AN MMSE BASED WEIGHTED AGGREGATION SCHEME FOR EVENT DETECTION USING WIRELESS SENSOR NETWORK

Bhushan G. Jagyasi, Bikash K. Dey, S. N. Merchant, U. B. Desai

SPANN Laboratory, Electrical Engineering Department,
Indian Institute of Technology Bombay, Powai, Mumbai-76
Email: {bhushan,bikash,merchant,ubdesai}@ee.iitb.ac.in

ABSTRACT

We consider design of wireless sensor network for event detection application. An MMSE based weighted aggregation scheme is proposed for event detection application using wireless sensor network. Accuracy and the network lifetime are the two performance evaluating parameters considered here. We compare the performance of the proposed scheme with the previously known schemes.

1. INTRODUCTION

In this paper, we explore the use of wireless sensor network (WSN) for an event detection application. Specifically, we extend the results of Steven [2], wherein we consider a weighted aggregation scheme as opposed a majority decision based aggregation scheme of [2]. Wireless sensor network [4] is a distributed network formed by deploying sensor nodes in the application area in large number. Each sensor node consists of the following components (i) multiple sensors to measure physical parameters, (ii) a micro-controller, (iii) small memory and (iv) a transceiver. These tiny sensor nodes are powered with a small battery having limited power. Basically, each sensor node has very small footprint. Multi-hop protocol is used to communicate the sensed data between different sensor nodes via the transceiver.

For event detection application, each sensor is assumed to sense the local information about the occurrence of global event correctly with some probability. The information sensed by all the nodes ultimately reaches the sink either in true form or after being processed by the intermediate aggregators. Thus, using all the information available, the sink makes the final global decision about an event. The performance of the network is thus determined by how accurately the event is detected by the sink. The objective is to make the event detection more accurate while maximizing the lifetime of the network. The basic requirements of WSN to be scalable and self organizable are also taken into consideration. All this is achieved under the constraint of limited power.

Recently, the use of WSN for landslide detection was proposed in [1]. Distributed detection algorithms for WSN have been talked of quite often in the past (for example see [3, 7, 8, 9]. In [7] a distributed algorithm was proposed to maximize the lifetime of the WSN. Whereas, [8, 9] discuss fundamental and advanced distributed algorithms with multiple sensor, where observation of each sensor is one bit. In [5, 6] optimization across routing, link and MAC layer is proposed to maximize the lifetime of WSN.

[2] considers event detection application for WSN. Our work relies heavily on the interesting results presented in [2]. In [2], aggregation scheme $M1$ and link cost for routing $C1$ is proposed. It uses Bellman-Ford routing algorithm algorithm with link cost $C1$ to obtain the spanning tree oriented towards the sink. $C1$ for link (i,j) is defined as $C1 = I_j/B_i$, where I_j is the number of nodes which can transmit to j^{th}node and B_i is the battery level of the i^{th}node. Considering I_j in the numerator and B_i in denominator, it is ensured that there are not too many children of the same node. This results in a balanced spanning tree. Aggregation scheme $M1$ states that each node makes one bit decision based on majority of the decisions received from its children. This one bit decision is then transmitted to its parent node which again follows the same aggregation scheme. This process goes on until the sink makes the final decision about an event. In was shown in [2] that use of aggregation scheme $M1$ and link cost $C1$ enables improved accuracy for event detection. Moreover, with aggregation scheme $M1$, we get better network lifetime since each nodes transmits only one bit to its parent.

