Identification in Sensor Networks

CONSTANTIN VOLOSENCU
Department of Automatics and Applied Informatics
“Politehnica” University of Timisoara
Bd. V. Parvan Nr. 2, 300223, Timisoara
ROMANIA

Abstract: - In the last years sensor networks have proved their huge viability in the real world, even if their resources in terms of energy, memory, computational power and bandwidth are strictly limited. One of the important problems related to the usage of wireless sensor networks in harsh environments is the identification of the states of the physical variables in the field, based on the measurements provided by the sensors. The sensor networks allow the usage of the multivariable estimation techniques in distributed systems. The paper presents a short survey of some characteristics of the sensor networks, distributed parameters systems and identification techniques. An examples of application of modeling of distributed systems in sensor networks and identification based on multivariable identification with auto-regression and neural network is presented.

Keywords: - wireless sensor network, system identification, distributed parameter systems, neural networks.

1 Introduction
The development of wireless sensor allows development of new methods and algorithms for identification of systems, especially in the case of distributed parameter systems. As a principle in this kind of identification the sensor network may be seen as a “distributed sensor” placed into the system with distributed parameters, allowing measurement in well-chosen points of the system. In the last years a lot of papers were published in the fields of identification of distributed parameter systems and sensor networks [1+5, 7, 9, 10, 12+14, 16, 18+20]. This paper presents a short synthesis of the main aspects of the concepts involved in identification of distributed parameter systems based on sensor networks.

Advances in scientific computation and developments in spatial sensor technology have enhanced the ability to develop modeling strategies and experimental techniques for the study of the spatiotemporal response of distributed nonlinear systems [3, 4, 6, 7, 9, 11, 14, 15, 17]. The simplifying methods for modeling of these systems, that are trying to capture the distributed system dynamics through lumped parameter models, can be developed. Robust implementations of distributed system identification algorithms, based on detailed spatiotemporal experimental data have now an important role to play.

Advances in hardware and wireless network technologies have created low-cost, low-power, multifunctional miniature sensor devices. These devices make up hundreds or thousands of ad hoc tiny sensor nodes spread across a geographical area. These sensor nodes collaborate among themselves to establish a sensing network. A sensor network can provide access to information anytime, anywhere by collecting, processing, analyzing and disseminating data. Thus, the network actively participates in creating a smart environment [1, 5, 12, 13, 16, 18+20].

Since for distributed parameter systems it is impossible to observe their states over the entire spatial domain, a possible solution is to locate discrete sensors to estimate the unknown system parameters as accurately as possible. There is recent original work on optimal sensor placement strategies for parameter identification in dynamic distributed systems modeled by partial differential equations [17, 19, 20]. New development of new techniques and algorithms or adopting methods, which have been successful in the field of optimal control and optimum experimental design, are reported in papers.
2. Related Work

A strategy by which sensor nodes detect and estimate non-localized phenomena such as boundaries and edges (e.g., temperature gradients, variations in illumination or contamination levels) is studied in [12]. A general class of boundaries, with mild regularity assumptions, is considered, and theoretical bounds on the achievable performance of sensor network based boundary estimation are established. A hierarchical boundary estimation algorithm is proposed that achieves a near-optimal balance between mean-squared error and energy consumption.

In many sensor networks applications, sensors collect correlated measurements of a physical field, e.g., temperature field in a building or in a data center. However, the locations of the sensors are usually inconsistent with the application requirements. The papers [19, 20] consider the problem of estimating the field at arbitrary positions of interest, where there are possibly no sensors, from the irregularly placed sensors. The sensor network on a graph is mapped, and by introducing the concepts of interconnection matrices, system digraphs, and cut point sets, real-time field estimation algorithms are derived. Simulations and real-world experiments on temperature estimation are done.

Theory of partial differential equations is presented in [6] and applications to some systems with distributed parameters in [15].

Developing low-order models of high fidelity is important if the objective is accurate control of the distributed parameter systems. The work [21] presents a method to develop a low-order model when there is no available exact model of the system. The foundations for this method are singular value decomposition theory and the Karhunen-Loève expansion. It is shown that satisfactory closed-loop performance of the nonlinear distributed parameter systems can be obtained using a dynamic matrix controller designed using the finite order model.

