ENERGY-EFFICIENT DECENTRALIZED EVENT DETECTION IN LARGE-SCALE WIRELESS SENSOR NETWORKS

Qing Ling* Fanzi Zeng† Zhi Tian‡

* Department of Automation, University of Science and Technology of China, Hefei, Anhui, P. R. China
† School of Telecommunications, Hunan University, Changsha, Hunan, P. R. China
‡ Department of Electrical and Computer Engineering, Michigan Technological University, Houghton, Michigan, USA

ABSTRACT

This paper addresses the problem of decentralized event detection in large-scale wireless sensor networks (WSNs). Compared with centralized or hierarchical solutions, decentralized algorithms are superior in terms of scalability and robustness. However, traditional decentralized optimization tools, such as consensus optimization, entail intensive information exchange of high-dimensional decision vectors and multipliers. This paper exploits the phenomenon of limited influence, namely, the influence of one event only affects its neighboring area. For this scenario, we let each sensor make decisions for its local area rather than for the entire network, and individual decisions seek to collaboratively reach the global optimum through iterative local communications at low network costs. An optimal solution based on the alternating direction method of multipliers (ADMM) is developed. To further reduce the network communication load, we also propose a heuristic decentralized linear programming (DLP) algorithm, which is shown to be efficient via simulations.

Index Terms— Wireless sensor network (WSN), decentralized event detection, alternating direction method of multipliers (ADMM), decentralized linear programming (DLP)

1. INTRODUCTION

For monitoring and control applications using large-scale wireless sensor networks (WSNs), primary considerations in information processing are energy efficiency, scalability and robustness. To address these issues, decentralized in-network processing has attracted considerable research interest in recent years [1, 2, 3, 4, 5]. In the absence of infrastructure or centralized control, sensors autonomously exchange information with neighboring nodes, aiming at optimizing a global objective function. Herein one of the most important optimization tools is consensus optimization [3, 4, 5]. Sensors optimize the global objective function individually; meanwhile, extra constraints are appended such that all local decision vectors will consent to an identical result. This augmented formulation is then decomposed and optimized in an iterative and decentralized manner. Under some assumptions on network connectivity, the decentralized consensus optimization approach is able to converge to the global optimum.

In this paper, we study the problem of decentralized event detection in large-scale WSNs. The goal is to identify the existence, locate the positions, and estimate the magnitudes of events or abnormalities in the sensing field. Applications include detecting nuclear radioactive sources, detecting fires in forests, monitoring structural health conditions, etc. A general approach to these problems is to select some sensor points or grid points in the sensing field, and recover the field map on these representative points [6, 7]. Length of the decision vector is hence equal to the number of selected points, while each element of the decision vector represents existence and magnitude of an event at a given position. In this case, the decision vector can be of high dimension for a large-scale network or a large sensing field with high detection resolution. Unfortunately, reaching consensus over the entire decision vector of large size is impractical due to excessive burden on network bandwidth, since sensors need to exchange intermediate decision vectors and Lagrange multipliers in each iteration of the decentralized consensus optimization algorithm.

To overcome the obstacle in large-scale WSNs, this paper develops energy-efficient decentralized algorithms. Different from consensus optimization, we set an alternative objective that is simple yet effective: sensors no longer optimize and consent on the entire decision vector as in the traditional consensus approach; in contrast, we let each sensor optimize its own decision variable corresponding to a scalar element in the decision vector, while communicating with its one-hop neighbors for global optimality. This new design is applicable as long as the influence of the events exhibits certain level of locality, that is, each event only influences its neighboring area in a large sensing field. In many monitoring problems the event sources of interest are indeed localized compared to the large sensing field, such as radioactive sources, fire sources, damages of structures, etc. In this case, the global detection task can be readily decomposed into related local tasks, which obviates the need for global information exchange of the en-
tire sensing field and hence considerably alleviates the burden on network bandwidth. Upon detecting events, sensors can either trigger alarms autonomously or transmit the refined information to a central console.

In the ensuing paper, we first provide an optimization model for detecting local events in the sensing field. Then, we derive two decentralized iterative solutions, in which each sensor optimizes its own decision variable other than the entire decision vector. Last, we provide simulation results to show the effectiveness of the proposed algorithms.

2. PROBLEM FORMULATION

Suppose that a large-scale WSN is deployed in a large two-dimensional area. The network has a set of \( L \) sensors, denoted as \( \mathcal{L} = \{ v_i \}_{i=1}^L \). Sensors have a common communication range \( r_C \). Each sensor can only communicate with its one-hop neighbors within the communication range. The network is connected given the communication range \( r_C \).