[2] also proposes the infinite precision aggregation scheme $M2$ which is used with spanning tree obtained by Bellman-ford routing using link cost $C2$. Link cost $C2$ for link (i,j) is defined as $C2 = P_{ij}/B_i$. Where, P_{ij} is the power required to transmit a bit from i^{th} node to j^{th} node which depends on the path loss of the link (i,j). Infinite precision aggregation scheme $M2$ requires transmission of multiple bits from a node to its parent. In $M2$ every i^{th} node finds out the number of nodes O_i (one's) observing the occurrence of event H and number of nodes Z_i (zero's) observing non-occurrence of event H in the subtree originating from itself. These O_i and Z_i are computed at every node based on its observation and based on information received from its children. O_i and Z_i are then communicated to their respective parent node using multiple bits. Finally sink node computes global O and Z based on the data received from its children and makes a final decision about the event. Aggregation scheme $M2$ will have poor network lifetime since it requires transmission of more number of bits from every node to its parent.

In order to motivate our weighted aggregation scheme, we present an example, where routing using link cost $C1$ does not results in a balanced tree. Figure 1 shows a spanning tree obtained by applying the Bellman-Ford routing algorithm using link cost $C1$ on a sensor network with 100 nodes in a square area. We assume a uniform distribution for randomly locating nodes. Here, different branches of any subtree emerging from any node have different number of nodes. Thus the majority decision rule will not result in optimum aggregation since it gives equal weights to the decisions coming from different children. However, decision made by any child may be based on data observed by different number of nodes. We propose a minimum mean square error (MMSE)

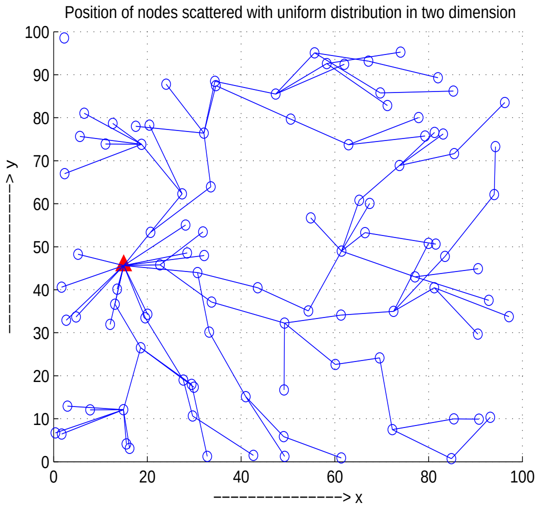

Figure 1: The Spanning tree as a result of Bellman-Ford routing algorithm using link cost $C1$

based weighted aggregation scheme (WAS), in which a node consider the decision from its children as before. However now, it also uses the number of descendants of each of its child for the purpose of weighing the decision. Here too, the decision of each node is one bit and consequently better network lifetime attribute of the aggregation scheme $M1$ is preserved. Because of the MMSE approach, we will see later, WAS gives better accuracy even for an unbalanced tree.

In Section 2, we present the construction of the MMSE based weighted aggregation scheme. In Section 3, we show how WAS can be used in WSN for event detection application. In Section 4 we carry out extensive simulation considering the proposed weighted aggregation scheme, the aggregation scheme $M1$ and the aggregation scheme $M2$. For $M2$ we use link cost $C2$, while for WAS and $M1$ we use link cost $C1$. From the simulation it is seen that: (a) The aggregation scheme $M1$ and weighted aggregation scheme both results in far better network lifetime as compared to the network lifetime of aggregation scheme $M2$, (b) Network lifetime of WAS and aggregation scheme $M1$ is comparable and (c) WAS outscores the aggregation scheme $M1$ in terms of accuracy. These results shows advantage of WAS over other aggregation schemes for event detection application. Finally, Section 5 concludes the paper.

2. THE DEVELOPMENT OF MMSE BASED WEIGHTED AGGREGATION SCHEME (WAS)

A local view of the wireless sensor network is shown in Figure 2. It shows parent node S_0 with its k children $S_1, \cdots, S_k$. In our work we consider the node as its own descendant along with other descendants. Let node S_i have N_i number of descendants (not only the children), denoted by $S_i, S_{i2}, \cdots, S_{iN_i}$. The sink is considered to be at level 0, whereas the level for any other node will be the number of hops it is away from the sink node. The node S_0 may be at any level in the network. We assume transmission of one bit from any sensor node to its parent node as its local decision. We assume a node knows the number of descendants its children have. For example the node S_0 knows the number of descendants N_i of its child S_i. Among all the descendants N_i of the node S_i, let the number of nodes deciding in favor of the occurrence of the event be n_i. Here we propose a MMSE based weighted aggregation scheme in which parent node S_0 computes the MMSE estimate $\hat{n}_i$ of n_i. Specifically, $\hat{n}_i$ is the MMSE estimate of the number of descendants of node S_i deciding in favor of the occurrence of the event.

For simplifying the analysis we assume the decision made by any node is the majority decision of all its descendants. Thus if the observations of the descendants of the node S_i are $X^i = \{X_{i1}, X_{i2}, ..X_{iN_i}\}$, then the decision made by the node S_i is given by

$$X_i = maj\{X^i\} \tag{1}$$

where for a set B of binary numbers, we define $maj\{B\} = 1$ if there are more or equal number of ones than zeros in B, while $maj\{B\} = 0$ otherwise. Similarly,

$$X = maj\{X^0 \cup X^1 ... \cup X^k\} \tag{2}$$

where $X^0 = \{X_0\}$.
We assume that the following information is available at S_0.

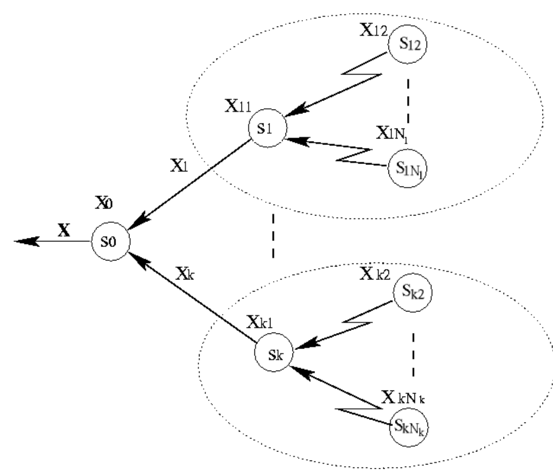

Figure 2: A local view of the network

1. Observation X_0 of the node S_0 itself.
2. Decision X_i made by its i^{th} child S_i, for $i = 1, 2, \cdots, k$.
3. Number of descendants N_i of S_i.
4. The probability p of correct sensing or accuracy of all the sensors.

The problem is to get an estimate $\hat{X}$ of X from the above information. The estimate

$$\hat{X} = \arg \max_{\beta \in \{0,1\}} P(X = \beta | \mathbf{X}) \tag{3}$$

will have the highest accuracy $P(X = \hat{X})$ where $\mathbf{X} = [X_0, X_1, ..X_k]$.

Recall that n_i is the actual number of descendants of node S_i observing in favor of the hypothesis (i.e., $H = 1$) and thus $n = n_0 + n_1 + \cdots + n_k$ will be the actual number of descendants of the node S_0 observing in favor of $H = 1$. Here we obtain the minimum mean square error (MMSE) estimate $\hat{n}_i$ of n_i and get the MMSE estimate of n as

$$\hat{n} = \hat{n}_0 + \hat{n}_1 + \cdots + \hat{n}_k. \tag{4}$$

This estimate can then be used to get an estimate of X. In the following, we obtain the MMSE estimate $\hat{n}_i$ of n_i given X_i and N_i for a particular node S_i. Now,

$$\begin{aligned} P(n_i|X_i=1) = P(n_i|X_i=1,H=0)P(H=0|X_i=1) \\ +P(n_i|X_i=1,H=1)P(H=1|X_i=1). \end{aligned} \tag{5}$$

Recall that p is the probability of correct sensing by the sensors. Then

$$\begin{aligned} P(X_i=1|H=0) &= P(X_i=0|H=1) \\ &= \sum_{l=1}^{\lfloor \frac{N_i-1}{2} \rfloor} {}^{N_i}C_l p^l (1-p)^{(N_i-l)} \end{aligned}$$

where ${}^nC_k = \frac{n!}{k!(n-k)!}$, and

$$\begin{aligned} P(X_i=1|H=1) &= P(X_i=0|H=0) \\ &= \sum_{l=\lceil \frac{N_i}{2} \rceil}^{N_i} {}^{N_i}C_l p^l (1-p)^{(N_i-l)}. \end{aligned} \tag{6}$$

Now,

$$\begin{aligned} P(X_i=1) &= P(X_i=1|H=0)P(H=0) \\ &\quad +P(X_i=1|H=1)P(H=1) \\ &= \frac{1}{2}[P(X_i=1|H=0)+P(X_i=1|H=1)]. \end{aligned}$$

Using equation 6, we have

$$\begin{aligned} P(X_i=1) &= \frac{1}{2}[P(X_i=1|H=0)+P(X_i=0|H=0)] \\ &= \frac{1}{2}. \end{aligned} \tag{7}$$

Note that this ensures transmission of maximum average information by each transmitted bit. Using Bayes' rule, we get

$$\begin{aligned} P(H=h|X_i=1) &= \frac{P(X_i=1|H=h)P(H=h)}{P(X_i=1)} \\ &= P(X_i=1|H=h) \end{aligned} \tag{8}$$

where $h \in \{0,1\}$. From equations (8) and (5) we get

$$\begin{aligned} P(n_i|X_i=1) &= P(n_i|X_i=1,H=0)P(X_i=1|H=0) \\ &\quad +P(n_i|X_i=1,H=1)P(X_i=1|H=1) \end{aligned}$$

$$= P(n_i, X_i=1|H=0) + P(n_i, X_i=1|H=1)$$

$$= \begin{cases} {}^{N_i}C_{n_i} p^{N_i-n_i}(1-p)^{n_i} & \\ \quad + {}^{N_i}C_{n_i} p^{n_i}(1-p)^{N_i-n_i} & \text{if } n_i \geq \frac{N_i}{2} \\ 0 & \text{otherwise.} \end{cases} \tag{9}$$

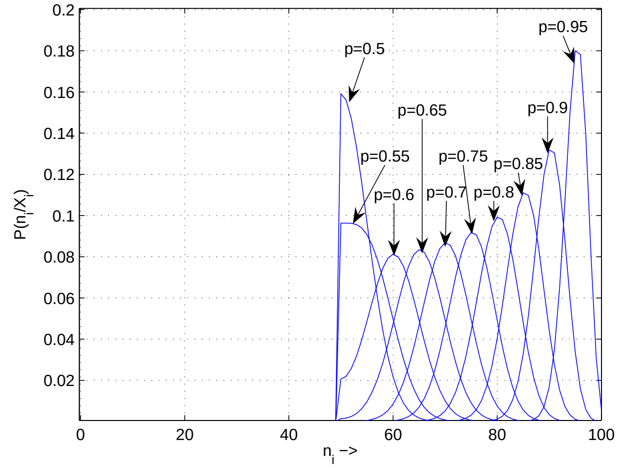

Figure 3: Conditional probability mass function of n_i for $X_i = 1$, N_i=100 and various values of p

Similarly,

$$P(n_i|X_i=0) = \begin{cases} {}^{N_i}C_{n_i} p^{N_i-n_i}(1-p)^{n_i} & \\ \quad + {}^{N_i}C_{n_i} p^{n_i}(1-p)^{N_i-n_i} & \text{if } n_i < \frac{N_i}{2} \\ 0 & \text{otherwise} \end{cases} \tag{10}$$

Figure 3 shows the conditional probability mass function of n_i given $X_i = 1$, obtained using equation 9. It shows multiple plots for different accuracies (p) and for 100 descendants (N_i=100). Here, for ease of visualization, the probability mass functions of the discrete variable n_i are shown as continuous plots. Similar plots can be obtained for $X_i = 0$ using equation (10), in which the curves will be nonzero in $n_i = 0$ to $\lfloor \frac{N_i-1}{2} \rfloor$ and will be the mirror image of the curves seen in Figure 3.

The MMSE estimate of n_i given $X_i = 1$ is given by the mean

$$\hat{n}_i|_{X_i=1} = \sum_{j=\lceil \frac{N_i}{2} \rceil}^{N_i} jP(n_i=j|X_i=1). \tag{11}$$

Similarly,

$$\hat{n}_i|_{X_i=0} = \sum_{j=0}^{\lfloor \frac{N_i-1}{2} \rfloor} jP(n_i=j|X_i=0). \tag{12}$$

The node S_0 finds the estimate $\hat{n}$ of n as

$$\hat{n} = \sum_{i=0}^{k} \hat{n}_i \tag{13}$$

and makes its decision as

$$\hat{X} = \begin{cases} 1 & \text{if } \hat{n} \geq \frac{N}{2} \\ 0 & \text{otherwise} \end{cases} \tag{14}$$

where $N = 1 + \sum_{i=1}^{k} N_i$.

3. THE PROPOSED ALGORITHM

Based on the above formulation we now show how WAS can be effectively used in WSN for event detection application. At first, we assume that the network topology does not change very often and thus the network is static and we show that improvement in accuracy is achieved with the same network lifetime as in the case of M1. The assumption of the static network topology can be relaxed with the help of minimal communication overhead whenever the topology changes.

3.1 The WAS Scheme for Static Network

Here we assume that once the nodes are deployed, the network topology does not change till the nodes start dying out. Initially each node is assigned the same battery level. We define network lifetime as the time before the first node in the network dies. Since the network is static, the number of descendants of a node remains fixed throughout the network lifetime.

The weighted aggregation scheme works as described below.

1. Using Bellman ford routing algorithm, the spanning tree is obtained as shown in Figure 1.
2. The sink node assigns level zero to itself and conveys its level information to its children. The children of the sink then assign level 1 to themselves and convey this information to their children and so on till the leaf nodes.
3. Each node is assigned time slot in such a way that any parent node gets a time slot after all its children nodes.
4. Before the start of first session the process of descendant update is initiated. This is done during the initial setup of a sensor network. Starting from the leaf nodes, each node S_i calculates the number of its descendants N_i as $N_i = 1 + \sum_{j \in R_i} N_j$ where R_i is the set of children nodes of S_i. N_i's thus calculated are transmitted to the parent node at lower level using few bits. These N_i's remain fixed till nodes starts dying. Thus before the first session begins, each node knows the number of descendants their children have.
5. Using equation 11 and 12 each sensor node S_i calculates once and keeps in memory the expected number of descendants deciding in favor of H=1 given $X_j = 1$ ($\hat{n}_j|_{X_j=1}$) and given $X_j = 0$ ($\hat{n}_j|_{X_j=0}$) using the corresponding values of N_j, $j \in R_i$. If this computation is undesired, these values can be preloaded in each sensor as a $2 \times L$ matrix, where L is the maximum number of descendants a node can have. The $(1,k)^{th}$ element of the matrix will contain the mean value $\hat{n}_j|_{X_j=1}$ for $N_j = k$ and the $(2,k)^{th}$ element will contain the mean value $\hat{n}_j|_{X_j=0}$ for $N_j = k$. After the initial setup, a node needs to keep only the columns corresponding to the number of neighbors of its children and free most of the memory occupied by the matrix for other use. During any session, if a node does not hear from all its children, there must be a dead child node and thus the network is assumed to be dead.
6. During every session, each node S_i obtains $\hat{n}_i$, the estimate of the number of its descendants deciding in favor of $H = 1$ as
$$\hat{n}_i = \sum_{j \in R_i} \hat{n}_j. \tag{15}$$
7. Now $\hat{X}_i$, the decision made by the node S_i, is obtained as
$$\hat{X}_i = \begin{cases} 1 & \text{if } \hat{n}_i \geq \lceil \frac{N_i}{2} \rceil \\ 0 & \text{otherwise} \end{cases}$$
and is communicated by S_i to its parent node.

Note that the communication overhead for each node S_i in this scheme over M1 is only due to the transmission of the number N_i. However these N_is are transmitted only once during the initial setup and so the overhead incurred at each node is negligible. Hence the lifetime of WAS is almost same as that of the aggregation scheme M1. Simulation results in Section 4 show that the accuracy obtained by WAS is better than that of M1.

3.2 Relaxing the Assumption of Static Network

After any node dies out or gets eliminated form the network, network topology changes. To keep track of change in network topology update of number of descendants is required. It can be done in any of the following ways: (i) The N_i's may be periodically updated after every few sessions as done during the initial setup. (ii) The N_i's are updated as soon as death of any node is detected by its parent. In this case update of the number of descendants takes place for all the ancesters of a dead node.

However, to keep the network alive even after some node die, the routing itself needs to be redone to make it more suitable for the changed topology.

4. SIMULATION RESULTS

Simulations have been performed for number of nodes M=100 uniformly deployed in a square grid area of size M^2. We make sure that the average number of neighbors per node is in between 7 and 9. We consider equal probability of occurrence of the event H, i.e., $P(H = 0) = P(H = 1) = 0.5$. We assume every sensor has same probability p of correct sensing. For the following results, we vary the probability p of correct sensing of all the sensors from 0.55 to 9, with increments of 0.05. Recall that the accuracy of the system is the probability of correct decision made by the sink node. The accuracy and network lifetime are plotted against p for the aggregation schemes M1, M2 and WAS. The plots are the average of 60 different random deployments. Here the initial battery level assigned to each node is 50000 units. For transmission of a single bit from a node to its parent, the transmission power required is proportional to the cube of the distance between the two. In our simulation, this has an average value of about 150 units. The power required by a node for receiving a single bit is fixed to 100 units.

In both *WAS* and $M1$ each node transmits a single bit and receives on average 4 bits during each session. Thus network lifetime for them is found to be comparable. However for the $M2$ aggregation scheme, the number of bits transmitted and received are significantly large and they depend upon the total number of nodes in the network. Here for present simulation network lifetime for $M2$ is found to be very less as compared to *WAS* and $M1$ aggregation schemes. Thus $M2$ gives very poor performance in terms of network lifetime, as

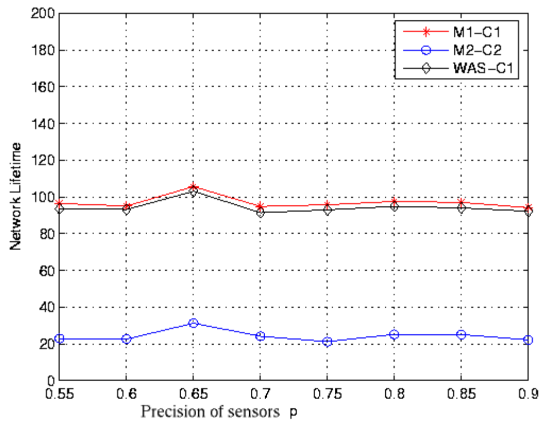

Figure 4: Comparison of lifetime for aggregation schemes M1-C1, M2-C2 and WAS-C1

seen in Figure 4, and hence is not suitable for wireless sensor network.

