Monitoring of Diffusion Processes with PDE Models in Wireless Sensor Networks

Lorenzo A. Rossi, Bhaskar Krishnamachari and C.-C. Jay Kuo

Department of Electrical Engineering, University of Southern California
Los Angeles, CA 90089-2564, USA.
E-mails: lrossi@usc.edu, bkrishna@usc.edu, cckuo@sipi.usc.edu.

ABSTRACT

The monitoring of a diffuse process, such as the propagation of a toxic gas in an area, using the partial differential equation (PED) model via autonomous wireless sensor networks is studied in this research. Sensor nodes update the base station with their estimates of PDE model parameters rather than raw sensor measurements. Then, the base station can reconstruct the phenomenon through model parameters and initial and boundary conditions. In-network processing techniques to estimate the PDE coefficients are presented. A scheme is presented to provide a hybrid combination of decision and data fusion to find a proper tradeoff between estimate accuracy and energy efficiency. Besides, several open issues in this research context, such as identifiability of parameters, monitoring of time varying boundary conditions and unknown sources, are discussed.

1. INTRODUCTION

Wireless sensor networks provide a very promising technology for security and defense applications, thanks to their potential to be deployed in a hostile environment at low cost for autonomous and intelligent measurements and monitoring. An important problem in this application context is the monitoring of a diffusion phenomenon, such as a cloud of toxic agents. Data gathering is a fundamental task for such a problem. Its purpose is to determine the optimal way (in the sense of least energy consumption) to transfer node measurements to a remote base station (BS). A typical assumption is that all nodes sample the sensor field uniformly in time and generate a packet for each round of measurements. All packets are delivered to the BS through various aggregation strategies [1].

The dense spatial-temporal signal sampling and transmission is not energy efficient due to strong spatial and temporal correlations of measured and transmitted data. In many real world applications, such correlations can be characterized by relatively simple mathematical models, *e.g.* one or a couple of partial differential equations (PDE's). Parabolic PDE's, called also diffusion equations, can be used to describe the gas diffusion phenomenon in a plant and/or an open environment. If this is the case, given proper initial and boundary conditions, the knowledge of PDE coefficients can be used to reconstruct data associated with the phenomenon.

In this work, we assume the existence of an underlying diffusion PDE model for the physical phenomenon being monitored by the sensor networks (e.g. propagating gasses). Then, the base station would be able to predict the evolution of the phenomenon using the model without the need of keep receiving raw measurements from the sensor field. Basically, it should gather the initial and boundary conditions as well as updated PDE parameters so that the dense data field can be reconstructed from the numerical solution of the PDE problem. It is apparent that transmitting a few PDE coefficients costs much less than sending the densely sampled data field, since these coefficients have a much slower variation rate in the spatial and temporal domains.

One of the most important problems in the aforementioned framework is the identification of unknown parameters through measured data of sensor nodes. For robust estimation of these parameters, it may be worthwhile to exchange sampled data among neighboring nodes at the expense of some communication cost. This is the data fusion approach. On the other hand, it is also possible to estimate the model parameters based on measurements at each sensor node and then exchange the estimated model parameters for further processing and refinement. This corresponds to the decision fusion approach. The latter may provide poorer results, expecially in the case of spatially varying coefficients, at a lower communication cost.

In this research, we provide a pioneering investigation on distributed identification of the PDE parameters. The technique is suitable for data as well as decision fusion implementations. In order to combine the advantages of both strategies, we also develop a hybrid data and decision fusion framework. Besides proposing and evaluating an approach for parameter identification, we discuss several relevant open issues in this research such as evaluating whether a given subset of sensor nodes is able identify the parameters; tracking time varying boundary conditions; localizing and monitoring eventual unknown sources in the field; determining a proper model for estimation; understanding the propagation phenomenon of data from sensor nodes to the base station.

The rest of this paper is organized as follows. Section 2 outlines some centralized (non-sensor network) approaches for the identification of coefficients of diffusion models. The problem formulation and the proposed hybrid data/decision fusion scheme are given in Section 3. Extensions of the method are explained in 4. Open issues are discussed in Section 6. Simulation results are shown in Section 5.

2. REVIEW OF RELATED WORK

A large amount of research on wireless sensor networks has been concerned with the data gathering problem [1,2]. As mentioned in the previous section, the general assumption under this context is that all nodes sample the sensor field uniformly in time, and then send raw measurements and/or aggregated information to the base station (BS) for further processing. Data aggregation is performed so as to minimize the overall energy demanded for communication among senor nodes and their base station.

In a general setting, there are no assumptions made about the correlation of the measured data. However, in many applications, there do exist some correlation of measured data, which can be exploited to develop a more effective data gathering system. One such example is the monitoring of the diffusion of some matter (e.g. the propagation of toxic gas). The sensor network sends to the BS the spatio-temporal samples of the phenomenon of interest, which have strong correlations. The exploitation of such correlations by means of a physical model, which corresponds to the diffusion PDE in this research, has the potential to reduce the amount of data to be transfered to the BS. Hence, a key problem in model-based data gathering is the identification of model parameters (the diffusion coefficients) via in-network distributed processing. Centralized algorithms for the identification of PDE coefficients have been presented in the literature (see the references in the next paragraph). However, to the best of our knowledge, there is little research on a distributed approach to achieve this goal.

The identification and control of systems modeled by PDE's have been studied by researchers in the fields of automatic control and applied mathematics [3], [4]. In this context, systems represented by PDE's are referred to as *infinite dimensional* or *distributed* systems* [5]. Moreover, a *lumped system* indicates a system that has discrete components (e.g. the discretized version of an infinite dimensional system). Here, we follow the terminology of system theory, but reserve the term *distributed algorithms* to indicate algorithms using multi-point independent computations and terms *distributed computation* and *distributed knowledge* as their meanings as commonly used in the sensor network context.