The sensing and event detection task is performed periodically. Within each sampling period, multiple events may occur. We make the following assumptions:

(A1): Event sources occur only at some sensor points. When the source of one event coincides with the position of \( v_i \), we denote the amplitude of the event by a nonnegative scalar \( c_i \).

(A2): Each event may influence a nearby region. The influence of a unit-amplitude event at sensor point \( v_i \) on a sensor point \( v_j \) is \( f_{ij} \). We suppose that \( f_{ij} = 0 \) if \( d_{ij} \geq r_C \), where \( d_{ij} \) is the distance between sensors \( v_i \) and \( v_j \).

(A3): The measurement of one sensor is the superposition of the influences of all events plus random noise. Since \( f_{ij} = 0 \) for \( d_{ij} \geq r_C \), the measurement \( b_i \) at sensor \( v_i \) can be written as \( b_i = \sum_{v_j \in N_i \cup v_i} f_{ij} c_j + e_i \), where \( e_i \) is random noise, and \( N_i \) denotes the one-hop neighboring sensors of \( v_i \).

The assumption (A1) selects sensor points as representative points of the sensing field. Hence the resolution of event detection is directly decided by the sensor density. Based on (A1), we can formulate the event-field mapping problem as recovering a nonnegative vector \( c = [c_1, \ldots, c_L]^T \) from the measurements.

The assumption (A2) describes the phenomenon of limited influence of one event on the whole sensing field. For example, in nuclear radioactive detection, the influence of a source decreases polynomially in distance. Similar distance-dependent influence can be observed from events in many practical applications such as fire sources and structural damages. By carefully setting the communication range \( r_C \), it is possible to make \( f_{ij} \approx 0 \) if \( d_{ij} \geq r_C \). This assumption is necessary for designing a simple decentralized algorithm as it eliminates the interrelationship between multi-hop sensors and hence avoids multi-hop communications.

We now introduce a key observation that the decision vector \( c \) is sparse, namely, the number of nonzero elements in \( c \) is much smaller than the vector size \( L \). This prior knowledge holds since event detection for a large-scale WSN is meaningless if the sensing field is full of event sources. Nevertheless, we do allow the influence of these sparse events to span over the large field. Hence, reconstruction of \( c \) boils down to minimizing a sparsity-imposing metric \( ||c||_1 \) that is the \( \ell_1 \) norm of \( c \) [8], subject to measurement constraints. From (A3), two similar formulations arise, one is a linear program that postulates a bounded measurement error \( \theta \) at each sensor:

\[
\begin{align*}
\min_{s.t.} & \quad ||c||_1 \\
& \quad |b_i - \sum_{v_j \in N_i \cup v_i} f_{ij} c_j| \leq \theta, \quad \forall v_i \in \mathcal{L} \\
& \quad c_i \geq 0, \quad \forall v_i \in \mathcal{L}
\end{align*}
\]

and the other is a second-order cone program that confines the total energy of the measurement errors to be lower than \( \epsilon \):

\[
\begin{align*}
\min_{s.t.} & \quad ||c||_1 \\
& \quad \sum_{v_i \in \mathcal{L}} (b_i - \sum_{v_j \in N_i \cup v_i} f_{ij} c_j)^2 \leq \epsilon \\
& \quad c_i \geq 0, \quad \forall v_i \in \mathcal{L}
\end{align*}
\]

Both (1) and (2) incorporate the prior knowledge of sparse events, which is important in alleviating the undesired false alarm rates of otherwise non-sparse solutions produced by general signal recovery approaches such as thresholding or least squares. For exposition simplicity, we only consider the linear program (1) in this paper.

3. DECENTRALIZED ALGORITHMS

This section introduces two decentralized algorithms to (1), an optimal one and a heuristic one.