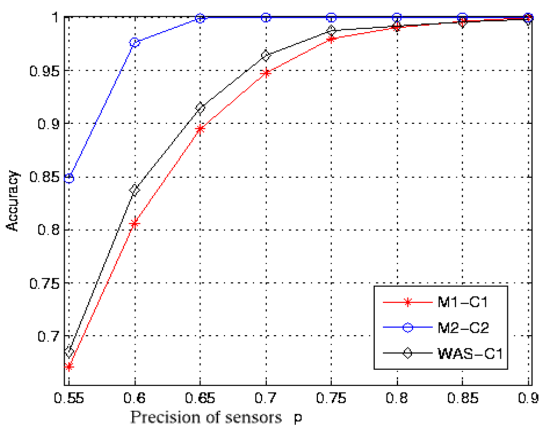

Figure 5: Comparison of accuracy vs probability of correct sensing for M1-C1, M2-C2 and WAS-C1

Figure 5 compares aggregation schemes M1, M2 and Weighted Aggregation Scheme (WAS) in terms of accuracy. It has been observed that aggregation scheme M2 has the best accuracy among all the three methods. This is expected from the fact that the sink has the exact information of the total number of observations in favor of $H = 1$ and $H = 0$ in the whole network. But this gain in accuracy for M2 comes at the significant loss of lifetime since each node transmits more number of bits per session. Since lifetime is the most critical aspect for wireless sensor network, M2 may not be the preferred aggregation scheme although it gives the best accuracy. In contrast, we have WAS and M1, both of them having very good network lifetime since in these aggregation schemes each node transmits a single bit per session. Also it is observed that the lifetime of WAS is comparable to that of M1. Now when we compare WAS and M1 in terms of accuracy WAS outscores M1.

5. CONCLUSION

An MMSE based weighted aggregation scheme is proposed for event detection using a wireless sensor network. The scheme requires transmission of a single bit by each sensor for every session as required by the previously proposed scheme M1 in [2]. So this scheme provides the same network lifetime as the scheme M1. This scheme outscores aggregation scheme M1 in terms of accuracy and M2 in terms of network lifetime. Thus the WAS is the most preferred aggregation scheme among those known so far for event detection using WSN.

REFERENCES

[1] A. Sheth, K. Tejaswi, P. Mehta, C. Parekh, R. Bansal, S.N.Merchant, U.B.Desai, C.Thekkhath, K. Toyama and T.Singh, "Poster Abstract-Senslide: A Sensor network Based Landslide Prediction System," in *ACM Sensys*, November 2005.

[2] Steven A. Borbash, "Design considerations in wireless sensor networks, " *Doctoral thesis submitted to University of Maryland*, 2004.

[3] R. Niu and P. K. Varshney, "Distributed detection and fusion in a large wireless sensor network of random size,"*EURASIP Journal on Wireless Communication and Networking* 2005, pp. 462-472.

[4] I.F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, "A survey on sensor networks, " in *IEEE Comm. Mag.*, Vol. 40, No. 8, August 2002, pp. 102-116.

[5] S.Cui, R.Madan, S. Lall, and A. J. Goldsmith, "A joint routing, MAC, and link layer optimization in sensor networks with energy constraints," in *IEEE International Conference on Communications*, 2005 , ICC 2005, Volume 2,16-20 May 2005 pp. 725 - 729.

[6] J.Xiao, S. Cui, Z.Q.Luo, and A.J.Goldsmith, "Joint estimation in sensor networks under energy constraint, " in *First annual IEEE Communicatins society conference on sensor and Ad-hoc communications and networks 2004* IEEE SECON, 2004. 4-7 Oct 2004 pp.264-271.

[7] R. Madan and S. Lall, "Distributed algorithms for maximum lifetime routing in wireless sensor networks, " in *Globecom'04*, Volume 2, 29 Nov- 3 Dec 2004, pp.748 - 753.

[8] R. Viswanathan and P. K. Varshney, "Distributed detection with multiple sensors: part I-fundamentals,"*Proceedings of the IEEE*, Vol. 85, Issue 1, Jan 1997, pp. 54-63.

[9] R. S. Blum, S. A. Kassam and H. V. Poor, "Distributed detection with multiple sensors: part II-advanced topics,"*Proceedings of the IEEE*, Vol. 85, Issue 1, Jan 1997, pp. 64-79.