Several adaptive algorithms for the identification of parameters of infinite dimensional systems in the continuous time-space domain have been presented [3], [4], [6]. One major application of this study is the automatic control of PDE-modeled plants. Adaptive control schemes that rely on the estimation system states and unknown parameters of a plant were proposed by Baumeister *et al.* [3] and Demetriou *et al.* [6]. Besides, they developed theory on the use of a finite dimensional approximation to the infinite dimensional estimator. Orlov and Bentsman [4] studied the problem of identifying an infinite dimensional system with spatial-varying parameters. They derived some constructive identifiability conditions that allow to determine persistently exciting input sources. In [7], Syrmos *et al.* formulated the problem via discrete-time state space equations and then estimated the parameters using nonlinear filtering techniques such as the of the extended Kalman filter (EKF).

All aforementioned algorithms are centralized and sensors are assumed to be connected to the computing center directly. Besides, the computing center is assumed to have perfect knowledge of boundary conditions and the manipulability of sources. These assumptions are reasonable in classic system control and identification

---

*Notice that the term *distributed systems* has a different meaning from those normally given in the sensor network literature.

problems, but not applicable in a wireless sensor network environment. In the latter context, the amount of data being exchanged among nodes is a critical parameter in system design due to the limited amount of energy in sensor nodes and the relatively high communicational cost. We will take these constraints into account in this work as described below.

3. DISTRIBUTED IDENTIFICATION OF DIFFUSION COEFFICIENTS

This section describes our approach to the identification of coefficients of diffusion equations. The problem formulation and several key assumptions are given in Section 3.1. The proposed discrete time-space model is presented in Section 3.2. The algorithm to identify system parameters is derived in Section 3.3.

3.1. Problem Formulation

We consider a set of sensor nodes $\{s_i\}$ over a sensor field S. The nodes are measuring discrete-time samples of a scalar field $x(t, \xi)$. For simplicity, we begin by discussing the one-dimensional (1-D) case. We will show that the framework can be extended to the higher dimensional case in a straightforward manner (Section 4.1).

It is assumed that all nodes know their absolute location in the field and the evolution in time and space of $x(.)$ can be modeled by the following diffusion equation:

$$x_t(t, \xi) = \theta_1 x_{\xi\xi}(t, \xi) + \theta_2 x_\xi(t, \xi) + \theta_3 x(t, \xi), \tag{1}$$

where $x_t(.)$ and $x_\xi(.)$ denote the first order derivat ives of $x(.)$ with respect to the time index t and the space index ξ, respectively, and where $t > t_0$ and $0 \leq \xi \leq L$. The initial and Dirichlet boundary conditions are given, respectively, as

$$x(0, \xi) = u_0(\xi),$$

and

$$x(t, 0) = 0 = x(t, L).$$

The function of the sensor network is to identify parameters $\{\theta_j\}$, $j = 1, 2, 3$, based on noisy measurements at nodes of $\{s_i\}$. These nodes send estimates, along with initial and boundary conditions, to the base station to supply the information necessary to reconstruct and predict a diffuse phenomenon without the need to send raw data continuously. The main focus of this work is about the estimation of diffusion coefficients. Note that we did not include a source signal in the model given in (1). We will address the source related issue in Section 6.

3.2. Model Discretization and Clustering

All nodes must be provided with a discretized version of the diffusion model described by Equation (1). We assume that nodes are partitioned in clusters and the identification task is performed at the clusters. The discretization of (1) in the space dimension is obtained by approximating the spatial derivatives with finite differences as

$$x_\xi(t, \xi)\big|_{\xi=ih} \approx \frac{x_{i+1}(t) - x_i(t)}{h}, \tag{2}$$

and

$$x_{\xi\xi}(t, \xi)\big|_{\xi=ih} \approx \frac{x_{i+1}(t) + x_{i-1}(t) - 2x_i(t)}{h^2}, \tag{3}$$

which lead to the lumped model described by the following state-space equations:

$$\dot{\mathbf{x}}(t) = \mathbf{A}(\boldsymbol{\theta})\mathbf{x}(t), \tag{4}$$

$$\mathbf{y}_{c_i}(t) = \mathbf{C}_{c_i}\mathbf{x}(t) + \mathbf{v}(t). \tag{5}$$

where c_i denotes a generic subset of nodes (cluster). The state vector $\mathbf{x}(t)$ represents the uniformly sampled version of the scalar field $x(t, \xi)$, with $x_i(k) := x(kt_s, ih)$, where h is the spatial sampling period. Note that the dimension of the state vector is $L/h + 1$. $\mathbf{A}(\boldsymbol{\theta})$ is a square matrix depending on the spatial discretization and $\boldsymbol{\theta}$ is the vector of parameters to identify.

Equation (5) represents the noisy sensor measurements at node cluster c_i. We assume the noise to be white and Gaussian. The matrix $\mathbf{C}_{c_i}$ is defined on the basis of geometrical relations between the sampling points of the scalar field $\mathbf{x}(t)$ and the sensor locations, $\mathbf{v}(t)$ is a noise term. Thanks to the knowledge of their absolute location in the field, all cluster leaders share the model of Equation (4), but have different measurement equations given by (5). Note that the system model in Equations (4) and (5) is valid even if the coefficients θ_i vary with time and space.

The measurement Equations (5) in the proposed model change from cluster to cluster. The in-network processing as defined by Equations (4) and (5) implies a clusterization of nodes. We may either preselect cluster heads or adopt some mechanism for distributed leader election. Only some of the clusters characterized by high SNR values should be involved in the estimation process for the energy consideration.

Besides clustering, there is a need of some information flooding mechanism so that all nodes involved in the identification process can share the knowledge of the model (4). The nodes participating in the estimation process need to know the location of model boundaries and the spatial sampling rate h, since the information is needed to determine matrix $\mathbf{C}_{C_i}$. Furthermore, all nodes have to share the same state matrix $\mathbf{A}(\boldsymbol{\theta})$. The knowledge of state equations (4) should be centralized in the sensor network in the sense that it must be shared by all nodes. In an application scenario, the selection of initial parameters could be led by the BS station or a node leader in the sensor network to avoid mismatch in model definition. This issue will be examined more thoroughly in Section 6.