3.1. An Optimal Decentralized Algorithm

Introducing slack variables \( \{s_{1i}, s_{2i}\} \) for the inequality constraints, (1) can be rewritten as the following linear program:

\[
\begin{align*}
\min_{s.t.} & \quad \sum_{v_i \in \mathcal{L}} c_i \\
& \quad \sum_{v_j \in N_i \cup v_i} f_{ij} c_j - s_{1i} + \theta - b_i = 0, \quad \forall v_i \in \mathcal{L} \\
& \quad \sum_{v_j \in N_i \cup v_i} f_{ij} c_j - s_{2i} - \theta - b_i = 0, \quad \forall v_i \in \mathcal{L} \\
& \quad c_i \geq 0, s_{1i} \geq 0, s_{2i} \leq 0, \quad \forall v_i \in \mathcal{L}
\end{align*}
\]

Based on the alternating direction method of multipliers (ADMM) [9], (3) has an iterative optimal solution. For any \( v_i \in \mathcal{L} \), we update the slack variables \( s_{1i} \) and \( s_{2i} \), Lagrange multipliers \( \gamma_{1i}, \gamma_{2i}, \lambda_{1i}, \) and \( \lambda_{2i}, \) and decision variable \( c_i \) as follows:

\[
\begin{align*}
s_{1i}(t + 1) &= \frac{1}{m_i} [\lambda_{1i}(t)]^T \\
s_{2i}(t + 1) &= -\frac{1}{m_i} [\lambda_{2i}(t)]^{-T}
\end{align*}
\]

\[
\begin{align*}
\gamma_{1i}(t + 1) &= \frac{1}{m_i} \left[ \sum_{v_j \in N_i \cup v_i} f_{ij} c_j(t) - s_{1i}(t + 1) - (b_i - \theta) \right] \\
\gamma_{2i}(t + 1) &= \frac{1}{m_i} \left[ \sum_{v_j \in N_i \cup v_i} f_{ij} c_j(t) - s_{2i}(t + 1) - (b_i + \theta) \right]
\end{align*}
\]

\[
\begin{align*}
\lambda_{1i}(t + 1) &= \lambda_{1i}(t) + \beta \gamma_{1i}(t + 1) \\
\lambda_{2i}(t + 1) &= \lambda_{2i}(t) + \beta \gamma_{2i}(t + 1)
\end{align*}
\]
gi(t +1) = ∑vj ∈Ni∪vi [−βfijγ1j(t +1)− fijλ1j(t +1) − βfijγ2j(t +1)− fijλ2j(t +1)]− 1
hi(t +1) = β ∑vj ∈Ni∪vi f2
j

ci(t +1) = [gi(t +1)/ hi(t +1) + ci(t)]+
(7)

Here [·]+ = max{·,0} and [·]− = min{·,0} are projection operators, and m is the number of sensors inside the communication range of sensor vi, and β is a constant positive coefficient.

According to (4)-(7), each sensor only needs to know the decision variables and Lagrange multipliers of its one-hop neighboring sensors to update its own slack variables, Lagrange multipliers, and decision variable. In summary, we have the following decentralized ADMM solution to (1):

Algorithm 1: ADMM

Step 1: Initialization. Each sensor vi holds a predefined common threshold θ and a constant β, and a measurement bi, ∀vi. It calculates mvi by counting the number of its one-hop neighbors. Initial decision variable, slack variables, and Lagrange multipliers are all set to 0.

Step 2: Information Exchange. At time t + 1, each sensor vi broadcasts its current decision variable ci(t), Lagrange multipliers λ1i(t) and λ2i(t), γ1i(t) and γ2i(t) to its one-hop neighboring sensors vj ∈ Nvi.

Step 3: Parameter Update. Each sensor vi updates its slack variables s1i(t +1) and s2i(t +1) via (4), its multipliers λ1i(t +1) and λ2i(t +1) via (5), its multipliers γ1i(t +1) and γ2i(t +1) via (6), and its decision variable ci(t +1) via (7).

Step 4: Iterative Optimization. Repeat Step 2 and Step 3 until reaching convergence.

It shall be noted that the ADMM technique has been applied to decentralized optimization, but for a different formulation where each sensor optimizes a local copy of the entire decision vector and consents on the vector through iterations [3]. In this paper, the distinct feature of the proposed algorithm is that by exploiting the phenomenon of limited influence, each sensor vi only needs to optimize its own scalable decision variable ci, not the entire vector c of large size, yet still iteratively converges to global optimality in decision making. In large-scale WSNs, this approach remarkably alleviates the communication burden of information exchange.

3.2. A Heuristic Decentralized Algorithm

The ADMM solution in Section 3.1 can be shown to be optimal for (1). It has lower computational complexity and communication load than conventional ADMM, but the convergence rate of the ADMM family generally can be slow, raising the communication costs of information exchange over many iterations. Moreover, sensors need to exchange not only decision variables but also Lagrange multipliers. Next we develop a heuristic decentralized linear programming (DLP) algorithm to solve (1) in a more energy-efficient manner.