In the paper [7] a methodology for the identification of distributed parameter systems, based on artificial neural network architectures, motivated by standard numerical discretization techniques used for the solution of partial differential equations is presented.

A new direct approach to identifying the parameters of distributed parameter systems from noise-corrupted data is introduced in [3]. The model of the system, which takes the form of a set of linear or nonlinear partial differential equations, is assumed known with the exception of a set of constant parameters. Using finite-difference approximations of the spatial derivatives the original equation is transformed into a set of ordinary differential equations. The identification approach involves smoothing the measured data and estimating the temporal derivatives using a fixed interval smoother. A least-squares method is then employed to estimate the unknown parameters. Three examples that illustrate the applicability of the proposed approach are presented and discussed.

3. Sensor Networks

Wireless sensor networks are extremely distributed systems having a large number of independent and interconnected sensor nodes, with limited computational and communicative potential. The sensors are deployed for data acquisition purposes on a wide range of locations, sometimes in resource-limited and hostile environments such as disaster areas, seismic zones, ecological contamination sites, and military combat zones. In this structure data processing is at the sensor level, data transmission is wireless, sensing mechanism is not necessarily and power supply is not necessarily wireless.

The sensors are smart, small, lightweight and portable devices, with a communication infrastructure intended to monitor and record specific parameters like temperature, humidity, pressure, wind direction and speed, illumination intensity, vibration intensity, sound intensity, power-line voltage, chemical concentrations and pollutant levels at diverse locations. The sensor number in a network is over hundreds or thousands of ad hoc tiny sensor nodes spread across different areas. Thus, the network actively participates in creating a smart environment. They are low cost and low energy devices, realized in nanotechnology. With them low cost wireless platforms, including integrated radio and microprocessors may be developed. The sensors are adequate for autonomous operation in highly dynamic environments as distributed parameter systems. When they fail new sensor may be added. They require distributed computation and communication protocols. They assure scalability, where the quality can be traded for system lifetime. They assure Internet connections via satellite.
Sensor network applications include: environmental monitoring, civil infrastructure monitoring, shared resource utilization, tracking, and military surveillance. Application are in microclimates, air quality, soil moisture, animal tracking, energy usage, office comfort, wireless thermostats, wireless light switches. In techniques they have as applications data acquisition of physical and chemical properties, at various spatial and temporal scales, as in distributed parameter systems, for automatic identification, measurements over long period of time.

The sensors have the following technical characteristics: a robust radio technology, cheap and energy efficient processors, lifetime energy source, on-board memory, flexible I/O for various sensors, common highly available components, efficient resource utilization – currently uses 10 µA average, high modularity, flexible open source platform. Some examples of their technical data are: 128 KB instruction EEPROM, 4 KB data EEPROM, 512 KB External Flash Memory, radio with 38 K or 19 K baud, at 900MHz, LEDs, µP at 7,3 MHz, JTAG, programming board, ISM Bands: 433-434,8 MHz in Europe, power consumption: 16 mA Tx, 9 mA Rx, 2 µA sleep, transmission range: 1m, off the floor 100m range, ground level 10 m range, interface block data to laptop, GPS, cost: $ 95.

Standards and protocols are imposed for sensor networks development.

The sensor networks have different structures. The star networks (point-to-point), are networks in which all sensors are transmitting directly with a central data collection point. The mesh networks are networks in which sensors can communicate with each other. In mesh networks sensor nodes can relay messages from other sensor nodes, there is no need for repeaters. Software controls the flow of messages through network with self-configuration. New nodes automatically detected and incorporate. Advantages of the mesh structures are: robustness, easily deployed, no RF site surveys needed, no repeaters needed, easily expanded. Their disadvantages are: more complicated software, energy consumption of nodes increases, each node must transmit other nodes messages as well as its own, potentially less bandwidth. A sensor has the following hardware: radio node, antenna, on-board board microprocessor contains code for managing mesh network. As hardware development board it contains pins for sensor connection, microprocessor for handling signal, power supply, serial port, radio node plugs onto top of board. The sensor contains software on board for data acquisition, signal processing, embedded programming, embedded C language, messages format up to user.