Example

In this example, we show an infinite dimensional model and its corresponding lumped model. Consider the PDE model described by the one-dimensional (1-D) diffusion equation

$$x_t(t,\xi) = \theta x_{\xi\xi}(t,\xi),$$

where $t > 0$ and $0 \leq \xi \ L$ with the Dirichlet boundary conditions

$$x(t,0) = 0 = x(t,L).$$

There two sensors in the field, respectively, in locations $\xi_1 = L/2$ and $\xi_2 = 4L/5$, the spacial sampling rate is $h = L/5$, and there is only one cluster formed by the two sensors.

The following 6×6 state matrix $\mathbf{A}(\theta)$ can be derived in a straightforward manner through the finite differences approximation presented in Equation (3). That is,

$$\mathbf{A}(\theta) = \frac{\theta}{h^2}
\begin{bmatrix}
-2 & 1 & 0 & \dots & 0 \\
1 & -2 & 1 & 0 & \dots \\
\vdots & \vdots & \vdots & \ddots & \ddots \\
0 & \dots & 0 & 1 & -2
\end{bmatrix}.$$

The measurement matrix $\mathbf{C}$ can be written as:

$$\mathbf{C} = \begin{bmatrix}
0 & 0 & \frac{1}{2} & \frac{1}{2} & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0
\end{bmatrix}.$$

3.3. Parameter Estimation via Kalman Filtering

The next step is concerned with the identification of unknown parameters $\boldsymbol{\theta}$. For this purpose, we adopt the extended Kalman filter (EKF) [8], [7]. The idea of the EKF is to treat unknown parameters as additional state variables, by defining an *augmented system*, where the state vector is defined as

$$\mathbf{z}(t) := \begin{bmatrix} \mathbf{x}(t) \\ \boldsymbol{\theta} \end{bmatrix}.$$

The augmented system is no longer linear since parameters $\boldsymbol{\theta}$ are multiplied with state variables $\mathbf{x}(t)$. Under this situation, the Kalman filter is used as the state predictor to the linearized version of the augmented system and the estimates of parameters $\boldsymbol{\theta}$ are obtained accordingly.

The augmented system can be written as

$$\dot{\mathbf{z}}(t) = \mathbf{f}(\mathbf{z}(t)) = \begin{bmatrix} \mathbf{A}(\boldsymbol{\theta})\mathbf{x}(t) \\ \mathbf{0} \end{bmatrix}, \tag{6}$$

$$\mathbf{y}_{c_i}(t) = \mathbf{C}_{c_i}\mathbf{x}(t) + \mathbf{v}(t). \tag{7}$$

Then, Equations (6) and (7) are discretized in time using the finite difference scheme given by Equation (2), leading to the discrete-time nonlinear system:

$$\mathbf{z}(k+1) = \begin{bmatrix} (\mathbf{I} + t_s \mathbf{A}(\boldsymbol{\theta}))\, \mathbf{x}(k) \\ \boldsymbol{\theta} \end{bmatrix}, \tag{8}$$

$$\mathbf{y}_{c_i}(k) = \mathbf{C}_{c_i}\mathbf{x}(k) + \mathbf{v}(k). \tag{9}$$

where t_s is the sampling time. Then, the Kalman filter can be applied. Given the mean of the state vector, $\bar{\mathbf{z}}_0$, the initialization for the covariance matrix $\mathbf{P}$ can be found conventionally as

$$\mathbf{P}_0 = E[(\mathbf{z} - \bar{\mathbf{z}}_0)(\mathbf{z} - \bar{\mathbf{z}}_0)^T]. \tag{10}$$

The linearized portions of Equations (8) and (9) are used:

$$\mathbf{J}(k) \quad := \quad \left. \frac{\partial \mathbf{f}(\mathbf{z}(k))}{\partial \mathbf{z}} \right|_{\mathbf{z} = \hat{\mathbf{z}}}, \tag{11}$$

$$\mathbf{H} \quad := \quad [\mathbf{C} \ \ \mathbf{0}]. \tag{12}$$

Therefore, at each measurement time, the filter equations are [8]:

$$\mathbf{L}(k) \quad = \quad \mathbf{P}(k/k-1)\mathbf{H}^T \left(\mathbf{H}\mathbf{P}(k/k-1)\mathbf{H}^T + \mathbf{R} \right)^{-1}, \tag{13}$$

$$\mathbf{P}(k/k) \quad = \quad \mathbf{P}(k/k-1) - \mathbf{L}(k)\mathbf{H}\mathbf{P}(k/k-1), \tag{14}$$

$$\hat{\mathbf{z}}(k/k) \quad = \quad \hat{\mathbf{z}}(k/k-1) + \mathbf{L}(k)\left(\mathbf{y}(k) - \mathbf{C}\hat{\mathbf{x}}(k/k-1)\right) \tag{15}$$

where $\mathbf{R}$ is the covariance matrix of noise measured by sensors the fact that $\mathbf{H}\hat{\mathbf{z}} = \mathbf{C}\hat{\mathbf{x}}$ is applied in the derivation of Equation (15). Then, the state estimate and the covariance are propagated to the next measurement time via

$$\mathbf{P}(k+1/k) \quad = \quad \mathbf{J}(k)\mathbf{P}(k/k)\mathbf{J}^T(k) \tag{16}$$

$$\hat{\mathbf{z}}(k+1/k) \quad = \quad \mathbf{f}(\hat{\mathbf{z}}(k)) \tag{17}$$

Ljung [9] proved that the extended Kalman filter may converge to biased estimates of parameters. This is essentially due to the fact that the state estimates are performed based on the linear projection of a nonlinear system. In our current case, some stability problem of the above algorithm was also observed in computer simulations. Hence, we modified the EKF algorithm by modifying the prediction equation (17). The algorithm becomes more robust by replacing Equation (17) with the prediction performed in the linearized domain, i.e.

$$\hat{\mathbf{z}}(k+1/k) = \mathbf{J}(k)\hat{\mathbf{z}}(k/k). \tag{18}$$

However, the price to pay is a higher bias.