The inspiration for our low-cost decentralized processing comes from [10], where Tseng proposed a parallel computing framework for optimization of linear programs satisfying a certain diagonal dominance conditions. Therein, the objective function is divided into uncoupled terms, and the constraints are assigned to different processors in parallel. This framework stimulates us to assign the objective function and constraints in (1) to individual sensors, as follows:

min ci, s.t. [bi − fii ci − ∑vj∈Nvi fiji cj] ≤ θ, ci ≥ 0.
(8)

Solution to Eq. (8) is ci = [bi − θ − ∑vj∈Nvi fiji cj] + fii, or null if bi + θ − ∑vj∈Nvi fiji cj < 0. Therefore we have the following heuristic DLP algorithm to solve (1):

Algorithm 2: DLP

Step 1: Initialization. Each sensor vi holds a predefined common threshold θ and a measurement bi. The initial decision variable is set as 0.

Step 2: Information Exchange. At time t + 1, each sensor vi broadcasts its current decision variable ci(t) to its one-hop neighboring sensors vj ∈ Nvi.

Step 3: Parameter Update. Each sensor vi updates decision variable via ci(t +1) = [bi−θ−∑vj∈Nvi fiji cj(t)] + fii.

Step 4: Iterative Optimization. Repeat Step 2 and Step 3 until reaching convergence.

Evidently, the DLP is simple to compute, and consumes minimal communication energy in information exchange during each iteration. Although proving convergence is nontrivial, the results in [10] alludes to the conjecture that the DLP can be optimal when {fij} meet certain diagonal dominance condition. As testified by simulation cases discussed next, the DLP not only performs well, but also exhibits much faster convergence behavior than the ADMM technique.

4. SIMULATION RESULTS

In the simulations, we divide the sensing field into an N × N lattice, and deploy L = N2 sensors at the grid points. The distance between two neighboring sensors is r, such that the lattice can be represented by sensor positions {(xr, yr), x, y = 1, 2, ..., N}. Suppose that events occur at several grid points. The influence from a unit-magnitude event at point vi to an energy-detection sensor at point vj is assumed to be fij = exp(−dij/σ2), in which σ is a known parameter. This Gaussian-shaped influence function approximately satisfies the condition of limited influence for properly chosen
communication range \( r_C \). For example, when \( \sigma = r \), setting \( r_C \) to be slightly larger than \( \sqrt{2r} \) makes \( f_{ij} \approx 0 \) if \( d_{ij} \geq r_C \). In this case one sensor can directly communicate with 8 neighboring sensors.

Let \( N = 10 \) and \( L = 100 \). Two events, with magnitude 1 and 0.5, occur at \((3r, 5r)\) and \((5r, 5r)\). The outputs are polluted by uniform random noises, ranging from -0.01 to 0.01. We set \( \theta = 0.05 \) and \( \beta = 1 \). The two decentralized algorithms, ADMM and DLP, both provide sparse solutions and accurately locate the events. Convergence behavior of the two algorithms is depicted in Figs. 1 and 2, respectively. The ADMM algorithm converges after 100 iterations, while the DLP algorithm converges after only 5 iterations.

Now we compare the communication load of the proposed decentralized algorithms and the centralized event detection scheme. In each iteration of the ADMM algorithm, each sensor needs to broadcast 1 intermediate decision variable and 4 Lagrange multipliers to neighbors. The communication load of each sensor is hence 5 times its number of iterations; that is 100 for sensors near to the events. The DLP algorithm achieves faster convergence, and each sensor needs to broadcast only 1 intermediate decision variable. Hence the resulting average communication load is 5, which is a remarkable improvement over the ADMM algorithm.

More importantly, the average communication load of the decentralized algorithms is almost invariant as the WSN scale \( L \) changes. As a comparison, we consider a centralized grid network, in which sensors need to transmit measurements to a fusion center located at sensor \((1r, 1r)\) via multi-hop communications. The average communication load is near to \( N/2 \), proportional to the square root of network size \( \sqrt{L} \).

Furthermore, the decentralized information processing scheme also improves robustness of the network. Due to the simple message passing protocol, the network is routing-free. Failures of several sensors only influence themselves and their neighbors, but not the entire network.

Finally, for energy- and bandwidth-constrained WSNs of large scale, we emphasize again the importance of optimizing individual decision variables by exploiting the property of limited influence. Otherwise under the general consensus optimization settings, the communication load of a single sensor in each iteration is proportional to the size of each decision vector, which is the same as the number of sensors. For practical applications of the decentralized event detection algorithms, readers are referred to [11].

5. REFERENCES