Different structure may be used in practice, for example. The sensor network may be static or mobile. For a static case each sensor node knows its own location, even if they were deployed via aerial scattering or by physical installation. If not, the nodes can obtain their own location through the location process. Moreover, all the sensors passed a one-time authentication procedure done just after their deployment in the field. The sensor nodes are similar in their computational and communication capabilities and power resources to the current generation sensor nodes. Every node has space for storing up to hundreds of bytes of keying materials in order to secure the transfer of information through symmetric cryptography.

In the network there is a base station, sometimes called access point, acting as a controller and also as a key server. It is assumed to be a laptop class device and it is supplied with long-lasting power.

An example of wireless cellular network (WCN) architecture is presented in Fig. 1.

![Cellular Network Architecture](image)

In this architecture, a number of base stations are already deployed within the field. Each base station forms a cell around itself that covers part of the area. Mobile wireless nodes and other appliances can communicate wirelessly, as long as they are within the area covered by one cell.

A versatile architecture is SENMA - SEnsor Network with Mobile Access architecture that is
presented in Fig. 2, used for large-scale sensor networks.

Fig. 2. SENMA Architecture

The main difference related to the cellular network architecture is that base stations are considered to be mobile, so each cell has varying boundaries which implies that mobile wireless nodes and other appliances can communicate wirelessly, as long as they are at least within the area covered by the range of the mobile access point.

Multiple sensor nodes can detect an event situated in the surrounding area, so redundancy of sensor networks is assured.

4 Distributed Parameter Systems

The distributed parameter systems, opposed to the lumped parameter systems, are systems whose state space is infinite dimensional. An object whose state is heterogeneous has distributed parameters. Such a system is described by partial differential equations. Partial differential equations are used to formulate problems involving functions of several variables, such as the propagation of sound or heat, electrostatics, electrodynamics, fluid flow, elasticity. Distinct physical phenomena have identical mathematical formulations, and the same underlying dynamic governs them. Some examples of distributed parameter systems are presented as follow [15].

One of the most important domain of applications of the partial differential equations is the process of heat conduction, with propagation of heat in anisotropic medium: propagation of heat in a porous medium, transference of heat in semi-space compound by two materials submitted to heating, processes of transference of heat between a solid wall and a flow of hot gas, estimation of the temperature field in space with fissured zone having the form of a circular disc. Applications related to electricity domain are: the propagation of electric current in cables, the heating of the electrical contacts. In the field of motion of fluid there are: plane motion of viscous fluids, running of viscous fluids in rectilinear tube, computation of losses of non-stationary heat in subterranean pipe, running of gases in water main. The processes of cooling and drying: cooling of clap, cooling of a sphere, drying of wood pieces, drying in vacuum. Phenomenon of diffusion: diffusion flow in a heavy sphere for chemical reactions happening with finite element on the sphere surface, the flames diffusion, which appears to the beginning of a tube, repartition density of particles loading by the meteorites. Other applications are: estimation of the ice height covering the snow the arctic seas, motion of underground waters, alloy of heavy fusible particles, investigation of the wave close to the single point of the board of a plane plate, the growing of the gas particles in a fluid, substances combustion, the temperature modification in the air mass.

Process of heat conduction. Let it be an object of a volume $V \subset R^3$. The frontier of dominium $V$ is a surface $S$, formed by a finite number of smooth surfaces. Let it be $\theta(P, t)$ the function of the object’s temperature, at the time moment $t$, where $P \in V$ is a point in the volume $V$. If different points of object have different temperatures, $\theta(P, t) \neq c t.$, then a heat transfer will take place, from the warmer parts to the less warm parts. Let it be a regular surface $\sigma$ placed in $V$, which contains the point $P$. From the theory of thermal conductivity through the $d\sigma$ in the time $dt$ a heat quantity $dQ$ is passing, proportional to the product $d\sigma dt$ and proportion to the function $\theta(P, t)$ derivative, along the normal $n$ to the surface $\sigma$ in the point $P$:

$$dQ = k \frac{d\sigma}{dt} \theta(P, t)$$

(1)

where $k$ is a proportionality factor, called coefficient of internal thermal conductivity of the object. The vector $\text{grad} \ \theta$ has its direction along the normal at the level surface for $\theta=ct.$, in the sense of $\theta$ rising.