To find a proper balance between the two, the two state prediction equations are used together to result in the latest version of the algorithm. First, the linearized prediction equation (18) is applied until a scalar function depending on the covariance of the parameters is lower than a given threshold. Then, equation (18) is adopted. This byhrid method exhibits stability as well as low bias. The threshold is defined to be the norm of the block $\mathbf{P}_{\boldsymbol{\theta}}$ of the covariance matrix $\mathbf{P}$ that is related to parameters $\boldsymbol{\theta}$:

$$\|\mathbf{P}_{\boldsymbol{\theta}}\|. \tag{19}$$

3.4. Data and Decision Fusion

The distributed algorithm described above implies that the member nodes in a cluster pass their measurements to the cluster head so that the head can perform the EKF. This is a data fusion process, i.e. the head collects raw samples from member nodes and process them together. The transmission of time series from member nodes may be expensive in terms of energy consumption due to a large number of data samples being transmitted, even if the size of a cluster is kept small. An alternative approach which is less expensive in energy consumption is to let each node in a cluster to process its own measurements and then to pass estimated parameter values at a lower rate to the cluster head. This is a decision fusion process. The cluster head only has to average the received estimates as

$$\hat{\boldsymbol{\theta}}_{C_i} = \frac{1}{|C_i|} \sum_{j:s_j \in C_i} \hat{\boldsymbol{\theta}}_{s_j}, \tag{20}$$

where $|C_i|$ is the number of member nodes s_j in cluster C_i. However, this approach may result in poor estimates since only the measurements of one sensor are available to the EKF algorithm (i.e. y_{s_i} is a scalar) at each node. Furthermore, as discussed in Section 6.1, the parameters may not be even identifiable from one measured time series.

To find a good balance between the estimate quality and the communication cost, we may consider a hybrid strategy that integrates data and decision fusion. The first iteration is performed within each small cluster (i.e. $\mathbf{y}_{c_i}(k)$ is a vector) via data fusion modality. Then, the algorithm switches to the decision fusion modality and the cluster head continues to perform the estimation from its own data. This allows some saving in the communication cost, since no time series are exchanged among nodes after the first iteration. The switching condition depends on two factors: (1) the norm of the sub-matrix of the covariance matrix associated with parameters estimates and (2) the number of sample being sent. The algorithm switches to decision fusion if either

$$\|\mathbf{P}_{\boldsymbol{\theta}}\| < t_1, \tag{21}$$

or

$$n_{it} > t_2, \tag{22}$$

where t_1 and t_2 are two thresholds and n_{it} is the number of samples sent to the fusion point since the beginning of the iterative algorithm.

Condition (22) is used to avoid excessive battery consumption for participating nodes. Threshold t_2 can be decreased as the remaining energy level of nodes is decreasing to extend the lifetime of the network with a soft degradation of the performance. At the switching point, the cluster head passes its covariance matrix $\mathbf{P}$ to member nodes so as to provide them with a better initialization state for EKF iterations. During the decision fusion phase, the estimated parameters are passed to the head from a generic member node only in case of relevant variations of its value.

4. EXTENSIONS OF THE PROPOSED APPROACH

This section presents extensions of the proposed algorithm. We consider the two dimensional (2-D) scalar field case in Section 4.1) and the case with spatial and temporal varying parameters in Section 4.2, respectively.

4.1. Extension to 2-D

The extension to the 2-D case has only impact on the definition of the state matrix $\mathbf{A}(\boldsymbol{\theta})$. The proper 2-D finite difference approximations should be adopted instead. Moreover, the samples of the 2-D scalar field should be mapped to the vector of state variables. Such a mapping depends on the geometry of the sensor field. We give a simple example to illustrate the procedure.

Example
Consider the PDE problem described by the diffusion equation:

$$x_t(t, \xi_1, \xi_2) = \theta \nabla x(t, \xi_1, \xi_2).$$

Recall that the Laplacian operator ∇^2 is defined as:

$$\nabla^2 x(t, \xi_1, \xi_2) := x_{\xi_1, \xi_1} + x_{\xi_2, \xi_2}.$$

The 5-point finite difference approximation for $\nabla^2 x$ is:

$$\nabla^2 x(t, \xi_1, \xi_2)\big|_{\xi_1 = ih, \xi_2 = jh} \approx \frac{x_{i+1,j}(t) + x_{i-1,j}(t) + x_{i,j+1}(t) + x_{i,j-1}(t) - 4x_{i,j}(t)}{h^2}. \tag{23}$$

By assuming a square geometry of the sensor field having $n \times n$ points, the state matrix is defined as the $n^2 \times n^2$ block matrix

$$\mathbf{A}(\theta) = \frac{\theta}{h^2} \begin{bmatrix} \mathbf{A}_n & \mathbf{I}_n & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{I}_n & \mathbf{A}_n & \mathbf{I}_n & \mathbf{0} & \dots \\ \vdots & \vdots & \vdots & \ddots & \ddots \\ \mathbf{0} & \dots & \mathbf{0} & \mathbf{I}_n & \mathbf{A}_n \end{bmatrix},$$

where $\mathbf{I}_n$ is a $n \times n$ identity matrix and

$$\mathbf{A}_n = \begin{bmatrix} -4 & 1 & 0 & \dots & 0 \\ 1 & -4 & 1 & 0 & \dots \\ \vdots & \vdots & \vdots & \ddots & \ddots \\ 0 & \dots & 0 & 1 & -4 \end{bmatrix}.$$

4.2. Time and Space Varying Parameters

Similar to the 2-D case, the structure of the algorithm does not change when the parameters are varying with time and/or space. The EKF is an adaptive algorithm so that it can track slow changes in time/spatial varying parameters. A parameter varying with the space ($\theta(\xi)$) will result in a set of parameters (θ_i) in the state space matrix $\mathbf{A}(\boldsymbol{\theta})$. Each parameter θ_i in the lumped system corresponds to the sample of the parameter $\theta(\xi)|_{\xi=ih}$ in the infinite dimensional system.

Example

Consider an infinite dimensional system described by the following diffusion equation. Here, the diffusion coefficient θ is space variable

$$x_t(t, \xi) = \theta(\xi) x_{\xi\xi}(t, \xi).$$

Then, the corresponding state matrix becomes

$$\mathbf{A}(\boldsymbol{\theta}) = \frac{1}{h^2} \begin{bmatrix} -2\theta_1 & \theta_1 & 0 & \dots & 0 \\ \theta_2 & -2\theta_2 & \theta_2 & 0 & \dots \\ \vdots & \vdots & \vdots & \ddots & \ddots \\ 0 & \dots & 0 & \theta_n & -2\theta_n \end{bmatrix}.$$

5. EXPERIMENTAL RESULTS

The experiments presented in this section were performed with respect to the 1-D scalar field. In the first experiment, we considered the diffusion equation

$$x_t(t, \xi) = \frac{1}{\pi^2} x_{\xi\xi}(t, \xi),$$

in interval $0 \leq \xi \leq 1$ and $t > 0$, with initial conditions: $x(0, \xi) = \sin \pi \xi$ and boundary conditions $x(t, 0) = 0$ and $\pi e^{-t} + x_\xi(t, 1) = 0$. The corresponding time varying scalar field is shown in Figure 1(a). The sampling space was set to $1/9$ while the sampling period to 0.01. The covariance of the additive white Gaussian noise at nodes was set to 0.05. Figure 1(b) shows noisy sensor measurements. The coefficient is estimated during the transient state of the system (i.e. no active sources in the field). Therefore, the SNR increases as time passes, because

the values of the scalar field are decaying. We performed the estimate of the diffusion coefficient by means of a cluster with 3 nodes. Figure 2(a) shows the estimate of the diffusion coefficient w.r.t its true value and the covariance of the estimate. Figure 2(b) shows the mean square error of the estimate for the scalar field. Note that the prediction of the scalar field is quite accurate. The bias in the estimate could be due to progressive increase in the SNR value.
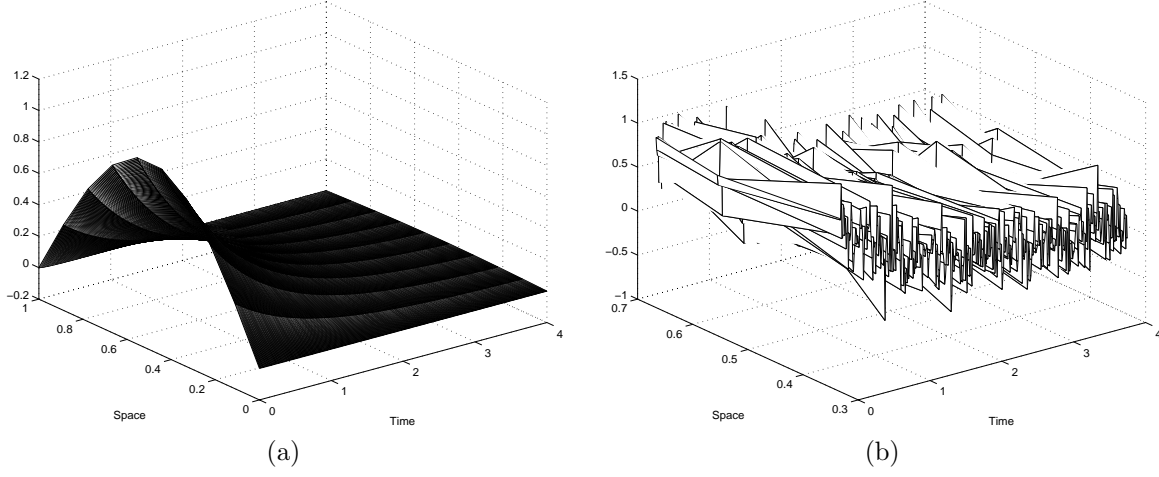

Figure 1. (a) the scalar field being monitored by the sensor network and (b) the noisy sensor measurements.
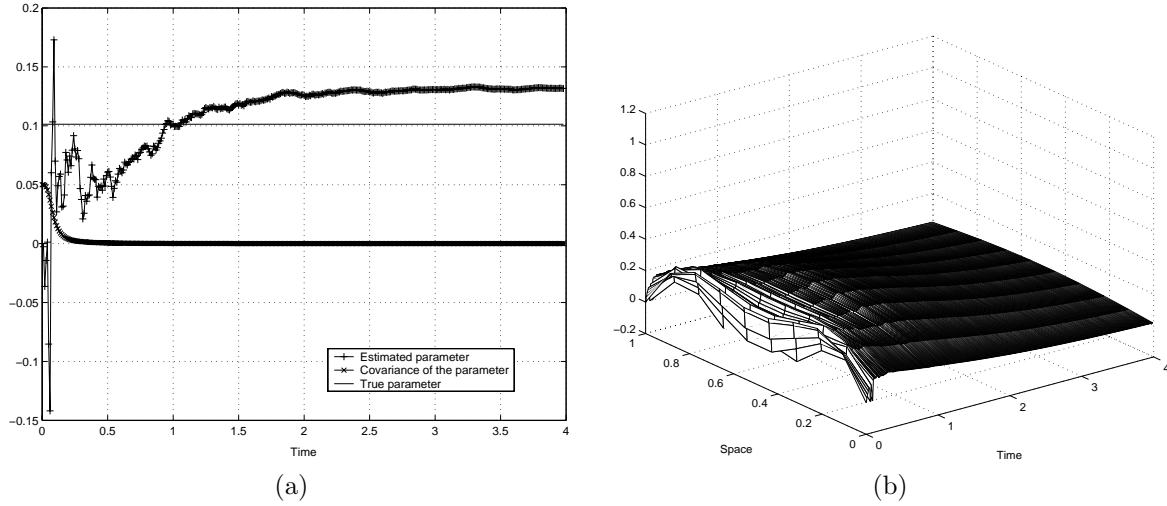

Figure 2. (a) The estimates of the diffusion coefficient and (b) the mean square error of the estimate for the scalar field.

In the second experiment, we considered the diffusion equation

$$x_t(t,\xi) = p_1 x_{\xi\xi}(t,\xi) + p_2 x_\xi(t,\xi) = \frac{1}{\pi} x_{\xi\xi}(t,\xi) + \frac{1}{10} x_\xi(t,\xi),$$

in interval $0 \leq \xi \leq 1$ and $t > 0$, with initial conditions: $x(0,\xi) = 5\sin\pi\xi$ and Dirichlet boundary conditions $x(t,0) = 0 = x(t,1)$. All other parameters (i.e. noise, sampling time and space) are the same as in the previous example. The estimates of the three parameters are shown in Figure 3 (a), (b) and (c). Figure 3(d) shows the mean square error of the estimate for the scalar field. In this case, the hybrid data/decision fusion approach was

adopted and the algorithm was switched to single node decision fusion after the first 10 iterations to save the communication cost.
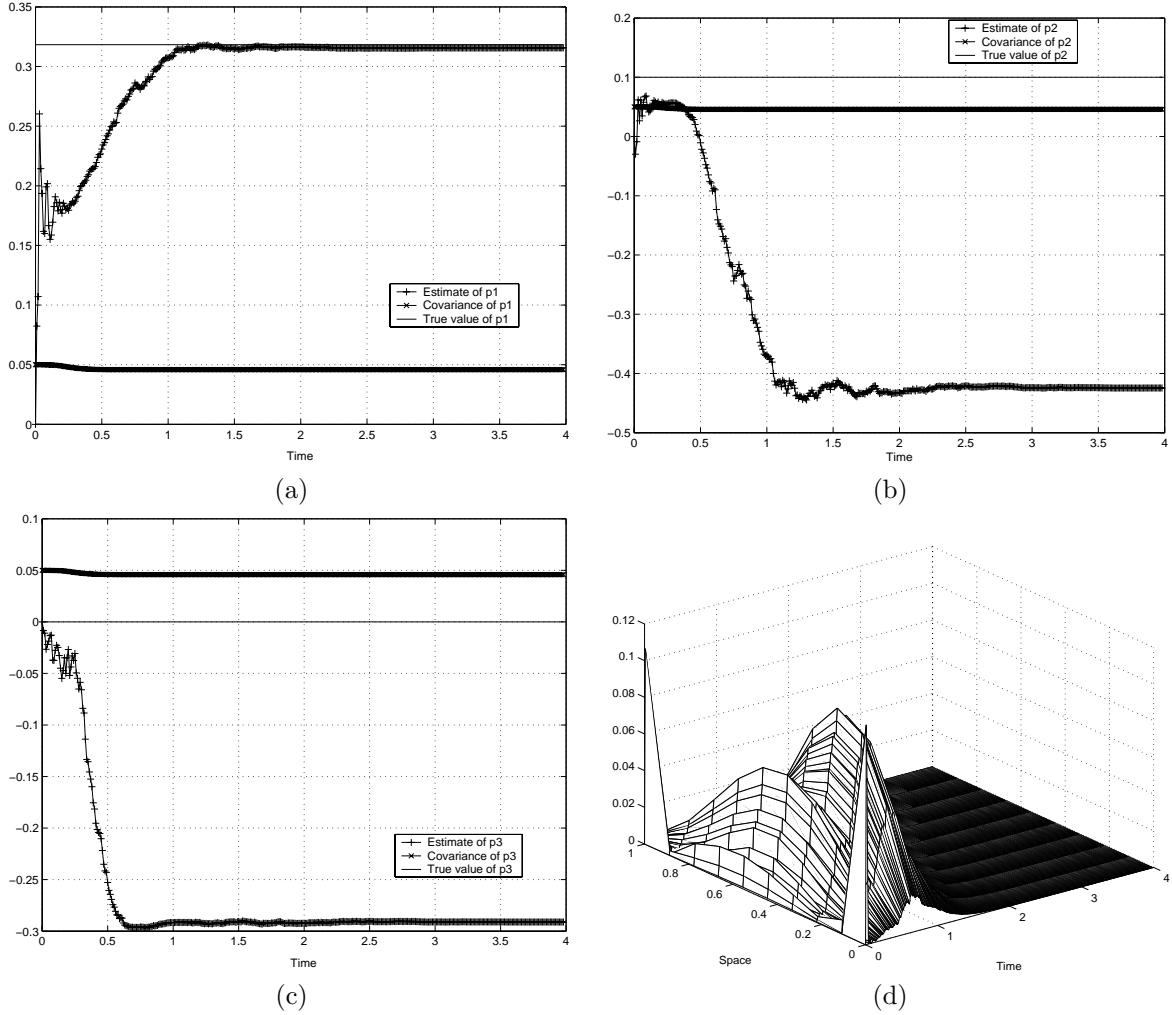

(a)

(b)

(c)

(d)

Figure 3. Estimates of parameters (a) p_1, (b) p_2, (c) p_3, and (d) the MSE for the estimate of the scaler field.

