\sum x_i} \tag{1}$$

Based on this probability, entropy H for a particular epoch is calculated as:

$$H = -\sum p_i \log_2 p_i \tag{2}$$

Maximum entropy means that we ideally have only a single sensor ON in each grid. If there are N grids then for maximum entropy p_i is defined as:

$$p_i = \frac{1}{N} \tag{3}$$

In this case entropy is calculated as in equation (4).

$$H_{\max} = \log_2 N \tag{4}$$

In the simulations it is assumed that 1600 sensors are randomly distributed in user defined grids. Initially p_t is set to 1. We ran the program for 50 times and checked that the numbers of sensors in each grid is randomly distributed.

To find the optimum value of the number of ON sensors, threshold probability changes adaptively. The goal is to obtain an entropy value which is as close to the maximum or reference entropy as possible. During the simulation, to get the highest possible value of entropy, each entropy value is compared to the previous one. Based on this comparison, if the entropy is larger than or the number of ON sensors is smaller than the previous values, threshold probability stays the same. Otherwise the threshold value is increased. If there are no active sensors in a grid then threshold probability is decreased to get new ON sensors.

Although sensors are alive there could be no sensors ON in a grid for a particular threshold probability value which means there can be no information recorded for that grid. Initially we reduced p_t when such a situation occurred. In another approach adopted, for such situations threshold probability is not decreased immediately. Instead, the same probability value is given a second chance. This way the rapid decrease in the threshold probability is avoided. If an active sensor is available in the proceeding epoch then the simulation continues with the existing p_t. On the other hand if there are no active sensors for a grid for the second time than threshold value is decreased.

Maximization of the entropy and minimization of total number of ON sensors depend on the threshold probability. While the number of ON sensors in one grid decreases, the possibility of having no sensors ON in other grids may increase. Another challenge is that for small grids (i.e. when we divide the area into high number of grids) the threshold probability should be high. Number of ON sensors is naturally reduced if the threshold probability is increased. However this may also lead to a reduction in the entropy. Although ideally maximum entropy value is assumed to be obtained for the cases where each grid contained only one active sensor for each epoch, same result would have been obtained for the cases where each grid had equal number of ON sensors other than 1. But as we tried to reduce the number of ON sensors alongside adjusting the entropy, the probability of this situation occurring is reduced.

IV. RESULTS

We tried different approaches to prolong the network lifetime as much as possible while maximizing coverage. In each of the cases we repeated the experiments 20 times and the values represent the averages of these 20 runs. We used $\Delta p_t=0.02$.

In the first case (Case 1) we just checked the number of ON sensors and tried to minimize this number while making sure that there is at least one sensor ON in each grid. The threshold probability is increased by $\Delta p_t/2$ if there were too many sensors ON; otherwise it was reduced by the same amount. The results of number of ON sensors for each epoch are shown in Figure 1.

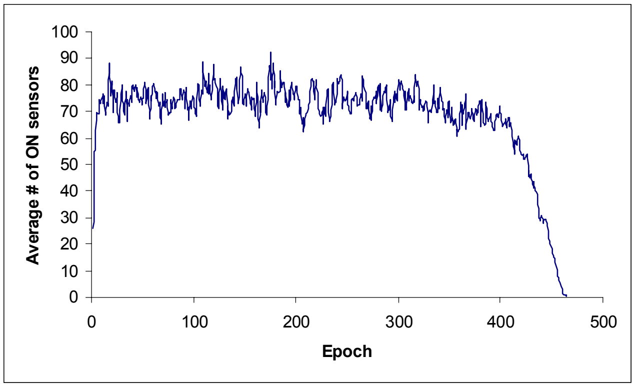

Fig. 1. Number of ON sensors vs. time for case 1 where the threshold probability is adjusted every epoch with 16 grids used

The values for each epoch represent the average of 20 runs. The coverage area was split to 16 grids in this case and the initial number of sensors deployed was assumed to be 1600. The average number of ON sensors per epoch was measured to be 72. The network lifetime was 447 epochs.