The law of heat propagation through an object in which there are no heat sources:

$$\frac{\partial \theta}{\partial t} = \text{div}(\mathbf{K} \ \text{grad} \ \theta)$$

(2)

The heat sources in the object have a distribution given by the function:

$$\mathbf{q}(x, t)$$

(3)
If the object is homogenous \( a = \sqrt{k/\gamma / \rho} = ct \). and the equation (2) is written:

\[
\gamma \frac{\partial \theta}{\partial t} = \frac{\partial}{\partial x} \left( k \frac{\partial \theta}{\partial x} \right) + F(t, x) \tag{4}
\]

In the case of heat propagation through a bar:

\[
\gamma \frac{\partial \theta}{\partial t} = \frac{\partial}{\partial x} \left( k \frac{\partial \theta}{\partial x} \right) + F(t, x) \tag{5}
\]

The initial conditions or of the limit conditions have physical significance. The equation does not determine completely the state of the object \( K \). We must take in considerations the initial state of the object, the temperature distribution in the object at the moment \( t=0 \):

\[
\theta(x, 0) = \theta_0(x) \tag{6}
\]

called initial condition.

### 5 Identification Techniques

System identification is for building accurate, simplified models of complex systems from noisy time-series data. It provides tools for creating mathematical models of dynamic systems based on observed input/output data. The identification techniques are useful for applications ranging from control system design and signal processing to time-series analysis. Actually, there is a huge amount written on the subject of system identification. The textbook [11] deals with identification methods and also describes methods for physical modeling. For more details about the algorithms and theories of identification there is [17]. But, only the experience with real data may help us to understand more. It is important to remember that any estimated model, no matter how good it looks on design, has only picked up a simple reflection of reality. So, in this aspect the sensor network is the powerful tool.

There are identification methods based on parametric model and on nonparametric models. From the second category we enumerate spectral methods, correlations methods, recursive state estimation with Bayes filters and Gaussian filters.

Models describe relationships between measured signals (Fig. 3).

Fig. 3. Identification signals

The outputs \( y(t) \) are then partly determined by the inputs \( u(t) \). In most cases, the outputs are also affected by more signals than the measured inputs. Such unmeasured inputs are called disturbance signals or noise \( e(t) \). The relationship is \( y(t) = f(u(t), e(t)) \).

In case of systems that cannot be modeled based on physical insights it is possible to use standard models, which by experience are known to be able to handle a wide range of different system dynamics. A family of such ready-made model, which tell the size and it is possible to find them to fit to measured data.

A general case model is the Box-Jenkins model, summarized as:

\[
\theta \quad \text{is the parameter vector thus contain the coefficients } b_i, c_i, d_i, f_i \text{ of the transfer functions and } e \text{ is the white noise.}
\]

The model is described by the structural parameters \( n_{b, c, d, f} \) and \( n_k \). When these have been chosen it remains to adjust the parameters \( b_i, c_i, d_i, f_i \) to data.

Some special cases are: the output error model (OE), where the properties of the disturbance signals are not modeled and \( H=1 \); the ARMAX model, the auto-regression and a moving average of white noise, when

\[
\theta = \quad \text{and the ARX model, where in the above case of ARMAX other simplification is done } C(q)=1.
\]

Starting from (9) it is possible to predict what the output \( y(t) \) will be, based on measurements of \( u(t) \), \( y(t) \). The prediction of \( y(t) \) is obtained with the general expression:
which have special cases for OE and ARX.

The *principle of minimizing the prediction errors* is used to fit the parameterized models to data.