In the last experiment, we considered the diffusion equation

$$x_t(t, \xi) = \frac{1}{5} x_{\xi\xi}(t, \xi) + \frac{1}{100} x_\xi(t, \xi) - \frac{1}{20} x(t, \xi).$$

The interval, initial and boundary conditions were the same as before while the sampling space was changed to 1/5. Figure 4 compares the mean square error of the prediction of the system state, for the cases of a single node cluster, a hybrid decision fusion with the first 10 and 80 iterations using the data fusion modality, respectively. It can be concluded that there is an advantage in performing at least a few preliminary iterations using the data fusion modality before switching to the decision fusion modality.

6. OPEN ISSUES FOR FUTURE RESEARCH

In Section 2, we explained why the problem of distributed identification of PDE-based systems is essentially new in the wireless sensor network area. Such challenges come from bringing a centralized problem to a distributed context but also subject to limitations and uncertainties of wireless sensor network systems.
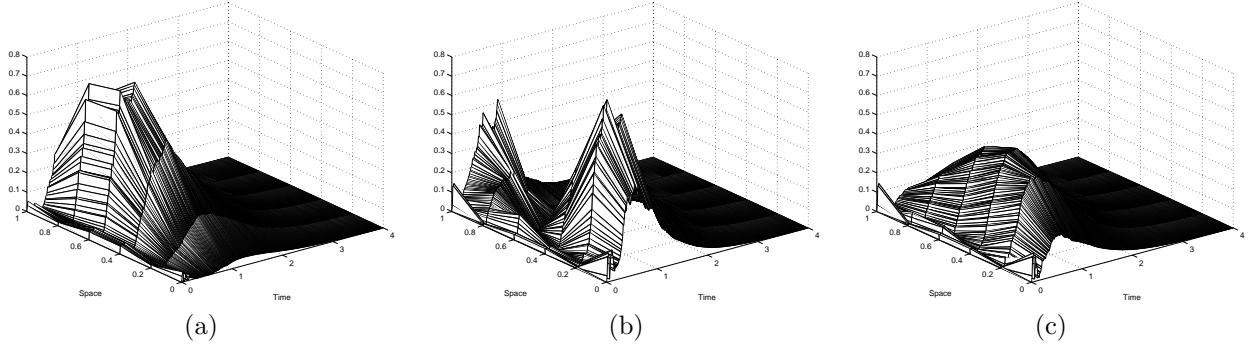
Figure 4. The mean square error of the estimate of parameter p_3: (a) a single cluster measurement, (b) a hybrid scheme with the first 10 iterations using the data fusion modality, and (c) a hybrid scheme with the first 80 iterations using the data fusion modality.

6.1. Indentifiability

It is relevant to study the proper conditions under which a system identification algorithm is able to converge to the correct parameter estimates. This is known as the identifiability problem. In this work, our approach is based on the EKF algorithm. This algorithm performs identification via state estimation so that identifiability is strictly related to the observability of the system state.

Observability is one of the fundamental problems in linear dynamical systems. A system is said to be *observable* at time k_0, if there is a finite $k_1 > k_0$ such that the knowledge of the outputs **y** from k_0 to k_1 is sufficient to determine state $\mathbf{x}(k_0)$. Conversely, an *unobservable* system is one where the values of some of the state variables may not be determined from the system output regardless of the number of samples taken.

By its nature, a Kalman filter applied to an unobservable system will not work. Similarly, the EKF approach will not work if it is applied to a nonlinear unobservable system. The observability test is simple to do for linear systems, but much more difficult for nonlinear ones. For the latter case, only some approximated methods are available [10]. Clearly, performing the EKF on the basis of only one measurement (as in the decision fusion) will present a higher probability of unobservability of the state and therefore of unidentifiability of parameters. It would be interesting to develop some approximate tests at nodes to assess the identifiability of model parameters.

6.2. Sources

The diffusion model in equation (1) does not include a term for the input source. This may correspond to some real scenarios, where the evolution of a phenomenon from some initial conditions is observed in the absence of sources. Sources may have been active in the field before the sensor network starts the monitoring activity (e.g. after the explosion of a gas container). In the absence of sources and with Dirichlet boundary conditions, the identification of coefficients is performed during the transient state of the model. Besides, the SNR at sensors decreases as time passes, because the measurement noise is stationary while measured signals are decaying.

In some other scenarios, one or more sources may be active in the field while the sensor network is sampling the environment. Therefore, the equations for the state space model and the corresponding algorithm should be modified to take input signals into account. The algorithmic extension is straightforward, we only have to modify the prediction equation (17) to be

$$\hat{\mathbf{z}}(k+1/k) = \mathbf{f}(\hat{\mathbf{z}}(k/k), \mathbf{u}(k)). \tag{24}$$

In such a case, source measurements must be provided to sensor nodes. In some scenarios such as the monitoring of industrial plants, source signals can be easily sampled and transmitted to nodes. However, in other situations, the information about the sources may not be readily available, which could bring out several research issues. It is either necessary to find a way to estimate the source signals or to design a blind prediction/identification approach.

A possible strategy to provide cluster heads with source signals would consist first in detecting, counting and localizing the sources. Then, the nodes closest to the sources could play the role of measuring the source signals and transmitting those to cluster heads. Here, we assume that the effects of secondary sources can be ignored at the node in charge of capturing the source signal and those nodes have enough energy to keep transmitting the source samples during the estimation process. Besides, the fusion and/or the source transmission nodes should take into account the propagation effects of sources signals.

The blind prediction/identification strategy seems to be challenging since it may be difficult to make statistical assumptions on sources that are realistic yet easy to treat mathematically.

6.3. Boundary Conditions

In the model presented in Section 3, we assumed homogeneous Dirichlet boundary conditions. This implies that the extension of the sensor field is greater than the one of the scalar field being measured. This may not be always true so that the boundary conditions can be different from zero. If this is the case, the sensor network has to measure boundaries and send samples to the base station since the information about the boundaries is necessary along with the model parameters to reconstruct the evolution of the observed phenomenon. To this end, e.g. in the self localization phase, the sensor network needs to identify the boundary sensors ($\{b_i\}$).

