Distributed Configuration of Sensor Network for Fault Detection in Spatio-Temporal Systems

Maciej Patan and Damian Kowalów
Institute of Control and Computation Engineering,
University of Zielona Góra,
ul. Podgórna 50, 65–246 Zielona Góra, Poland
E-mail: {m.patan,d.kowalow}@issi.uz.zgora.pl

Abstract. The problem of fault detection in spatio-temporal systems is formulated as that of maximizing the power of a parametric hypothesis test verifying the nominal state of the process under consideration. Then, adopting a pairwise communication schemes, a computational procedure is developed for the spatial configuration of the observation locations for sensor network which monitor changes in the underlying parameters of a distributed parameter system. As a result, the problem of planning the percentage of experimental effort spent at given sensor locations can be solved in a fully decentralized fashion. The approach is verified on a numerical example involving sensor selection for a convective diffusion process.

1. Introduction
Recently, one can observe an extremely fast development of methods of Fault Detection and Isolation (FDI) for dynamical systems. Although a vast amount of existing approaches (for surveys reader can be referred to [3, 7, 9, 24, 39]), most of contributions focus on the methods for lumped parameter systems with just a few approaches dedicated to spatio-temporal systems also known as distributed-parameter systems (DPSs). Since in DPS it is usually impossible to observe the system states over the entire spatial domain, the problem of sensor location and observation scheduling for proper system monitoring become of vital importance. But within the framework of FDI systems for DPSs, the optimization of measurement process which increases the reliability of the diagnosis is usually neglected and there is a substantial lack of techniques dedicated to this area.

Technological advances in communication systems and the development of inexpensive mobile measurement systems made it possible to deploy a group of networked devices in a number of environments, cf. [13, 4]. A cooperated and scalable network of sensors has the potential to substantially enhance performance of the observation systems. Although the number of sensor placement techniques to manage the configuration for sensor networks of practical scale is very limited, some effective approaches has been proposed to cover a number of different experimental settings, including stationary [12, 34, 22, 26, 17, 10, 21, 11], scanning [22, 33, 15] or moving observations [29, 27, 34, 19, 36]. Their adaptation to the sensor-network setting was reported in [32, 31].

However, most techniques communicated by various authors rely on centralized techniques, which assume an existence of some superior entity to maintain the whole network and responsible
for optimization of the observation strategy. The distributed nature of the design problem is taken into account very occasionally. Recent advancements in sensor networks necessitate effective, distributed and fault-tolerant algorithms for computation and information exchange.

In the background of diagnostic problems, another difficulty is a development of proper relation between the performance of diagnostic system and the observational strategy. For this purpose, a classical hypothesis testing is used for model-based diagnosis in DPSs. This allows us to construct a qualitative criterion of sensor allocation using the notion of D-optimum designs originated in optimum experiment design theory. Then, proper fault detection and localization schemes are developed based on the results reported by Patan [22] for stationary sensors and further extended for scanning and mobile sensor networks [16, 22, 37, 23, 14, 30].

In this paper we wish to focus on certain computational aspects of sensor placement problems for fault detection in DPSs. Given a finite set of stationary located sensors, the observation horizon is divided into consecutive stages. Then, the non-negative weights are assigned to each sensor at given stage so as to maximize the determinant of the Fisher Information Matrix (FIM) related to the parameters to be identified. The weight assigned to a measurement point can be interpreted as the proportion of observations performed at this point, or the percentage of experimental effort spent at this point. The potential solutions are of considerable interest while assessing which sensors are more informative than the others and allow for complexity reduction of the measurement system. The solution proposed here is close in spirit to the classical optimum experimental design theory for lumped systems [1, 38]. Its novelty relies on a generalization of a known numerical algorithm for computation of a D-optimum design on a finite set together with its efficient distributed implementation dedicated to sensor network.

As a result, an extremely simple computational scheme is developed to determine best sensor configurations. Its idea is based on the class of so-called gossip algorithms in which each node communicates with no more than one neighbor at each time instant [40, 5, 6]. The different nodes of the network independently calculate and store the desired quantities and their final estimates are obtained in a fully decentralized fashion. The advantage of such information exchange is clear as it is robust with respect to individual sensor faults and the global estimate of information matrix is stored at all sensor nodes so easily can be recovered.

