

Optimum Routing and Scheduling for Estimation in Wireless Sensor Networks

Amir Hossein Mohajerzadeh, Mohammad Hossein Yaghmaee
Department of computer engineering, Ferdowsi university of
Mashhad, Mashhad, Iran
ah.mohajerzadeh@stu-mail.um.ac.ir, yaghmaee@ieee.org

Vahid Fakoor, Elham Mirfarah
Department of Statistics, Ferdowsi university of Mashhad,
Mashhad, Iran
fakoor@math.um.ac.ir, el_mirfarah@stu-mail.um.ac.ir

Abstract—In this paper, the distributed estimation in wireless sensor network using optimal task scheduling is considered. The main goal is to maximize network lifetime where it consists of a fusion center and a set of sensor nodes. Lifetime defined as the number of rounds accomplished before network becomes nonfunctional. In order to prolong network lifetime, we determine the optimal number of active sensor nodes and the number of samples provided by each of them using linear optimization problem. Sensor observations are quantized into messages and then forwarded to a fusion center where a final estimation is performed based on degree of certainty. Simulation results show that our optimal proposed algorithm has achieved significant efficiency when compared with other heuristic methods.

Keywords; *Degree of Certainty; Estimation; Task scheduling; Wireless sensor networks; Network lifetime*

I. INTRODUCTION

Wireless Sensor Networks (WSN) consists of a set of geographically distributed sensor nodes which perform their tasks as an integrated system. Sensor nodes have constraints such energy sources, computational power, storage capacity, etc [1]. WSN have current applications such as, environment monitoring, healthcare, battlefield surveillance, home automation, etc [2] and growing future applications such as distributed estimation, distributed detection and tracking.

In this paper, distributed estimation of unknown deterministic parameter has performed using set of observations whom provided by distributed sensor nodes. In distributed estimation each node delivers a subset of observations from environment to a central node called *Fusion Center* (FC), directly or indirectly. One of the main goals of FC is to reconstruct the underlying physical phenomenon based on input data gathering from sensor measurements. Estimation literature attracts a great deal of attention in computer networks [3, 4], also nowadays, it becomes an attractive topic in signal processing in wireless sensor networks [5, 6]. In WSN, sensor nodes collect real data, perform a local data compression and then send them to the FC. FC gathers data and produces a final estimation of the observed parameter.

Most of works which is done on estimation in WSN [7, 8], assume that the joint distribution of sensor's observations is known and the real valued messages can be sent from the sensors to the FC without distortion.

A common problem for all WSNs with different applications is resource constraints. Sensor nodes have only small batteries where replacement can be costly if not possible. In comparison with sensing and computation, communication is the most energy-consuming operation of the sensor nodes. Hence, in order to extend lifetime, reducing the communications between every type of sensor nodes is vital requirements of WSN's. Various methods have been proposed to increase network lifetime and efficiency in wireless sensor networks [9].

With respect to the WSN's characteristics [10], various distributed estimation algorithms have been designed for them [11, 12]. They addressed various design and implementation to digitize transmitted signal into several bits. The problem of decentralized estimation has been studied, in distributed control [13], in tracking [14] and data fusion [15]. [16] proposes optimal power scheduling problem for the decentralized estimation of a noise-corrupted deterministic signal in an inhomogeneous sensor network. They have determined the optimal quantization and transmit power level at local sensors so as minimize the total transmit power while ensuring a given mean square error (MSE) performance. [17] studies the optimal tradeoff between the number of active sensors and the quantization bit rate for each active sensor to minimize the estimation MSE. In [18] the estimation of a scalar field over a bidimensional scenario through a WSN with energy constraint is investigated. The paper provides a mathematical framework to analyze the independent aspects of WSN communication protocols and signal processing design. [19] has studied the performance-energy tradeoff for distributed estimation in a WSN. It has used Best Linear Unbiased Estimation (BLUE) to estimate observed phenomena. Like so many other works [19, 20] uses optimization problem to achieve the best possible functionality.