The problem is more complicated when the boundary conditions are varying with time, for example, due to an external source. This requires boundary sensors to keep updating the network and the base station with their measurements. Some simplifications can be achieved by considering the problem in idealized geometrical conditions. Notice that, given a sensor network having m nodes, the number of boundary nodes is $O(\sqrt{m})$. Hence, if the boundary conditions are time varying, $O(\sqrt{m})$ will have to transmit their data to the base station, while m nodes should transfer their data to the BS in a general data gathering scenario.

6.4. Distributed Knowledge of the Model

In Section 3, the state-space model was defined with the same state equations for all nodes. This may imply a centralized definition of the initial model, i.e. the spatial sampling rate, the number of parameters to identify, boundary locations, etc. There are two issues related to such a model definition: (1) how to select the proper model characteristics such as the spatial sampling rate and the boundary locations and (2) how to determine a strategy to flood the model information to all nodes of the network. Actually the two problems can be treated jointly. The model could be defined either at the base station or at the sink and then broadcast to all nodes. A model validation phase might be necessary. Otherwise, this information could be decided a priori, without measurements in the field and just embedded in the node.

As shown earlier, a space varying parameter $\theta(\xi)$ in the PDE model is mapped into a vector of parameters $\boldsymbol{\theta}$ in the lumped model. Thus, one of the consequences of having defined a unique state space model for all nodes is that at each cluster all parameters need to be identified together, including the ones associated to other spacial locations. Besides requiring a higher complexity, this could bring some identifiability problems, if the area covered by the cluster is much smaller than the sensor field. Thus, it may be interesting to consider local state space models at each cluster, rather than a global one for all the cluster, i.e.

$$\dot{\mathbf{x}}(t) = \mathbf{A}_{c_i}(\boldsymbol{\theta}_{c_i})\mathbf{x}(t), \tag{25}$$
$$\mathbf{y}_{c_i}(t) = \mathbf{C}_{c_i}\mathbf{x}(t) + \mathbf{v}(t). \tag{26}$$

This should lead to a reduced number of parameters to estimate. Hence, the dimension of $\boldsymbol{\theta}_{c_i}$ would be smaller than that of $\boldsymbol{\theta}$. In this scenario, each cluster should deal with time varying boundary conditions. This kind of model would be more complicated to handle to reconstruct and predict the observed phenomenon. Furthermore, the base station should reconstruct a unique PDE model from the local lumped models which have been identified by the sensor field.

6.5. Data Propagation

The timely transmission to the base station of all data needed to predict the phenomenon (e.g. parameters, initial and boundary conditions) is another important issue in the proposed framework since we want the BS to predict the phenomenon. For this task, existing data gathering techniques can be used. However, the propagation of estimated parameters through the network could be associated to some estimate updating. Thus, the design of joint flooding/estimation parameter strategies is another relevant issue to consider for this research.

7. CONCLUSION

The problem of monitoring diffuse phenomena via autonomous sensor networks was studied. This subject can have security related applications, for example, when the observed phenomenon is a toxic gas. Parabolic partial differential equations (PDE's) were adopted to model the time and space correlations of the diffusion phenomena observed by the sensor field. In this scenario, the sensor network sends to the base station (BS) only the model parameters, the initial and boundary conditions. This information suffices to predict the evolution of the phenomenon without the need of continuous updates of raw data from the sensor field to the BS. Our main focus here was on the distributed identification of parameters of PDE's. To this end, a method based on extended Kalman filtering (EKF) was presented. The proposed scheme was characterized by a hybrid combination of data and decision fusion modalities, and simulation results were presented. Finally, several open issues in this context, such as identifiability of the parameters, monitoring of time varying boundary conditions and unknown sources, were discussed.

REFERENCES

1. K. Kalpakis, K. Dasgupta, and P. Namjoshi, "Maximum lifetime data gathering and aggreagation in wireless sensor networks," in *IEEE International conference on networking*, pp. 685–696, Aug. 2002.
2. W. R. Heinzelman, A. Chandrakasan, and H. Balakrishnan, "Energy efficient communication protocol for wireless sensor networks," in *Proc. International Conference on System Sciences*, **33**, January 2000.
3. J. Baumeister, W. Scondo, M. Demetrioiu, and I. Rosen, "On-line parameter estimation for infinite dimensional dynamical systems," *SIAM J. Control Optimization* **35**, pp. 678–713, March 1997.
4. Y. Orlov and J. Bentsman, "Adaptive distributed parameter systems identification with enforceable identificability conditions and reduced-order spatial differentiation," *IEEE Trans. on Automatic Control* **45**, pp. 36–50, February 2000.
5. R. Curtain and H. Zwart, *An Intoduction to Infinite-Dimensional Linear Systems Theory*, Springer Verlag, 1995.
6. M. Demetriou and I. Rosen, "Robust adaptive estimation schemes for parabolic distributed parameter systems," in *36th IEEE Conference on Decision and Control*, **4**, pp. 3448–3453, December 1997.
7. J. Yin, V. Syrmos, and D. Yun, "System identification using nonlinear filtering methods with applications to medical imaging," in *39th IEEE Conference on Decision and Control*, **4**, pp. 3313–3318, December 2000.
8. J. M. Mendel, *Lessons in Estimation Theory for Signal Processing, Communications and Control*, Prentice Hall, 1995.
9. L. Ljung, "Asymptotic behavior of the extended kalman filter as a parameter estimator for linear systems," *IEEE Trans. on Automatic Control* **AC-24**, pp. 36–50, February 1979.
10. C.-T. Chen, *Linear Systems Theory and Design*, Oxford University Press, 3rd ed., 1999.