The output $y$ at time $t$ is thus computed as a linear combination of past outputs and past inputs. It follows, for example, that the output at time $t$ depends on the input signal at many previous time instants. This is the system dynamics. The identification problem is then to use measurements of $u$ and $y$ to figure out: the coefficients in this equation; how many delayed outputs $n_f$ to use in the description. If the time delay in the system is $n_f$ that, from the equation we may see that it takes $n_h$ sample periods before a change in $u$ will affect $y$. How many delayed inputs $n_b$ to use.

The same identification algorithms may be implemented using *neuronal networks*. In this case a feedforward neural network, with continuous values, with two hidden layers, working as a time series estimator is recommended [8].

6 Malicious Node Detection
In this example a strategy based on antecedent values provided by each sensor for detecting their malicious activity is presented. At each time moment the sensor’s output is compared with its estimated value computed by a robust *autoregressive neural predictor*. In case that the difference between the two values is higher then a chosen threshold, the sensor node becomes suspicious and a decision block is activated.

The strategy of detection considers an *autoregressive (AR) model* that approximates the time evolution of the measured values provided by each sensor:

$$\text{(12)}$$

where $x(t)$ is the series under investigation (in our case it is the series of values measured by the same sensor), $a_i$ are the auto regression coefficients, $n$ is the order of the auto regression and $e$ is the noise. The model (12) may be implemented using a *feedforward neural network* with continuous values. The inputs of the neural network are the measured values of the sensor, at previous $n$ time moments. The coefficients $a_i$ will be given by the weights and biases of the neurons from the hidden layers of the neural network. The output of the neural network will be the estimate at the time $t$ of the sensor value:

$$\text{(13)}$$

The weights and biases values are computed by training, knowing on-line or off-line a set of training under the form of a time series $x(t), x(t-1), \ldots, x(t-n)$.

The strategy uses the time series of measured data provided by each sensor and relies on an autoregressive neural predictor placed in base stations (Fig. 4).

Fig. 4. Malicious node detection scheme
The *detection principle* is: comparing the value of a malicious sensor node, that will try to enter false information into the sensor network, that will be identified by its output value, $x(t)$ with the value $\hat{x}(t)$ predicted.

The *proposed methodology* is described as follows. At every instant $t$ the estimated value $\hat{x}_A(t)$ is computed relying only on past values $x_A(t-1), \ldots, x_A(0)$ and parameter estimation and prediction is used, as in the following steps. First the parameters $w_i, b_i$ of the neural network are determinate, using the Levenberg Marquardt method. A *training set* including all the possibilities of the sensor network behavior is used. The neural network is trained to obtain a small training error and a high degree of generalization. The neural network is test with a *test set*. Second, the *prediction value* $\hat{x}(t)$ is obtained using the following equation:

$$\text{(14)}$$

After that, the present value $x_A(t)$ measured by the sensor node is compared with its estimated value $\hat{x}_A(t)$ by computing the error:

$$\text{(15)}$$

If this *error* is higher than a *threshold* $\varepsilon_A$ then the sensor $A$ will be considered to be a potentially corrupted sensor and the decision block will be activated (Fig. 4). Here, based on a database containing the known attacks models, a knowledge-
based system can take the decision to expel the malicious node from the network topology.

There is no simple to establish the correct model order in case of an AR model. In this case there are two parameters that influence the decision: the type of data measured by sensors and the computing limitations of the base stations. Because both of them are a priori known an off-line methodology is recommended. Realistic values are between 3 and 6.

The structure of the neural network is established after iterative trainings.

The distributed process. The propagation of a temperature wave, in a homogenous planar field, is considered, where several sensor nodes $S_{ij}$ with $i=1,\ldots,N$ and $j=1,\ldots,M$, being a part of a sensor network, have been deployed. These sensors are measuring the local temperature $\theta \, \text{[°C]}$. A possible malicious node, to be detected, is denoted by $S_A$. An auto-regression neural network model is developed to estimate the temperature value provided by the sensor $A$: $\hat{\theta}_A(t)$, by taking into consideration the previous values of the data provided by sensor $x_A(t-1), x_A(t-2), \ldots, x_A(t-n)$.

The time distribution of the temperature $\theta$ through the homogenous medium in space is $\theta(z,t)$, at the moment $t$, at distance $z$ from the heat source.