Generally, one of the most efficient ways to deploy a WSN over a target area is to cover the whole terrain using minimum possible number of sensor nodes. In this paper we will consider one of the easiest methods of communication is WSN: direct forwarding (DF). In DF all the sensor nodes forwards data to FC directly in one hop. As we know, by using higher number of sensor nodes, estimation process will be performed more precisely. When a network has less than enough wireless sensor nodes to achieve desired precision,

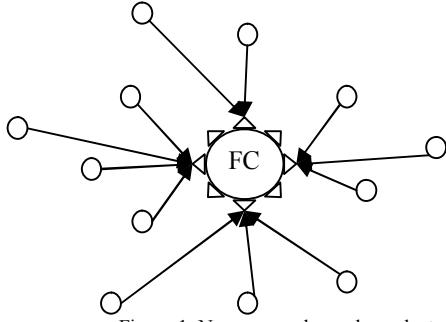


Figure 1. N sensor nodes make a cluster

adding more wireless sensor nodes can improve network functionality.

The main objective of this paper is to design an efficient scheduling scheme to control sensor nodes tasks (sensing + communication) in order to extend network lifetime as much as possible. Basically, scheduling is classified into four main categories [21]; they are the “*always alive*”, “*random on-off*”, “*adaptive on-off*” and “*periodic on-off*”. The proposed scheme acts as an adaptive on-off scheduling scheme in which FC creates a total scheduling program and other sensors only follow that. [22]

The paper is organized as follows. Section II introduces the system model of estimation in single hop WSN. We have used interval estimation; it also will be discussed in section II. In Section III the network lifetime maximization problem for the estimation in WSN is formulated as a linear programming (LP) problem. In order to demonstrate the performance of the proposed algorithm, simulation results have been shown in section IV. Finally section V concludes the paper.

II. PROBLEM FORMULATION

In figure 1, an example of a cluster of a wireless sensor network has been shown. WSN's structure is considered hierarchical in this paper. Each cluster consists of a set of N distributed cluster member sensor nodes and a FC (cluster head acts as fusion center) designed to cooperate to estimate an unknown parameter θ . Each cluster member observes the event, quantize and transmit its collected information to the fusion center. FC makes the final estimation based on all the received messages from cluster members. The observations are corrupted by additive noise and described using equation 1:

$$x_{ki} = \theta + \varepsilon_{ki} ; k = 1, 2, \dots, N ; i = 1, 2, \dots, n_k \quad (1)$$

x_{ki} refers to the i^{th} observation of sensor node k. Each sensor node can send many samples (observations) depended on problem conditions. n_k is the sample size which is provided by node k. Sensor noise variables, ε_{ki} , are considered to be independent, mean zero Gaussian random variables with $\text{var}(\varepsilon_k) = \sigma_k^2$ ($k = 1, 2, \dots, N$). θ is the parameter to be estimated. At the first step of estimation process, due to severe bandwidth and energy limitation of WSN, each sensor locally quantizes the real valued analog observation y_k into an unbiased discrete message $m_k = Q_k(y_k)$ of length L_k bits as in [3]. $Q_k(y_k)$ is quantization

function, and final message m_k is transmitted to the FC via direct wireless channel.

Network lifetime is divided into the different rounds. In each round based on the scheduling program which is advertised by the FC, cluster member nodes deliver their observations ($x_{11}, x_{12}, \dots, x_{Nn_k}$) to the FC. Then FC makes final estimation of collected observation using a fusion function $f: \bar{\theta} = f(x_{11}, x_{12}, \dots, x_{Nn_k})$. The precision of the estimation of unknown parameter θ is a crucial parameter for proposed algorithm efficiency. In this paper, using interval estimation, we wish the error of estimation to be under control. Estimation process is done independently in each cluster.

FC determines the number of sent messages, the time whom they sent and “*On-Off*” mode of every node, using scheduling program. The main goal of this paper is to maximize network lifetime and keep error in desired bound by designing a scheduling program to manage cluster member activities.

Different methods have been proposed for transmitting data inside the clusters [23]. Due to limited geographical area of clusters, in most of WSNs, direct forwarding is still one of the most applicable solutions. Cluster head sends final estimation results to the sink.