In the second case (Case 2) we tried to control the number of ON sensors using both entropy as defined above and the criteria we employed in Case 1. Entropy values and also total number of ON sensors are measured for each epoch and new values are compared with the old ones. If the entropy is higher than the old one or the value of total number of ON sensors is smaller than the previously measured value, then the threshold probability stays the same. Otherwise p_t is modified as $p_t=p_t+\Delta p_t/2$. When there are no ON sensors in any of the grids, the value of p_t is adjusted as $p_t=p_t-\Delta p_t$ to obtain new threshold value. The change in the number of ON sensors for each epoch is presented in Figure 2. Again it was assumed that there were 16 grids and initially 1600 sensors were deployed. The average number of ON sensors for each epoch was 67 in this case with a network lifetime of 474 epochs.

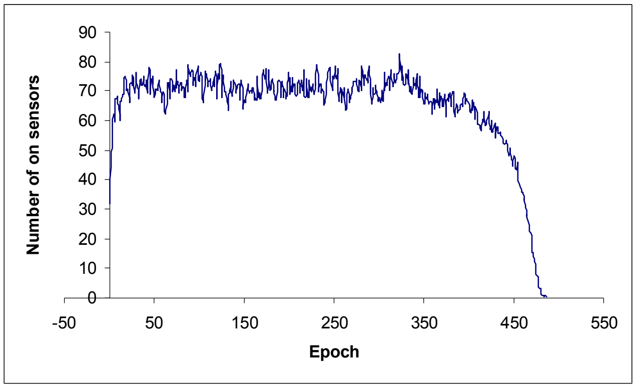

Fig. 2. Number of ON sensors vs. time for Case 2 where the threshold probability is adjusted every epoch with 16 grids used. Entropy is used to adjust the probability.

The network lifetime which is defined as the instance when there is no live sensor left in any one of the grids and the average number of ON sensors per epoch are similar in both cases although there is slight improvement for Case 2.

We then changed our adjustment policy for p_t and adjusted this value every 2 epochs to allow for occasional situations where one grid may not be covered (Case 3). The results show an improved network life as shown in Figure 3. In this case the average number of sensors per epoch and the network lifetime was 45 and 719 respectively for the case where entropy was used. These figures were 52 and 620 when entropy was omitted as shown in Figure 4(Case 4).

We also checked the case where we changed the number of grids from 16 to 9 keeping the number of sensors initially placed at 1600. The result of number of ON sensors for each epoch is presented in Figures 5 and 6 for the case where entropy is omitted (Case 5) and Case 6 where entropy is considered respectively. Introducing entropy as a parameter not only reduces the average number of ON sensors but also the epochs where there are no ON sensors are eliminated. This can be observed from Figures 7 and 8 which are representative single runs from Case 5 and Case 6 respectively. The network lifetime and average number of ON sensors for each epoch are summarized in Table I.

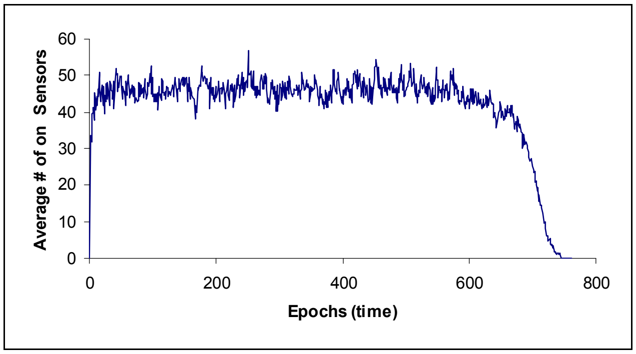

Fig. 3. Number of ON sensors vs time for Case 3 where the threshold probability is adjusted every other epoch using entropy for a 16 grid network.

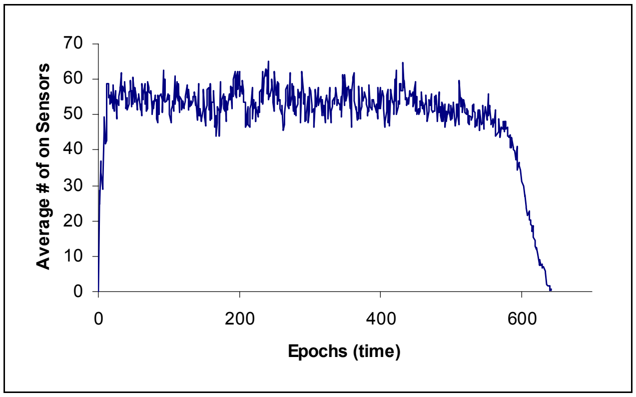

Fig. 4. Number of ON sensors vs time for Case 4 where the threshold probability is adjusted every other epoch, entropy omitted, 16 grids used.

We have also tried the measure the network lifetime for different number of grids and tested this relation using the case where we used entropy and adjusted p_t every other epoch. The results are shown in Figure 9.

TABLE I
AVERAGE NUMBER OF ON SENSORS AND NETWORK LIFETIME FOR DIFFERENT CASES

	Average Number of ON Sensors	Network Lifetime (epochs)
Case 1	72	447
Case 2	67	474
Case 3	45	719
Case 4	52	620
Case 5	27	1262
Case 6	21	1528

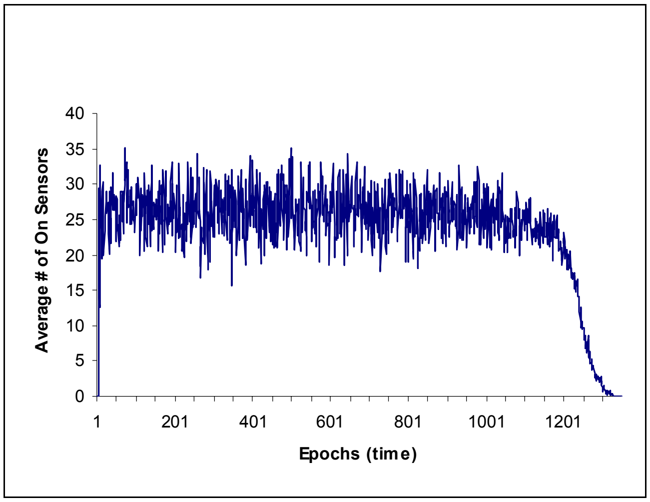

Fig. 5. Number of ON sensors vs time for case 5 where threshold probability is adjusted every other epoch and number of grids is set to 9.

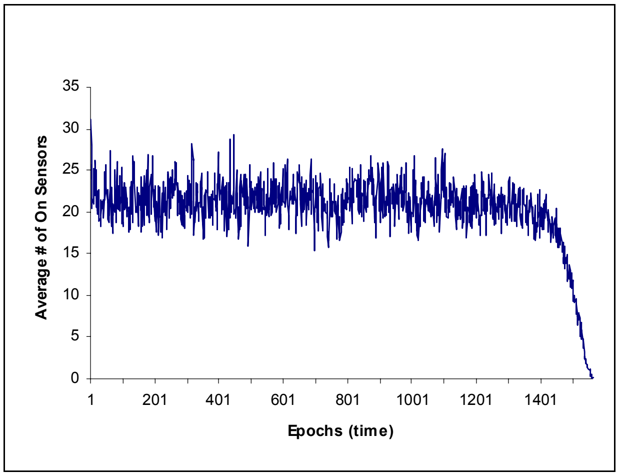

Fig. 6. Number of ON sensors vs time for case 6 where threshold probability is adjusted every other epoch and number of grids is set to 9.

V. CONCLUSIONS

In this study, we attempted to obtain a uniform coverage in a sensor network through the use of entropy. This is also expected to yield a prolonged network lifetime as the number of ON sensors at each epoch is limited. As expected number of active sensors at each epoch and network lifetime are inversely proportional. We, therefore, used different approaches to control the number of ON sensors and introducing entropy as a control parameter improved the network lifetime.

We used a threshold probability to turn the sensors ON and depending on the coverage we adjusted this probability. In one approach we did this every epoch while in another case we changed this probability every other epoch if so needed. The former scheme produced a 6% improvement over the case where entropy is not used and the latter case 16%.