2. Model-based fault detection in DPS

Parameter estimation is one of the fundamental approaches in the framework of analytical techniques of fault detection [9]. It is especially useful in situations when the faulty system state does not only changes system response but leads also to perturbation in values of model parameters. In practice, this is quite common if only parameters have a physical interpretation

Let \( y = y(x; t; \theta) \) denote the scalar state of a given DPS at a spatial point \( x \in \Omega \subset \mathbb{R}^d \) and time instant \( t \in T = [0; t_f], t_f < \infty \). Here \( \theta \) represents an unknown constant \( m \)-dimensional parameter vector which must be estimated using observations of the system.

The state \( y \) is observed by \( N \) scanning sensors, which possibly change their positions at time instants \( 0 < t_0 < t_1 < \cdots < t_K = t_f \) and will be remaining stationary for the duration of each subinterval \( T_k = [t_{k-1}, t_k], k = 1, \ldots, K \). The observations can be formally represented as [20, 15]

\[
  z_j(t) = y(x^j_k, t; \theta) + \varepsilon(x^j_k, t), \quad x^j_k \in X, \quad t \in T_k,
\]

for \( j = 1, \ldots, N \) and \( k = 1, \ldots \), where \( x^j_k \) is element selected from defined \( a-priori \) set of sensor locations \( X = \{x^1, \ldots, x^N\} \) and \( \varepsilon(x^j_k, t) \) is measurement noise which is assumed that is zero-mean, Gaussian, spatial uncorrelated and white [14, 20], i.e. \( \mathbb{E}[\varepsilon(x^j_k, t)] = 0 \) and \( \text{var}(\varepsilon(x^j_k, t)) = \sigma^2 \) where \( \sigma > 0 \) denotes standard deviation of the measurement noise.
A further assumption is that the estimation of the unknown parameter vector $\theta$ is performed via minimization of the least-squares criterion

$$J(\theta) = \sum_{j=1}^{N} \sum_{k=1}^{K} \int_{T_k} |z^j(t) - \hat{y}(x^j_k, t; \theta)|^2 \, dt,$$

with $\theta \in \Theta_{ad}$ and $\Theta_{ad}$ is the set of admissible parameters and $\hat{y}(\cdot, \cdot; \theta)$ denotes the system model response suitable to $\theta$. The estimate of the true value of $\theta^*$ is a vector $\hat{\theta}$ which minimize $J(\theta)$.

An elementary idea of fault diagnosis is to compare the resulting parameter estimates with the corresponding known nominal values, treating possible differences as residuals which contain information about potential faults. Then using thresholding techniques, the appropriate decision making system could be constructed to detect abnormal situations in system functioning, see [22, 18, 37, 23]. Based on the observations, it is possible to test the simple parametric statistical hypothesis $H^0 : \theta = \theta^*$, where $\theta^*$ is the nominal value of the vector $\theta$ corresponding to the normal system performance.

Introducing the generalization of the likelihood function for the considered experiment[8]:

$$L(z|\theta) = \left(\frac{1}{2\pi\sigma^2}\right)^{N/2} \exp \left(-\frac{1}{2\sigma} J(\theta)\right),$$

and setting $\Theta_0 = \{\theta \in \Theta_{ad} : \theta = \theta^*\}$ it could be define the following generalized log-likelihood ratio:

$$\lambda(z) = 2 \ln \sup_{\theta \in \Theta_{ad}} \frac{L(z|\theta)}{L(z|\theta_0)} = J(\hat{\theta}) - J(\hat{\theta}),$$

where

$$\hat{\theta} = \arg \min_{\theta \in \Theta_{ad}} J(\theta), \quad \hat{\theta} = \arg \min_{\theta \in \Theta_0} J(\theta).$$