In the execution of estimation process for parameter θ in each round, each cluster member node delivers its observations to the FC, based on scheduling program. Therefore variables n_1, n_2, \dots, n_N are the number of delivered samples of nodes 1, 2, ..., N, respectively. Each sample consists of the last observation of the node. BLUE scheme is not efficient for wireless sensor networks because of high communication cost. Therefore instead of sending real-valued observations, using quantization procedure, data volume will be decreased. Different methods of quantization have been proposed [25], for example *uniform randomized quantizer* [26]. Quantizers are applicable for noises with different Probability Density Function (*pdf*). They finally generate unbiased message function. It is worth mentioning that, in this paper we assume that environmental noises corrupt data only in observation phase.

Different wireless sensor network applications have various limitations with respect to the error. Depended on the application requirements, different level of error is acceptable. In proposed algorithm, *Error Bound of Estimation (EBE)*, is an essential parameter. User can determines EBE based on application characteristics. In other words, experimenter must specify a desired bound on the error of estimation, called γ , and associated confidence level, $(1 - \alpha)$. With respect to the EBE, sample size required to estimate parameter θ is determined. Equation (4) uses BLUE estimator, $\bar{\theta}$, to estimate unknown parameter θ based on interval estimation.

$$P(|\bar{\theta} - \theta| < \gamma \cdot \text{MSE}) > 1 - \alpha \quad (4)$$

γ is determined based on EBE and $(1 - \alpha)$ is the confidence level. If $\gamma = 0.01$ and $\alpha = 0.01$ the concept of the equation (4) is “*the error of estimation for parameter θ is less than 0.01 with at least probability of 0.99*”.

$$\gamma^2 / \left(\frac{z\alpha}{2} \right)^2 \leq \frac{\sum_{i=1}^N n_i \sigma_i^2}{(\sum_{i=1}^N n_i)^2} \quad (6)$$

In equation (6) parameters γ^2 and $z\alpha^2$ are known and the goal is to find desirable values for variables n_1, n_2, \dots, n_N . In other words, equation (6) determines the number of samples required to perform estimation process based on parameters γ and α . Different values of variables n_1, n_2, \dots, n_N are eligible in equation (6). In section 3, the best values will be selected using an optimization problem.

III. SCHEDULING MODEL FOR SENSOR NODES TASKS

One of the most important challenges of the wireless sensor networks is to decrease energy consumption while nodes do their tasks effectively. In this paper WSN's nodes are divided into different independent clusters. Cluster member nodes deliver their observations to the FC using direct links based on the scheduling program provided by FC itself.

In equation (7), ∂ is path loss exponent depending on the channel characteristics. Parameter ∂ is usually set 2. $\omega(b)$ is a function determining the required energy to transmit a b-bit message one meter.

In scheduling program two following concepts should be determined for each round. 1) The nodes mode in round, "on-off", and 2) number of instants delivered to the FC (only for active nodes). With respect to the points mentioned before, matrixes S and AC have been considered. S is a $N \times R$ matrix. N is number of cluster member nodes, R is number of performed rounds during network lifetime. All the rounds have the same length, T time units. $S(i, j)$ element from matrix S determines the number of instants provided by j^{th} node in i^{th} round for the FC. Elements of AC matrix determine whether nodes are active or not in a round. If $AC(i, j) = 1$, it means that j^{th} node is active in i^{th} round and If $AC(i, j) = 0$, it means that j^{th} node is inactive in i^{th} round.

In equations (9)-(12), optimization function that finds the best values for elements of the AC and S matrixes are presented.

$$\text{Min } F = \sum_{i=1}^R \sum_{j=1}^N S(i, j). E(j) - R \quad (9)$$

$$S.T: \forall j \in N, \sum_{i=1}^R (S(i, j). E(j)) + (E_{ac}. AC(i, j)) < E_{pri} \quad (10)$$

$$S.T: \forall i \in R, \forall j \in N, AC(i, j) = \left\lceil \frac{S(i, j)}{\sum_{k=1}^N S(i, k)} \right\rceil \quad (11)$$

$$S.T: \forall i \in R, \left(\frac{\gamma^2}{\left(\frac{z\alpha}{2} \right)^2} \right) \leq \left(\frac{\sum_{j=1}^N S(i, j). \sigma_i^2}{(\sum_{j=1}^N S(i, j))^2} \right) \quad (12)$$

The goal of the optimization problem is to minimize function F . First component of function F is " $\sum_{i=1}^R \sum_{j=1}^N S(i, j). E(j)$ ". It declares the total energy

consumption of all network nodes in all rounds. It has positive coefficient, therefore it should be minimized. Second component of function F is R . It declares number of performed rounds. It has negative coefficient, therefore it should be maximized. Network lifetime is equal to " $R \times T$ ".

Equation (10), determines that each node can consume energy at most E_{pri} units. E_{pri} is initial energy of the node. Node energy consumption is consists of two parts. First part discusses that node consumes energy proportional to number of sent messages and its relative distance to the FC. Second part is relevant to the node's energy consumption due to being active in the rounds. E_{ac} is amount of energy consumed by the node in each round when it is active (without considering sent messages). With respect to the equation (10), optimization problem tries to determine nodes mode ("on-off") in order to maximize network lifetime. In inactive "off" mode node's energy consumption is very low, and therefore we have ignored it. Equation (11) presents relation between S and AC matrixes. If $S(i, j)$ has non zero value, $AC(i, j)$ will be 1, otherwise, $AC(i, j)$ will be 0.

Equation (12) is derived from equation (6) by only changing variables. It is essential condition about EBE. As discussed before, the least number of instants in each round is calculated based on EBE. Also equation (12) is essential about determining network lifetime. In a round, if cluster member nodes have no sufficient residual energy in order to satisfy EBE, practically network lifetime is over.

IV. EFFICIENCY EVALUATION

Before In this section, we present some simulation results to compare efficiency of the proposed algorithm and following heuristic methods:

1) Heuristic 1: in each round, nodes send samples to the sink depending on their remaining energy. Assume that, node i has e_i unit remaining energy, then the number of samples provided by node i is " $\left(\frac{e_i}{\sum_{k=1}^N e_k} \right) \times N_b$ ". N_b is the number of samples determined using BLUE estimator [24] besides degree of certainty [28].

2) Heuristic 2: in each round, nodes provide samples equally, in other words all the nodes participate in estimation process with uniform energy scheduling.

In considered scenarios, different number of network nodes has been deployed. The noise variance σ_k^2 and the initial energy source E for all the nodes are the same. We consider all the variances the same, but this means not the observation error is the same in all nodes. It means that average observation error in all the nodes is the same, but in each individual sample different amount of error exists.

Figure 3 shows the network lifetime achieved by proposed algorithm under different values of γ (we call it G, in figure 3). As discussed in section 2-1, γ is determined based on EBE. This means that user can determine desirable error bound by determining γ . The bigger values of γ causes more flexible estimation process with respect to EBE. As noted before, we have estimated unknown parameter θ , using degree of certainty. Therefore, when estimation

process precision is more flexible (bigger values for parameter γ is acceptable), by collecting lower number of samples, desirable precision is achievable. Lower number of samples leads to average network lifetime extension.

Horizontal axis in figure 3 shows parameter W which is calculated based on parameters $\frac{z\alpha}{2}$ and γ . Based on equation (12), we can easily define relationship between W and the two $\frac{z\alpha}{2}$ and γ parameters (see equation (16)).

$$W = \gamma^2 / \left(\frac{z\alpha}{2} \right)^2 \quad (16)$$

As you can see in figure 3, network lifetime becomes larger for bigger values of γ . This is because of lower number of samples. It is worth mentioning that the main goal of proposed algorithm is to prolong network lifetime by scheduling nodes activity with respect to desirable estimation precision.

Figure 4 shows network lifetime achieved by proposed algorithm, heuristic 1 and heuristic 2 methods versus parameter W . In figure 4, results have been gained by a network with $N=10$ sensor nodes. Parameters W is described in details before in this section. As clear in figure 4, proposed algorithm is more efficient rather than the two other methods. The lower values of parameter W , the more difference between proposed algorithm and heuristic 1 is. But for higher values of parameter W , efficiency of algorithms is closer. In real world, big value for W is not applicable.

V. CONCLUSION

In this paper we have proposed a task scheduling method for estimation in wireless sensor networks, which is rarely addressed in the literature. We consider the distributed estimation in energy-limited wireless sensor networks while our main goal is to maximize network lifetime. From the application perspective, the estimation task cycles which accomplished before the time network becomes nonfunctional is considered as network lifetime. We have used linear programming in order to find the best possible scheduling program, which can be easily solved by any LP solver. Task scheduling program determines the number of samples which should be provided by each sensor node in each round. In this paper, we have assumed that the observation noises among different sensors are independent and mean zero Gaussian random variables and the channels from the local sensors to the fusion center are error free. For the future work, we have planned to extend the optimization problem for multihop wireless sensor networks. We have simulated proposed protocol using Opnet simulator; results confirm that proposed protocol manages to achieve its goals.

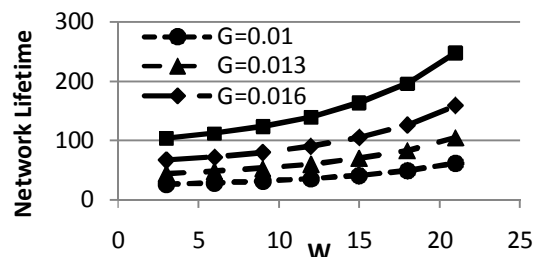


Figure 3, Network lifetime versus Parameter W for different values of G

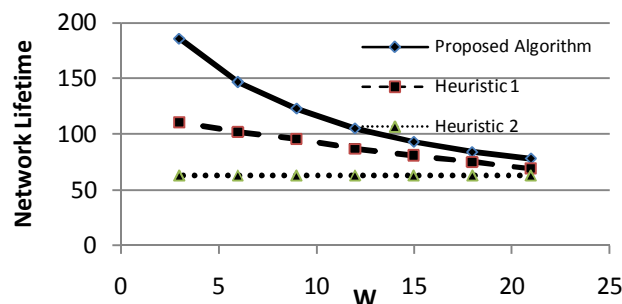


Figure 4, Network lifetime versus parameter W for all the algorithms

VI. REFERENCE

- [1] M. Tubaishat, S. Madria, "Sensor Networks: An Overview." IEEE POTENTIALS April/ May, pp 23, 2003
- [2] S. Kumar, F. Zao, and D. Shepherd, Eds., IEEE Signal Processing Magazine (Special Issue on Collaborative Information Processing), March. 2002, vol 19.
- [3] E. Ayanoglu, "On optimal quantization of noisy sources," IEEE Trans. Inf. Theory, vol. 36, no. 6, pp. 1450–1452, Nov. 1990.
- [4] W. Lam and A. Reibman, "Quantizer design for decentralized systems with communication constraints," IEEE Trans. Commun., vol. 41, pp. 1602–1605, Aug. 1993.
- [5] T. C. Aysal, K. E. Barner, "Constrained Decentralized Estimation Over Noisy Channels for Sensor Networks", IEEE TRANSACTIONS ON SIGNAL PROCESSING, VOL. 56, NO. 4, pp. 1398-1410, APRIL 2008
- [6] J.-J. Xiao, A. Ribeiro, Z.-Q. Luo, and G. B. Giannakis, "Distributed compression-estimation using wireless sensor networks," IEEE Signal Processing Magazine, vol. 23, no. 4, pp. 27–41, July 2006.
- [7] I. F. Akyildiz, W. Su, Y. Sankarsubramaniam, and E. Cayirci, "Wireless sensor networks: A survey," Computer Netw., vol. 38, pp. 393–422, Mar. 2002.
- [8] Z.-Q. Luo, "Universal Decentralized Estimation in a Bandwidth Constrained Sensor Network", IEEE TRANSACTIONS ON INFORMATION THEORY, VOL. 51, NO. 6, JUNE 2005
- [9] J.-H. Chang, L. Tassiulas, "Maximum Lifetime Routing in Wireless Sensor Networks", IEEE/ACM TRANSACTIONS ON NETWORKING, VOL. 12, NO. 4, AUGUST 2004
- [10] I. F. Akyildiz, W. Su, W. Sankarasubramaniam, and E. Cayirci, "A survey on sensor networks," IEEE Communication magazine, pp. 102–114, 2002.
- [11] J.-J. Xiao, A. Ribeiro, Z.-Q. Luo, and G. B. Giannakis, "Distributed compression-estimation using wireless sensor networks," IEEE Signal Process. Mag., vol. 23, no. 4, pp. 27–41, Jul. 2006.
- [12] A. Ribeiro, G. B. Giannakis, "Bandwidth-Constrained Distributed Estimation for Wireless Sensor Networks—Part II: Unknown

- Probability Density Function", IEEE TRANSACTIONS ON SIGNAL PROCESSING, VOL. 54, NO. 7, JULY 2006
- [13] S. Cui, A. Goldsmith, A. Bahai, "Joint modulation and multiple access optimization under energy constraints," Proc. IEEE Global Telecomm. Conf., Dallas, Texas, Dec. 2004, pp. 151–155.
- [14] A. S. Willsky, M. Bello, D. A. Castanon, B. C. Levy, and G. Verghese, "Combining and updating of local estimates and regional maps along sets of one-dimensional tracks," IEEE Trans. Autom. Control, vol. AC-27, pp. 799–813, Aug. 1982.
- [15] Z. Chair and P. K. Varshney, "Distributed bayesian hypothesis testing with distributed data fusion," IEEE Trans. Syst., Man, Cybern., vol. 18, pp. 695–699, Sep.–Oct. 1988.
- [16] J-J Xiao, S. Cui, Z-Q Luo, A. J. Goldsmith, "Power Scheduling of Universal Decentralized Estimation in Sensor Networks", IEEE TRANSACTIONS ON SIGNAL PROCESSING, VOL. 54, NO. 2, FEBRUARY 2006
- [17] J. LI, G. AlRegib, "Rate-Constrained Distributed Estimation in Wireless Sensor Networks", IEEE TRANSACTIONS ON SIGNAL PROCESSING, VOL. 55, NO. 5, MAY 2007
- [18] D. Dardari, A. Conti, C. Burrati, R. Verdone, "Mathematical Evaluation of Environmental Monitoring Estimation Error through Energy-Efficient Wireless Sensor Networks", IEEE TRANSACTIONS ON MOBILE COMPUTING, VOL. 6, NO. 7, JULY 2007
- [19] H. Chen, "Performance-Energy Tradeoffs for Decentralized Estimation in a Multihop Sensor Network", IEEE SENSORS JOURNAL, VOL. 10, NO. 8, AUGUST 2010
- [20] J. LI, G. AlRegib, "Network Lifetime Maximization for Estimation in Multihop Wireless Sensor Networks", IEEE TRANSACTIONS ON SIGNAL PROCESSING, VOL. 57, NO. 7, JULY 2009
- [21] C-T Cheng, C. K. Tse, F. C. M. Lao, "An Energy-Aware Scheduling Scheme for Wireless Sensor Networks", IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY, VOL. 59, NO. 7, SEPTEMBER 2010
- [22] M. L. Sichitiu, "Cross-Layer Scheduling for Power Efficiency in Wireless Sensor Networks," Proceedings of INFOCOM 2004, Hong Kong, Vol2, pp.266-276, March 2004.
- [23] Qiangfeng, et al, "Routing Protocols for Sensor Networks.", 1st IEEE Consumer Communications and Networking Conference, 2004.
- [24] S. Kay, Fundamentals of Statistical Signal Processing: Estimation Theory. Englewood Cliffs, NJ: Prentice–Hall, 1993.
- [25] H. Papadopoulos, G. Wornell, and A. Oppenheim, "Sequential signal encoding from noisy measurements using quantizers with dynamic bias control," IEEE Trans. Inf. Theory, vol. 47, pp. 978–1002, Mar. 2001.
- [26] J-J. Xiao and Z-Q. Luo, "Universal decentralized estimation in an inhomogeneous sensing environment," IEEE Trans. Inf. Theory, vol. 51, no. 10, pp. 3564–3575, Oct. 2005.
- [27] J. E. Beasley, "Advances in Linear and Integer Programming", Oxford Science press, 1996.
- [28] W. Mendenhall, D. D. Wackerly, R. L. Scheaffer, "Mathematical statistics with applications", PWS-KENT, 1989.