The heat conduction, when neglecting the heat loses in the environment, is described by the heat equation:

$$ c_0 \frac{\partial^2}{\partial z^2} \theta(z,t) = \frac{\partial}{\partial t} \theta(z,t) \tag{16} $$

where $c_0$ is the heat conductivity coefficient of the medium.

In order to investigate how the strategy works the function $\theta = \theta(z,t)$ is discretised into the aggregates $\theta_{j,k}$ (temperature value provided by $S_{j,k}$) measured at the distance $z_{j,k}$ from the origin. The goal is to obtain the temperature $\theta_A$ measured by the corresponding sensor ($S_A$).

The energy conservation is governed for each point in the field by the following equation:

$$ \frac{d}{dt} W_{j,k} = P_{in}^{j,k} - P_{out}^{j,k} \tag{17} $$

where $W_{j,k}$ is the energy stored in point $(j,k)$, $P_{in}^{j,k}$ is the input power in the point and $P_{out}^{j,k}$ is the output power from the point. The space model of the sensor deployed in the field with the heat sources is presented in Fig. 5.

$$ \sum_{i=1}^{r} P_{i} \left[ \theta_{i}^{j,k}(t) - \theta_{i}^{j,k}(t) \right] $$

A discrete time equivalent equation of (8), with a chosen adequate sample period $h$ is used. Each cell of sensors is receiving inputs from the around medium, from $r$ sources with powers $P_i$, $i=1, \ldots, r$, positioned around the network. The heat sources $P_i$ are positioned in different points in the coordinate system $xOy$. Some coordinate transformations may be done and the sources may be moved in the adjacent points of the network.

The neural network used for estimation is a feedforward neural network, with continuous values. It has 4 inputs, the sensor values at 4 antecedent time moments: $x_A(t-1), x_A(t-2), x_A(t-3)$ and $x_A(t-4)$. The output layer has one neuron for the estimated temperature. The structure of it is presented in Fig. 6.

According to Kolmogorov’s theorem two hidden layers are used, with biases, to obtain a reduced error of approximation of the estimate. The first and the second hidden layers have a reduced number of neurons, 32 and 16 neurons, respectively. These numbers resulted after some iterative training. The activation functions of the neural network are the hyperbolic tangent function for the hidden layers and the first-order linear function for the output layer.
The training of the neural network with a training set, which cover the entire possible scenario in the field. The input data for estimation is a time series of the temperature sensor \( S_A \) as the state of the heat diffusion model (18). This time series is obtained using sums of the traveling temperature waves, generated by the heat sources \( P_i \). The temperatures propagate through to the sensor \( S_A \).

The training set was obtained using present and anterior values of the sensors, \( \theta_A(t-1), \theta_A(t-2), \theta_A(t-3), \theta_A(t-4); \theta_A(t) \) taken from the transient responses of the model (18). The training was made with the Levenberg-Marquardt method.

The sum square error after 9 training epochs is presented in Fig. 7.

At a specific time moment (\( t=400 \)) the sensor was corrupted. Some different sets of candidate models for the model structure could be experimented. A 4th order estimation general model was chosen. With weights and biases it has the expression:

\[
\hat{x}_A(t) = \sum_{i=1}^{4} a_i(w,b) \cdot x_A(t-i) + \xi
\]  
(19)

This autoregressive neural estimation is applied for the sensor \( S_A \).

The estimated temperature \( \hat{x}_A(t) = \hat{\theta}_A(t) \) for the sensor \( S_A \) is presented in Fig. 8, over the original time series \( x_A(t) \).

Fig. 7. The training error

Fig. 8. The sensor value and the estimate

7 Conclusion

In this paper a short survey of the following topics is presented: -sensors, as small and smart measuring devices, (some technical characteristics of the sensors are presented), which in a high number they are forming sensor networks, -sensor network structures and characteristics, which acts as a distributed sensor in distributed parameter systems; -identification techniques based on parameter models, which are possible to be implemented using neural networks.

A case study of malicious nod detection using a
neural auto-regression method in the process of plane heat propagation is presented.

References: