Journal of Network and Computer Applications 57 (2015) 142–155

Contents lists available at ScienceDirect

Journal of Network and Computer Applications

journal homepage: www.elsevier.com/locate/jnca

Efficient data collecting and target parameter estimation in wireless sensor networks

Amir Hossein Mohajerzadeh, Mohammad Hossein Yaghmaee, Afsane Zahmatkesh

Computer Department, Ferdowsi University of Mashhad, Iran

ARTICLE INFO

Article history:
Received 30 June 2013
Received in revised form
17 October 2014
Accepted 18 June 2015
Available online 31 July 2015

Keywords:
Estimation
Optimization problem
Random variable target parameter
Scheduling
Single hop and multihop routing
WSN

ABSTRACT

Nowadays, one of the attractive topics in Wireless Sensor Networks (WSN) is Estimation. In WSN, sensor nodes collect locally processed data and then send them to a Fusion Center (FC). The FC collects data to produce a final estimation of the observed target parameter. In this paper, a distributed estimation method in a WSN will be proposed. The network consists of a FC and a number of network member nodes. The proposed method addresses two main challenges: (1) estimating target parameter according to a required application's precision and (2) collecting and forwarding data to the FC. The main goal of the proposed method is to prolong the network lifetime and to estimate the target parameter efficiently. The FC estimates unknown target parameter using the collected data. This work considers the various types of target parameters based on the environment characteristics; besides, using the credible confidence interval, the required sample size for the estimating target parameter with desirable precision is calculated. According to the requirements of estimation process, this paper will propose efficient scheduling/routing algorithms to collect data. Therefore, by the joint scheduling and routing algorithms, the proposed estimation method experiences high efficiency.

1. Introduction

Wireless Sensor Networks (WSN) consists of sensor nodes which are distributed geographically. These sensor nodes perform their tasks as an integrated system. The sensor nodes have many various constraints such as: energy sources, computational power, storage capacity, etc. (Tubaishat and Madria, 2003). During the past few years, WSN have found many applications including environment monitoring, healthcare, battlefield surveillance, home automation and so on (Rezaee et al., 2014). Recently, WSN has been applied for distributed estimation, distributed detection and tracking.

In this paper, using a set of observations provided by distributed sensor nodes, the distributed estimation of unknown target parameter is performed. In the distributed estimation, each node delivers a subset of observations from the environment to a central node called *Fusion Center* (FC), directly or indirectly. The FC is responsible to reconstruct the underlying physical phenomenon based on input data gathering from the sensor measurements. The estimation literature attracts a considerable attention in computer networks (Ayanoglu, 1990; Lam and Reibman, 1993); besides, today it becomes one of the attractive topics in signal processing in WSNs (Aysal and Barner, 2008; Huang and Hua, 2012). In WSN, sensor nodes collect locally processed data and then send them to a FC. The FC collects data to produce the final estimation of the observed target parameter. Most of the previous studies on estimation in WSNs (Akyildiz et al., 2002; Luo, 2005) assume that the joint distribution of sensor's observations is known and the real observation messages can be delivered from the sensors to the FC without distortion.

Resource constraint is a common problem for all WSNs with different applications. Sensor nodes have only small batteries where replacement can be costly and complicated. Compared to sensing, communication is the most energy-consuming operation of the sensor nodes. As a result, in order to extend lifetime, reducing the communications between sensor nodes is one of the essential requirements of WSNs. Various methods have been proposed to increase network lifetime and efficiency in WSNs (Chang and Tassiulas, 2004).

Respecting the WSN's characteristics (Akyildiz et al., 2002), different distributed estimation algorithms have been proposed (Chaudhary and Vandendorpe, 2012; Yuh-Ren and Cheng-Ju, 2011). All of them address design and implementation methods in order to digitize transmitted signal into several bits. The problem of decentralized estimation has been studied, in distributed control (Cui et al., 2004), target tracking (Roseveare and Natarajan, 2012) and data fusion (Behbahani et al., 2012). An optimal power scheduling problem for the decentralized estimation of a noise-corrupted

E-mail addresses: mohajerzadeh@um.ac.ir (A.H. Mohajerzadeh), yaghmaee@ieee.org (M.H. Yaghmaee), zahmatkesh.afsane@stu-mail.um.ac.ir (A. Zahmatkesh).

http://dx.doi.org/10.1016/j.jnca.2015.06.005
1084-8045/

deterministic signal in an inhomogeneous sensor network is proposed in Xiao et al. (2006); moreover, the optimal quantization and transmission power level at local sensors is determined in order to minimize the total transmission power, while ensuring a given Mean Square Error (MSE) performance. LI and AlRegib (2007) study the optimal tradeoff between the number of active sensors and the quantization bit rate for each active sensor to minimize the estimation MSE. Dardari et al. (2007) investigate the estimation of a scalar field over a bidimensional scenario through a WSN with energy constraint. This paper provides a mathematical framework to analyze the independent aspects of WSN's communication protocols and signal processing design. Chen (2010) looks over the performance-energy tradeoff for distributed estimation in a WSN; in addition, the Best Linear Unbiased Estimation (BLUE) is utilized to estimate the observed phenomena. Like other abundant works, Chen (2010) and LI and AlRegib (2009) use 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 area with the minimum possible number of sensor nodes. In this paper, in order to extend the network lifetime, the main objective is designing an efficient scheduling scheme to control sensing and communication tasks in the sensor nodes. Basically, scheduling is classified into four main categories (Cheng et al., 2010) including: "*always on*", "*random on–off*", "*adaptive on–off*" and "*periodic on–off*". In the proposed algorithm, the scheduling program is an adaptive on–off scheme; in addition, the FC is responsible to create a total scheduling program so the other sensors only follow this program (Sichitiu, 2004). Routing in hierarchical networks consists of two main parts: (1) intra-cluster routing and (2) inter-cluster routing. Delivering the sensor node's observation to the cluster head and communicating with other cluster heads to send data to the sink is done by intra-cluster routing. In intra-cluster routing, because of the small size of cluster's area, both direct (single hop) and indirect (multihop) forwarding are applicable. In direct routing, nodes deliver their data to the cluster head or sink with using the intermediate nodes, while in multihop routing, nodes cooperate with each other to transfer the collected data to the base station. In this paper, a linear programming (LP) is used to find the best schedule for direct routing; in addition, a nonlinear programming (NLP) is used to select multihop routes between each network node and the FC. Providing fairness is one of the main subjects in the proposed NLP. One of the open research areas in WSN's field is joint routing and scheduling algorithm. Some of the prior works schedule nodes tasks depended on MAC layer protocol (Liu et al., 2010; Wu et al., 2010). In Incel et al. (2012), a number of various methods are proposed which used realistic simulation models under the many-to-one communication paradigm known as convergecast; moreover, the time scheduling on a single frequency channel is considered to minimize the number of the required time slots to complete a convergecast. This work combines scheduling with transmission power control to reduce the effects of interference. Finally, Incel et al. (2012) constructs the degree-constrained and the capacitated minimal spanning trees. Some of the previous works consider specific applications. For instance, a combined scheduling and routing mechanism is proposed to provide a deterministic guarantee of end-to-end deadlines leveraging (Ryu et al., 2009). It regards a battlefield at the border of South and North Korea. A lifetime-aware routing with considering desired sensing spatial coverage is proposed in Karkvandi et al. (2011). Cohen and Kapchits (2009) suggest a routing and scheduling algorithm for maximizing the lifetime of mesh sensor networks, while it guarantees an upper bound on the end-to-end delay. In most of the prior works, energy-efficient routing and sleep scheduling are considered as two separate tasks. They assume that one component is pre-given, whereas they optimize the other one (Wang et al., 2008). In this paper, in order to achieve the highest efficiency, the joint routing and scheduling algorithms are performed.

The reminder of this essay is organized as follows. Section 2 introduces the system model for estimating unknown target parameter, and interval estimation, too. In Section 3, the network intra-cluster routing and scheduling for collecting estimation process data is formulated as optimization problem. In order to demonstrate the performance of the proposed method, simulation results have been presented in Section 4. Finally, Section 5 concludes the paper.

2. Problem formulation

This paper considers the hierarchical WSN's topology which consists of different clusters. Each cluster contains N sensor nodes and a FC, which are designed to cooperate with each other to estimate an unknown parameter θ; moreover, cluster head acts as a FC. The FC is responsible to reconstruct the underlying physical phenomenon based on input data gathering from the sensor measurements. The FC is a role which possibly any sensor node can perform its tasks. In several papers, sink is considered to be as the FC, but in the current study cluster head is FC. In each cluster, estimation process is performed independently, and the result would be transferred to the sink. Based on the information of the user about the target parameter, two different estimation models will be proposed. The estimation model also needs a required precision which should be provided by the user. The member of each cluster monitors the events, quantizes and transmits the collected data to the FC. Using an optimization problem, data collection (consists of routing and scheduling) is performed according to the estimation model, the network nodes position and remained energy. Then, the FC processes the data and applying inter-cluster routing algorithm, sends estimation results to the sink. Sink is able to deliver final value of target parameter θ to the user. Each part of the estimation method will be explained in detail in the rest of the paper.

In this paper, firstly, the proposed estimation models are explained. Estimation model is fully dependent on the user information about the target parameter. If the user does not have any information about the target parameter, the estimation model explained in Section 2.1 should be applied. In other mode, if the target parameter distribution function is Normal, the estimation model explained in Section 2.2 should be utilized.

2.1. Estimation model for the target parameter possessing unknown distribution function

In this article, the hierarchical WSN is considered. Each cluster member observes the event, quantizes and transmits its collected information to the FC. The final estimation is made based on all the received messages from cluster members. The observations are corrupted by additive noise and described by using the following equation:

$$x_{ki} = \theta + \varepsilon_{ki}; \quad k = 1, 2, \ldots, N; \quad i = 1, 2, \ldots, n_k \tag{1}$$

x_{ki} is the ith observation of the sensor node k. Depended on the problem conditions, each sensor node is able to send many samples (observations). The sample size provided by node k is shown by n_k. Sensor noise variables, ε_{ki}, are considered to be independent, mean zero Gaussian random variables with $\mathrm{var}(\varepsilon_k) = \sigma_k^2 (k = 1, 2, \ldots, N)$. θ is the parameter to be estimated. Since WSNs have severe bandwidth and energy limitation, at the first step of estimation process, the real valued analog observation is quantized locally into an unbiased discrete message in each