Increasing the grid size (or reducing the number of grids) also helped to achieve longer network lifetime. When 9 grids are used, over 100 % improvement is achieved over the 16 grid case even when the entropy is not used. Using entropy introduced a further 21 % improvement. Our measurements show that as the number of grids is reduced, the network life is increased as shown in Figure 9. This is because as there are more sensors in a grid there is a better chance of having ON sensors hence higher probability values may be used. But in this case we are assuming that each sensor has a wider sensing area.

As a future work location information will also be obtained using an approach similar to the ones described in section 2.

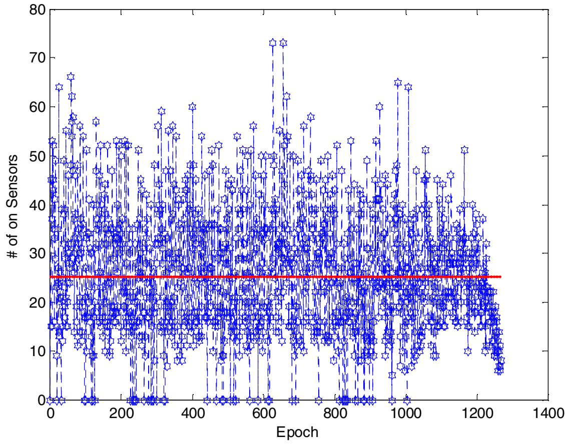

Fig. 7. Number of ON sensors vs time for a single run of Case 6 where threshold probability is adjusted every other epoch and number of grids is set to 9.

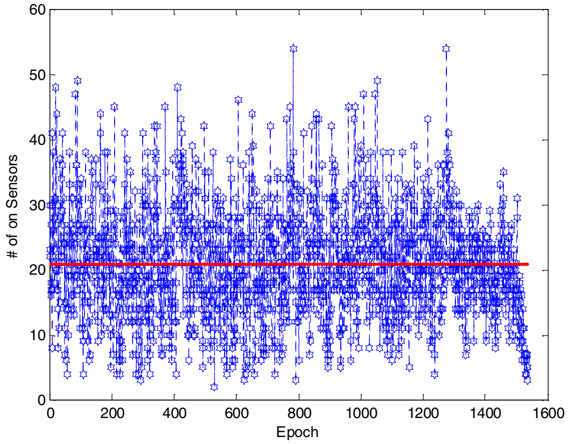

Fig. 8. Number of ON sensors vs time for a single run of Case 6 where threshold probability is adjusted every other epoch and number of grids is set to 9.

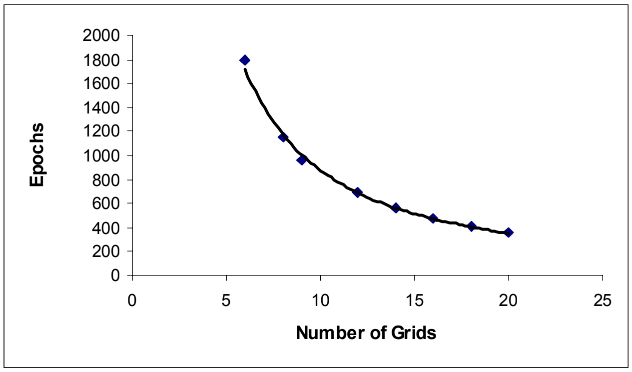

Fig. 9. Network life in epochs versus number of grids for the case where the threshold probability is adjusted every other epoch, initially 1600 sensors deployed in each case.

REFERENCES

[1] Iyer, R., Kleinrock, L., QoS Control for Sensor Networks, Communications, IEEE International Conference (ICC '03),1, 517-521, May. 2003.

[2] Meguerdichian, S., Koushanfar, F., Potkonjak, M., Srivastava, M. B., Coverage Problems in Wireless Ad-hoc Sensor Networks, Proc. of IEEE (INFOCOM 2001), 1380-1387, Anchorage, Alaska, April 2001.

[3] Frolik, J., QoS Control for Random Access Wireless Sensor Networks, Wireless Communications and Networking Conference (WCNC. 2004), 3, 1522-1527, Atlanta, GA Mar. 2004.

[4] Kay, J. Frolik, J., Quality of Service Analysis and Control for Wireless Sensor Networks, IEEE International Conference(MASS2004), 359-368, San Jose, CA,Oct. 2004.

[5] Wang, H., Yao, K.,Pottie, G., Estrin, D., Entropy-based Sensor Selection Heuristic for Target Localization, Proceedings of the third international symposium on Information processing in Sensor Networks (IPSN'O4), 36-45, Berkeley, California, April, 2004.

[6] Johansson, M., Sternad, M., Resource Allocation Under Uncertainty Using the Maximum Entropy Principle, IEEE Transactions on Information Theory, 51 (12), 4103-4117, December,2005.

[7] Ammari, H.M., Das, S.K., Data Dissemination to Mobile Sinks in Wireless Sensor Networks: An Information Theoretic Approach, IEEE International Conference(MASS2005), 8, Washington DC, Nov. 2005.

[8] Nama, H. Mandayam, N. , Sensor Networks Over Information Fields: Optimal Energy and Node Distributions, Wireless Communications and Networking Conference, IEEE(WCNC 2005), 3, 1842-1847,New Orleans, LA, March 2005.

[9] Chen, H., Grid-based Approach for Working Node Selection in Wireless Sensor Networks, ACM Symposium on Applied Computing (SAC 2002), 1134-1139, Madrid, Spain, March 2002.

[10] Akl, R., Sawant, U., Grid-based Coordinated Routing in Wireless Sensor Networks, IEEE Consumer Communications and Networking Conference (CCNC 2007), 5,. January 2007.

[11] Tan H., Maximizing Network Lifetime in Energy-Constrained Wireless Sensor Network, International Conference On Communications and Mobile Computing (IWCMC 2006), Vancouver, British Columbia, Canada, 1091-1096, July 2006.

[12] Stoleru, R., Stankovic, J. A., Probability Grid: A Location Estimation Scheme for Wireless Sensor Networks, The First IEEE Communications Society Conference on Sensor and Ad Hoc Communications and Networks (SECON, 2004), 430-438, Santa Clara, CA, October 2004.

[13] Stoleru, R., He, T., Stankovic, J. A., Walking GPS: A Practical Solution for Localization in Manually Deployed Wireless Sensor Networks, 29th Annual IEEE International Conference on Local Computer Networks (LCN'04), 480-489, November 2004.

[14] Bahl, P., Padmanabhan, V.N., RADAR: An In-Building Rf-based User Location and Tracking System, Proceedings of IEEE (Infocom), 2, 775-784 Tel Aviv, Israel, March 2000.

[15] Pryiantha,N.B., Chakaborty, A., Balakrishnan, H., The Cricket Location-Support System, The Sixth Annual ACM International Conference on Mobile Computing and Networking(MobiComm2000), 32-43, Boston, Massachussetts, August 2000.

[16] Bulusu, N., Heidemann, J., Estrin, D., GPS-Less Low Cost Outdoor Localization for Very Small Devices, IEEE Personal Communications, 7(5), 28-34, October 2000.

[17] He, T. Huang, C., Blum, B., Stankovic, J.A., Abdelzaher, T., Range-Free Localization Schemes for Large Scale Sensor Networks, The Ninth Annual International Conference on Mobile Computing and Networking, 81-95, San Diego, CA, September 2003.

[18] Nagpal, R., Shrobe, H. Bachrach, J., Organizing A Global Coordinate System for Local Information on an Ad Hoc Sensor Network, MIT A.I Memo, 1666, 1999.

[19] Niculescu, D., Nath, B., DV-Based Positioning in Adhoc Networks, Telecommunication Systems, 22(1-4), 267-280, 2003.

[20] Zhang, R., Zhao, H., Labrador, M. A., A Grid-Based Sink Location Service for Large-scale Wireless Sensor Networks, Int. conference on Communications and mobile computing, Vancouver, British Columbia, Canada, 689-694, July 2006.