The generalized log-likelihood ratio test is widely used in statistics, because it can be shown that assuming the validity of the null hypothesis $H^0$ the sequence $\lambda(z)$ for $N \to \infty$ is weakly convergent to $x^2$ random variable on $m$ degrees of freedom [8, Thm. 3.6.1, p 55]. The meaning of this fact is that for the given significance level $\gamma$ representing a fixed range of model uncertainty we can compare the observed value of $\lambda(z)$ with some threshold $k_\gamma$ obtained from the cumulative chi-squared distribution with $m$ degrees of freedom, where $k_\gamma$ is ‘$\gamma$-quantile’ of the distribution. Then, the decision rule for fault detection takes the following form:

$$S = \begin{cases} S^1 & \text{if } 2 \ln \lambda(y) \geq k_\gamma \quad (\text{reject } H^0), \\ S^0 & \text{if } 2 \ln \lambda(y) < k_\gamma \quad (\text{accept } H^0). \end{cases}$$

The potential rejection of $H^0$ indicates a significant deviation of the vector from the nominal value of this parameter and forms a base for proper detection of abnormal states in the system. When the null hypothesis $H^0$ is true, we do not want to reject it. The mistake in rejecting null hypothesis $H^0$ when it is true is called the Type I error. Similarly, accepting null hypothesis when the alternative hypothesis of the form $H^1 : \theta \neq \theta^*$ is true is called the Type II error [8]. In fault diagnosis, there is a direct connection between these errors and the probability of a false alarm and missed detection, respectively.
3. Optimal measurement problem

In order to achieve a reliable diagnosis on the DPS state we have to keep the probabilities of false alarms and missed detection low. In practice, in order to avoid the multi-criteria optimization, a typical approach is to minimize the Type II error with a constraint imposed on the maximum level of the Type I error.

Based on the parameter vector \( \theta \) it can be shown that the power of the presented hypothesis test (i.e. the probability of not committing the Type II error) can be increased by taking a large number of measurements \( N \) or, alternatively, by maximizing the D-optimality criterion (see [18] for details):

\[
\Psi(M) = -\log \det(M),
\]

defined on the so-called *average* Fisher Information Matrix given by [15]

\[
M = \frac{1}{NT} \sum_{j=1}^{N} \sum_{k=1}^{K} g(x_j^k, t)g^T(x_j^k, t) \, dt,
\]

where

\[
g(x, t) = \left( \frac{\partial y(x, t; \theta)}{\partial \theta} \right)^T_{\theta=\theta^0},
\]

stands for the so-called *sensitivity vector*, \( \theta^0 \) being a prior estimate to the unknown parameter vector \( \theta \). The D-optimality leads to minimization of the volume of the uncertainty ellipsoid for the parameter estimates. Since in real world problems, the number of available observations has to be always limited, thus in practice, such strategy is important alternative to obtain the high powered test which leads directly to the low probability of missed detection.

The introduction of an optimality criterion renders it possible to formulate the sensor location problem as an optimization problem

\[
\Psi[M] \rightarrow \min,
\]

with respect to \( x_j^k, \ j = 1, \ldots, N, \ k = 1, \ldots, K \) belonging to the admissible set \( X \).

Owing to assumption on independent measurement noise, we admit of replicated measurements, i.e. some values \( x_j^k \) may appear several times in the optimal solution (we can use multiple sensors located at the same position or we can repeat the process). Consequently, we introduce \( r_j^k \) as the numbers of replicated measurements corresponding to the sensor nodes locations \( x^1, \ldots, x^N \) at consecutive time subinterval \( T_k \). The collection of variables

\[
\xi_k = \{(x_j^1, p_{kj}), \ (x_j^2, p_{kj}), \ (x_j^N, p_{kjN})\},
\]

where \( p_{kj} = r_j^k/N, \ N = \sum_{j=1}^{N} r_j^k \) is called the *exact design* of experiment for subinterval \( T_k \). The proportion \( p_{kj} \) of observations performed at \( x_j^k \) can be considered as the percentage of experimental effort spent at that point.

Introducing the \( \xi = (\xi_1, \xi_2, \ldots, \xi_K) \) as the total experimental design, we rewrite the FIM in the form

\[
M(\xi) = \frac{1}{K} \sum_{j=1}^{N} \sum_{k=1}^{K} p_{kj} \int_{T_k} g(x_j^k, t)g^T(x_j^k, t) \, dt,
\]

Here the \( p_{kj} \)'s are rational numbers, since both \( r_j^k \)'s and \( N \) are integers. This discrete nature of \( N \)-observation exact designs causes serious difficulties and cannot be solved easily by standard
optimization techniques, particularly when $N$ is large. A reasonable approach is to relax the definition of the design. When $N$ is large, the weights $p_{kj}$ can be approximated with a real numbers in the interval $[0, 1]$, not necessarily integer multiples of $1/N$. Obviously, we must have $\sum_{j=1}^{N} p_{kj} = 1$, for $k = 1, \ldots, K$, so we may think of the designs as probability distributions on $X$. This leads to the so-called continuous designs which constitute the basis of the modern theory of optimal experiments [2, 28]. It turns out that such an approach drastically simplifies the design. Thus, we shall operate on designs which concentrates approximately $Np_{kj}$ measurements at each location $x_k$.

Then we may redefine an optimal design as a solution to the optimization problem

$$\xi^* = \arg \min_{\xi \in \Xi(X)} \Psi[M(\xi)],$$

where $\Xi(X)$ denotes the set of all probability distributions on $X = \{x^1, \ldots, x^N\}$.

**4. Characterization of the optimal solutions**

In the remainder of this paper we shall assume that $g \in C(X \times T; \mathbb{R}^m)$. A number of characterizations of the optimal design $\xi^*$ can be derived in a rather straightforward manner from the general results given in [35, 16].

**Lemma 1.** For any $\xi \in \Xi(X)$ the information matrix $M(\xi)$ is symmetric and non-negative definite. Let us introduce the notation $\mathcal{M}(X)$ for the set of all admissible information matrices, i.e. $\mathcal{M}(X) = \{M(\xi) : \xi \in \Xi(X)\}$.

**Lemma 2.** $\mathcal{M}(X)$ is compact and convex.

The next result constitutes the necessary and sufficient condition for the optimality of designs. It is usually called an equivalence theorem [35, 16].

**Theorem 1.** The following conditions are equivalent:

- the design $\xi^*$ minimizes $\Psi[M(\xi)]$,
- the design $\xi^*$ minimizes $\max_{x \in X} \phi_k(x, \xi^*) = \varsigma_k(\xi^*)$, $k = 1, \ldots, K$, where

$$\phi_k(x, \xi) = \frac{1}{l_f} \int_{T_k} g^T(x, t)M^{-1}(\xi)g(x, t) \, dt,$$

$$\varsigma_k(x, \xi) = \frac{1}{l_f} \sum_{j=1}^{N} \int_{T_k} g^T(x, t)M^{-1}(\xi)g(x, t) \, dt.$$

The sensitivity function $\phi_k(x, \xi)$ is of vital importance here, providing us with a simple test for the optimality of designs. In particular, if it takes the values less than or equal to $\varsigma_k(\xi)$ for all $x \in X$ then $\xi$ is optimal, otherwise $\xi$ is not optimal.

**5. Distributed algorithm**

Analytical determination of optimal designs is possible only in simple situations and for general systems it is usually the case that some iterative design procedure will be required. In the case considered in the paper, i.e. the design for fixed sensor locations, a computational algorithm can be derived based on the mapping $T: \Xi(X) \rightarrow \Xi(X)$ defined by

$$T\xi = \left\{ \left( x_k^1, p_{k1} \frac{\phi(x_k^1, \xi)}{\varsigma_k(\xi)} \right), \ldots, \left( x_k^N, p_{kN} \frac{\phi(x_k^N, \xi)}{\varsigma_k(\xi)} \right) \right\}, \quad k = 1, \ldots, K.$$
From Theorem 1 it follows that a design $\xi^*$ is D-optimal if it is a fixed point of the mapping $\mathcal{T}$, i.e.

$$\mathcal{T}\xi^* = \xi^*.$$  

(17)

Also, we have the valuable property that the sum of weights $p_{kj}$ is invariant with respect to $\mathcal{T}$ for any $k = 1, \ldots, K$ (i.e. it is always equal to 1). As for interpretation of this mapping, let us note that in case of $\xi$ not being the optimal design, it increases the weights for those sensor locations where we observe the high values of sensitivity function at the cost of decreasing the weights of non-informative locations with small values of $\phi$. Using this simple idea, it is possible to develop a decentralized configuration algorithm for sensor network being a distributed generalization of that proposed in [25, p. 139] for the classical optimum experimental design problem consisting in iterative computation of a D-optimum design on a finite set (and further extended in [35, 16]).

In the following we assume the asynchronous time model for the configuration process. Let $r = 0, 1, 2, \ldots$ be the discrete time index, which partition the continuous configuration time axis into time slots $Z_r = [z_{r-1}, z_r)$.

Owing to Theorem 1 we can use the mapping $\mathcal{T}$ to iteratively improve the experimental effort distribution. Unfortunately $\phi_k(\cdot, \xi)$ and $\varphi_k(\xi)$ computed according to (14) in a centralized fashion are not very useful in our setting. But the only component which cannot be calculated independently of other network nodes is the global information matrix $M(\xi)$. Moreover, from (12) it is clearly seen the information matrix is a weighted average of the subinterval information matrices

$$M_k = \frac{1}{t_f} \sum_{j=1}^{N} p_{kj} \int_{T_k} g(x^j_k, t) g^T(x^j_k, t) \, dt.$$  

(18)

Thus, our task is closely related to the problem of distributed averaging on a network which appears in many applications and has been a subject of extensive studies [40, 5, 6].

Distributed averaging can be achieved in many ways. One of straightforward techniques is a pairwise communication flooding, also known as a gossip scheme, which in its classic version assumes that at the $k$-th reconfiguration time slot the $i$-th sensor contacts some neighboring node $j$ with probability $Q_{ij}$ i.e. a pair $(i-j)$ is randomly and independently selected. At this time, both nodes set their values equal to the average of their current values. Denoting the estimate of (18) for subinterval $T_k$ maintained by $j$-th sensor at configuration time slot $Z_r$ as $M_{kj}^l(\xi^{(r)})$ we have ($\leftarrow$ is an update operator).

$$M_{kj}^l(\xi^{(r)}) \leftarrow \frac{p_{ki}M_{kj}^i(\xi^{(r)}) + p_{kj}M_{kj}^j(\xi^{(r)})}{p_{ki} + p_{kj}}, \quad \ell \in \{i, j\}, \quad k = 1, \ldots, K.$$  

(19)

However, in our setting the averaging problem is not typical as the contribution of the nodes with weights tending to zero as $r$ increases should be eliminated from the averaging. This can be achieved by increasing the contribution of the local information matrix via the concept of so-called running consensus [6]

$$M_{kj}^l(\xi^{(r)}) \leftarrow \frac{r - 1}{r} M_{kj}^i(\xi^{(r)}) + \frac{1}{r} \Upsilon_k(x^j_k),$$  

(20)

where $\Upsilon_k(x^j_k) = \int_{T_k} g(x^j_k, t) g^T(x^j_k, t)$. The first term enforces consensus among the nodes (represents the average information from the rest of network), while the second accounts for the increase in the total contribution of the local node. Note, that there is no need to record the whole design structure of $\xi^{(r)}$ at each node as the information about the current solution is propagated via local estimates of global FIM, i.e. $M_{kj}^i(\xi^{(r)})$. The whole idea of the communication process is given by the Algorithm 1.
For all subintervals $T_k$ data are calculated separately to reach most informative configuration for mentioned part of experiment. At $t = 0$ each network node starts with a global FIM estimate $M_k^i(\xi^{(r)})$ initialized with its local information matrix $Y_k(x_k^i)$ and starting with non-zero weights (it is important since once the weight attains zero, a multiplicative update forces it to remain zero). Then at each of asynchronous configuration time slot $Z_r$ an appropriate pair of nodes exchanges information according to Algorithm 1. It can be shown that algorithm converges to the optimal solution, what is a consequence of independent weight update for each network node and the stochastic convergence of gossip algorithms[40].

Obviously, the sensor network topology and the choice of proper communication scheme significantly influence on the convergence rate. In general, under some assumptions on the network connectivity graph, a suitable gossip algorithm can be provided via the semidefinite programming with an exponential convergence rate. Since discussion on those very important issues is far beyond the scope of this work, therefore we refer the reader to the seminal papers [40, 5] for details.

6. Simulation example

As an illustration of the presented approach, consider the problem of sensor configuration for parameter estimation in the process of air pollutant transport over a given urban area. Within this domain, which has been normalized to the unit square $\Omega = (0, 1)^2$, an active source of pollution is present, which yields the pollutant spatial concentration $y = y(x, t)$. The evolution of $y$ over the normalized observation interval $T = (0, 1]$ is described by the following advection-diffusion equation:

$$\frac{\partial y(x, t)}{\partial t} + \nabla \cdot (v(x, t)y(x, t)) = \nabla \cdot (\alpha(x)\nabla y(x, t)) + f(x), \quad x \in \Omega$$  

subject to the boundary and initial conditions:

$$\frac{\partial y(x, t)}{\partial n} = 0, \text{ on } \Gamma \times T,$$

$$y(x, 0) = y_0, \text{ in } \Omega,$$

where term $f(x) = 50 \exp(-50||x - c||^2)$ represents the pollutant source located at the point $c = (0.3, 0.3)$ and $\partial y/\partial n$ stands for the partial derivative of $y$ with respect to the outward normal.
to the boundary $\Gamma$. The mean spatiotemporal changes of the wind velocity field over the area were approximated by $v(x,t) = (2(x_1 + x_2 - t), x_2 - x_1 + t)$. The assumed functional form of the spatial-varying diffusion coefficient is $a(x) = \theta_1 + \theta_2 x_1 x_2 + \theta_3 x_1^2 + \theta_4 x_2^2$. The subject of interest here is a detection of significant increase in the intensity of the pollutant emission from the source. As the symptom of this abrupt fault we can observe the excessive deviation of the parameters $\theta$ from their nominal values. Therefore, these parameters need estimation based on measurement data from monitoring stations.

A MATLAB program was written using a PC equipped with Intel Core i5 processor (1.9GHz, 12 GB RAM) running Windows 10. The nominal values of the systems parameters were assumed to be $\theta_1^0 = 0.02$, $\theta_2^0 = 0.01$, $\theta_3^0 = 0.005$. Calculations were performed with a finite element method for a spatial mesh composed of 682 triangles, 378 nodes and an evenly partitioned time interval (3 subintervals $T_k = ((k-1)/3,k/3)$, $k = 1,2,3$). The sensitivity coefficients were then linearly interpolated and stored within the sensor nodes repositories. In our scenario, the observation grid was assumed to be created at locations selected from among those elements of above-mentioned 378-point triangulation mesh which do not lie on the outer boundary (there were 312 such nodes, which are indicated with dots in Fig. 2). Given $N = 312$ prospective sites in $\Omega \cup \Gamma$ we aim at finding the locations at which the measurements made by sensors would lead to D-optimum least-squares estimates of the parameters $\theta$. In simulation it was assumed that the network is fully connected with uniform probability distribution for the connection between selected pair of nodes. Sensor configurations at different stages of algorithm are shown in Fig. 2. The sensors tend to form the pattern reflecting the areas of greatest changes in the pollutant concentration but the observations are averaged over time and it is not trivial to follow the dynamics of the observation strategy. It can be observed that during each time subinterval the system dynamics enforce the distribution of experimental effort. In the first stage ($T_1$), the pollutant concentration takes significant values only in the very close vicinity of the source, thus all measurement effort is spent at this location (cf. Figs. 2(a)-(c)). As time elapses the pollutant spreads over the whole spatial domain, therefore the sensors tends to measure the system state in other regions, where we observe the greatest changes in concentration and gather the rich information about the system dynamics (Figs. 2(d)-(f)). Note, that algorithm approximate the pattern of final design very fast after reaching relative small number of $r$. Then, the algorithm slows down and achieving a very high accuracy of the solution becomes expensive in the sense of the number of pairwise communications. An application of a more efficient communication scheme at the latter stage of configuration process would be helpful here to overcome this effect.

7. Concluding remarks
The problem of selecting optimal observation points in view of reliable detection of abnormal state of DPS has been addressed. The main contribution of this work is a proper characterization of problem relaxation and its adaptation to a distributed scheme of computation and information exchange. The advantage of introducing continuous designs lies in the fact that the problem dimensionality is dramatically reduced. A decided majority of existing approaches to the sensor location problem for DPSs is designed only for centralized monitoring systems, and in the context of modern sensor networks often this comes out impractical. On the other hand, numerous effective routines for distributed information fusion designed for sensor, peer-to-peer or wireless ad hoc networks are not directly applicable for the purpose of parameter estimation in DPSs. The decided advantage of the proposed distributed calculation scheme is a great simplicity and the ability to produce solutions reasonably fast.

Acknowledgement
This work is supported by National Science Center in Poland under grant 2014/15/B/ST7/03208.
Figure 1. Temporal changes in the wind velocity field and pollutant concentration

References
[1] A. Atkinson, A. Donev, and R. Tobias. *Optimum experimental designs, with SAS*, volume 34. Oxford University Press, 2007.
[2] A. C. Atkinson, A. N. Donev, and R. D. Tobias. Optimum experimental designs, with sas. 2007.
[3] M. Basseville, I. V. Nikiforov, et al. *Detection of abrupt changes: theory and application*, volume 104. Prentice Hall Englewood Cliffs, 1993.
[4] P. H. Bauer. New challenges in dynamical systems: The networked case. 18(3):271–277, 2008.
[5] S. Boyd, A. Ghosh, B. Prabhakar, and D. Shah. Randomized gossip algorithms. *IEEE/ACM Transactions on Networking (TON)*, 14(SI):2508–2530, 2006.
[6] P. Braca, S. Marano, and V. Matta. Enforcing consensus while monitoring the environment in wireless sensor networks. *IEEE Transactions on Signal Processing*, 56(7):3375–3380, 2008.
[7] L. H. Chiang, R. D. Braatz, and E. L. Russell. *Fault detection and diagnosis in industrial systems*. Springer Science & Business Media, 2001.
[8] G. C. Goodwin and R. L. Payne. Dynamic system identification: experiment design and data analysis. 1977.
[9] J. Korbicz and J. M. Kościelny. *Modeling, diagnostics and process control: implementation in the diaster*
Figure 2. Allocation of active sensors (red circles) with experimental effort (blue stem plot) in consecutive time subintervals for different number of communications $r$. 

system. Springer Science & Business Media, 2010.

[10] D. Kowalów, M. Patan, W. Paszke, and A. Romanek. Sequential design for model calibration in iterative learning control of dc motor. In Methods and Models in Automation and Robotics (MMAR), 2015 20th International Conference on, pages 794–799. IEEE, 2015.

[11] D. Kowal´ ow, M. Patan. Optimal sensor selection for model identification in iterative learning control of spatio-temporal systems. In Methods and Models in Automation and Robotics - MMAR 2016 : 21 international conference; ISBN: 9781509018673, pages 70–75, Midzyzdroje, Polska, 2016. [B. m.], [B. w.].

[12] A. Nehorai, B. Porat, and E. Paldi. Detection and localization of vapor-emitting sources. IEEE Transactions on Signal Processing, 43(1):243–253, 1995.

[13] P. Ögren, E. Fiorelli, and N. E. Leonard. Cooperative control of mobile sensor networks: Adaptive gradient climbing in a distributed environment. 49(8):1292–1302, 2004.

[14] M. Patan. A parallel sensor scheduling technique for fault detection in distributed parameter systems. In Euro-Par 2008–Parallel Processing, pages 833–843. Springer, 2008.

[15] M. Patan. Distributed scheduling of sensor networks for identification of spatio-temporal processes. International Journal of Applied Mathematics and Computer Science, 22(2):299–311, 2012.

[16] M. Patan. Optimal sensor networks scheduling in identification of distributed parameter systems, volume
425. Springer Science & Business Media, 2012.

[17] M. Patan and D. Kowalów. Robust sensor scheduling via iterative design for parameter estimation of distributed systems. In Methods and Models in Automation and Robotics (MMAR), 2014 19th International Conference On, pages 618–623. IEEE, 2014.

[18] M. Patan and K. Patan. Optimal observation strategies for model-based fault detection in distributed systems. International Journal of Control, 78(18):1497–1510, 2005.

[19] M. Patan, C. Tricaud, and Y. Chen. Resource-constrained sensor routing for parameter estimation of distributed systems. In Proc. 17th IFAC World Congress, 2008.

[20] M. Patan and D. Uciński. Time-constrained sensor scheduling for parameter estimation of distributed systems. In Decision and Control (CDC), 2010 49th IEEE Conference on, pages 7–12. IEEE, 2010.

[21] M. Patan and D. Uciński. Cost-constrained d-optimum node activation for large-scale monitoring networks. In Decision and Control (CDC), 2016 55th IEEE Conference on, pages 1643–1648. IEEE, 2016.

[22] M. Patan and D. Uciński. Optimal activation strategy of discrete scanning sensors for fault detection in distributed-parameter systems. In Proceedings of the 16th IFAC world congress, Prague, Czech Republic, pages 4–8, 2005.

[23] M. Patan and D. Uciński. Configuring a sensor network for fault detection in distributed parameter systems. International Journal of Applied Mathematics and Computer Science, 18(4):513–524, 2008.

[24] R. J. Patton, P. M. Frank, and R. N. Clark. Issues of fault diagnosis for dynamic systems. Springer Science & Business Media, 2013.

[25] A. Pázman. Foundations of optimum experimental design, volume 14. Springer, 1986.

[26] N. Point, A. V. Wouwer, and M. Remy. Practical issues in distributed parameter estimation: Gradient computation and optimal experiment design. Control Engineering Practice, 4(11):1553–1562, 1996.

[27] B. Porat and A. Nehorai. Localizing vapor-emitting sources by moving sensors. Signal Processing, IEEE Transactions on, 44(4):1018–1021, 1996.

[28] L. Pronzato and E. Walter. Identification of parametric models from experimental data. Springer, Berlin.

[29] E. Rafajlowicz. Optimum choice of moving sensor trajectories for distributed-parameter system identification. International Journal of Control, 43(5):1441–1451, 1986.

[30] A. Romanek, M. Patan, and D. Kowalów. Decentralized scheduling of sensor networks for parameter estimation of spatio-temporal processes. Advanced and Intelligent Computations in Diagnosis and Control, 386:145, 2015.

[31] Z. Song, Y. Chen, C. R. Sastry, and N. C. Tas. Optimal observation for cyber-physical systems: a fisher-information-matrix-based approach. Springer Science & Business Media, 2009.

[32] C. Tricaud and Y. Chen. Optimal mobile sensing and actuation policies in cyber-physical systems. Springer Science & Business Media, 2011.

[33] C. Tricaud, M. P. Dariusz, U. Yang, and Q. Chen. D-optimal trajectory design of heterogeneous mobile sensors for parameter estimation of distributed systems. In 2008 American Control Conference, pages 663–668. IEEE, 2008.

[34] D. Uciński. Optimal selection of measurement locations for parameter estimation in distributed processes. International Journal of Applied Mathematics and Computer Science, 10(2):357–379, 2000.

[35] D. Uciński. Optimal measurement methods for distributed parameter system identification. CRC Press, 2004.

[36] D. Uciński. Sensor network scheduling for identification of spatially distributed processes. International Journal of Applied Mathematics and Computer Science, 22(1):25–40, 2012.

[37] D. Uciński and M. Patan. Sensor network design for the estimation of spatially distributed processes. International Journal of Applied Mathematics and Computer Science, 20(3):459–481, 2010.

[38] É. Walter and L. Pronzato. Qualitative and quantitative experiment design for phenomenological models a survey. Automatica, 26(2):195–213, 1990.

[39] M. Witzczak, M. Buciaikowski, V. Puig, D. Rotondo, and F. Nejjar. An lmi approach to robust fault estimation for a class of nonlinear systems. International Journal of Robust and Nonlinear Control, 2015.

[40] L. Xiao and S. Boyd. Fast linear iterations for distributed averaging. Systems & Control Letters, 53(1):65–78, 2004.