G Model
JJPC-2143; No. of Pages 12
ARTICLE IN PRESS
Journal of Process Control xxx (2017) xxx–xxx
Contents lists available at ScienceDirect
Journal of Process Control
journal homepage: www.elsevier.com/locate/jprocont

A scalable design of experiments framework for optimal sensor placement

Jing Yu [a], Victor M. Zavala [b,*], Mihai Anitescu [a,c]

[a] Department of Statistics, University of Chicago, 15734 S. University Avenue, Chicago, IL 60637, USA
[b] Department of Chemical and Biological Engineering, University of Wisconsin-Madison, 1415 Engineering Dr., Madison, WI 53706, USA
[c] Mathematics and Computer Science Division, Argonne National Laboratory, 9700 S. Cass Ave., Lemont, IL 60439, USA

ARTICLE INFO

Article history:
Received 16 April 2016
Received in revised form 8 January 2017
Accepted 23 March 2017
Available online xxx

Keywords:
Sensor data
Sensor placement
Estimation
Scalable

ABSTRACT

We present a scalable design of an experiments framework for sensor placement in systems described by partial differential equations (PDEs). In particular, we aim to compute optimal sensor locations by minimizing the uncertainty of parameters estimated from Bayesian inverse problems. The resulting problem is a computationally intractable mixed-integer nonlinear program constrained by PDEs. We approach this problem with two heuristics used in compressed sensing and optimal control literature: a sparsity-inducing approach and a sum-up rounding approach. We also investigate metrics to guide the design of experiments (the total flow variance and the A-optimal design criterion) and analyze the effect of different noise structures (white and colored). Using an application in natural gas pipelines, we conclude that the sum-up rounding approach gives the best results and produces shrinking gaps with increasing mesh resolution. We also observe that convergence for the white noise measurement error case is slower than for the colored noise case. For A-optimal design, the solution is close to a uniform distribution of sensors along the pipeline while for the flow variance design the distribution is unstructured.

1. Introduction

The sensor placement problem seeks to determine the optimal number, locations, and types of sensors that would maximize information about a dynamical system. Because information can often be expressed in terms of the posterior covariance matrix of the states or parameters of the system, the problem can often be cast as an optimal design of experiments problem. Such a problem is computationally challenging, particularly in the infinite-dimensional case, because one must solve a mixed-integer and bilevel optimization problem constrained by differential algebraic equations or by partial differential equations (PDEs). This problem has been addressed by using mixed-integer programming techniques for contaminant detection in water networks [3,2,22,10]. In these studies, an optimal set of sensor locations is selected from a set of candidate locations to minimize a certain engineering metric such as contaminant detection time, population exposure, or likelihood of detection. Likelihoods are assigned based on contamination scenarios, and not on information content of the sensor data recorded, as in a traditional experimental design setting. As a result, these approaches fail to provide statistically meaningful sensor network designs. Moreover, because the formulations capture flow dynamics by using surrogate representations such as transportation delays, they fail to capture physical effects.

Sensor placement problems have also been addressed in a more general control setting where one seeks to optimize a measure of observability such as the covariance matrix, Kalman estimator gain, or the so-called observability Grammian matrix. This problem is again a bilevel optimization problem. The covariance matrix approach in [6] bypasses this by assuming that the dynamic model is linear, thus allowing the inner minimization problem to be formulated as a linear matrix inequality. The approach in [20] models the dynamics of the covariance matrix directly as a Riccati differential equation, which implicitly assumes linearity and thus enables the use of semidefinite programming algorithms. This approach, however, is focused on control policy design to extract maximum information, and not on sensor placement design. Consequently, the authors do not consider discrete decisions associated to placement. A rigorous treatment of nonlinear dynamics is presented in [19] by casting the problem as a mixed-integer nonlinear program. The authors use a genetic algorithm to deal with the inner minimization problem that computes the observability metric. A similar approach is used in [14] to address the inner minimization problem. Mixed-integer techniques have also been used in the context

* Corresponding author.
E-mail address: victor.zavala@wisc.edu (V.M. Zavala).

http://dx.doi.org/10.1016/j.jprocont.2017.03.011

of information maximization for Gaussian processes and for designing Latin hypercube samples [11,8]. These approaches do not use physical models.

Recently, the sensor placement problem for systems described by PDEs has been cast as an A-experimental design problem in which the number of sensors (i.e., the design cost) is controlled by using an sparsity-inducing ℓ_0 regularization norm that is in turn approximated by using a smoothing function [1]. This compressed sensing approach was shown to be scalable and applicable to infinite-dimensional systems, but it requires tuning and can be numerically unstable. One can also formulate and solve the problem as a mixed-integer programming problem directly, but this is computationally intractable because the PDEs are in general nonconvex and because the problem has a bilevel nature.

An important application of optimal sensor location techniques is infrastructure networks (oil, water, gas, and electricity) in which large amounts of sensor data need to be processed in real time in order to reconstruct the state of the system or to identify leaks, faults, or attacks. In this work we focus on natural gas networks, which are used to transport fuel to power generation facilities and urban areas from storage and processing facilities. These networks comprise pipelines that span thousands of miles and exhibit complex dynamics. An interesting property of natural gas networks is that significant amounts of gas can be stored inside the pipelines. The stored gas is distributed spatially along the pipelines and is normally referred to as *line-pack* [5]. Line-pack is used by pipeline operators to modulate variations of gas demands at multiple spatial points in intraday operations. Some of the strongest variations in gas demands are the result of on-demand startup and shutdown of gas-fired power plants [15]. Modulating these variations is challenging because the fast release of line-pack at multiple simultaneous locations can trigger complex spatiotemporal dynamic responses that propagate hundreds to thousands of miles and that can take hours to stabilize. Therefore, line-pack management is performed by using sophisticated optimal control and pipeline simulation tools. To use these automation tools, one must reconstruct spatiotemporal state fields (flows, pressures, temperatures) [16] and detect natural gas leaks [4]. This task is challenging from a practical stand point given the limited amounts of sensor data (often limited to pressure and flow signals at a finite set of locations), the infinite-dimensional nature of pipeline systems, and the complex physical behavior of these systems. Such challenges are not unique to natural gas networks but also arise in other domains such as geophysics and contaminant source detection in water networks.

In this work we present a scalable design of experiments framework to compute optimal sensor locations for systems described by PDEs. This is done by minimizing the uncertainty in the state and of parameters estimated from Bayesian inverse problems. The resulting problem is a mixed-integer infinite-dimensional optimal control problem. We approach this problem by using two efficient heuristics that have the potential to be scalable for such problems: a sparsity-inducing approach used in machine learning [1] and a sum-up rounding approach used in optimal control [17]. We investigate two objectives: the total flow variance and the A-optimal design criterion. Using a natural gas pipeline case study, we conclude that the sum-up rounding approach produces shrinking gaps with finer meshes. We also observe that convergence for the white noise measurement error is slower than for the colored noise case. For the A-optimal design the solution is close to the uniform distribution, but for the total flow variance the pattern is noticeably different.

The paper is structured as follows. In Section 2 we define the physical system model that we use for sensor placement. In Section 3 we provide the formulations of the design of experiments problems that we aim to solve. In Section 4 we present numerical experiments with machine learning and the sum-up rounding procedure for solving the design of experiments problem. In Section 5 we summarize our conclusions and briefly describe future work.

2. Distributed system modeling

We illustrate the complexity of the optimal sensor placement problem by focusing on the physical equations describing the dynamics of natural gas networks. Details on the model derivation, nomenclature, and units used in this section can be found in [24].

2.1. Problem physics

The isothermal flow of gas through a horizontal pipeline is described by the conservation and momentum equations:

$$\frac{\partial \rho(\tau,x)}{\partial \tau}+\frac{\partial(\rho(\tau,x)\nu(\tau,x))}{\partial x}=0 \tag{2.1a}$$

$$\frac{\partial(\rho(\tau,x)\nu(\tau,x))}{\partial \tau}+\frac{\partial p(\tau,x)}{\partial x}=-\frac{\lambda}{2D}\rho(\tau,x)\nu(\tau,x)|\nu(\tau,x)|. \tag{2.1b}$$

Here, $\tau\in\mathcal{T}:=[0,T]$ is the time dimension with final time T (planning horizon), and $x\in\mathcal{X}:=[0,L]$ is the axial dimension with length L. The pipeline diameters are denoted as D, and the friction coefficients are denoted as λ. The states of the link are the gas density $\rho(\tau,x)$, the gas speed $\nu(\tau,x)$, and the gas pressure $p(\tau,x)$. The transversal area A, volumetric flow $q(\tau,x)$, and mass flow $f(\tau,x)$ are given by

$$A=\frac{1}{4}\pi D^2 \tag{2.2a}$$

$$q(\tau,x)=\nu(\tau,x)A \tag{2.2b}$$

$$f(\tau,x)=\rho(\tau,x)\,\nu(\tau,x)A. \tag{2.2c}$$

For an ideal gas, pressure and density are related as follows:

$$\frac{p(\tau,x)}{\rho(\tau,x)}=c^2. \tag{2.3}$$

Here, c is the gas speed of sound. The speed (assuming an ideal gas behavior) and the friction factor λ can be computed from

$$c^2=\frac{\bar{\gamma}ZRT}{M} \tag{2.4a}$$

$$\lambda=\left(2\log_{10}\left(\frac{3.7\,D}{\epsilon}\right)\right)^{-2}, \tag{2.4b}$$

where Z is the gas compressibility factor, R is the universal gas constant, T is the gas temperature, M is the gas molar mass, ϵ is the pipe rugosity, and $\bar{\gamma}$ is the adiabatic constant. Often one desires to transform (2.1) into a more convenient form in terms of mass flow and pressure by using (2.3) and (2.2):

$$\frac{\partial p(\tau,x)}{\partial \tau}+\frac{c^2}{A}\frac{\partial f(\tau,x)}{\partial x}=0 \tag{2.5a}$$

$$\frac{1}{A}\frac{\partial f(\tau,x)}{\partial \tau}+\frac{\partial p(\tau,x)}{\partial x}=-\frac{\lambda\rho(\tau,x)}{2D}\frac{f(\tau,x)}{\rho(\tau,x)A}\left|\frac{f(\tau,x)}{\rho(\tau,x)A}\right|. \tag{2.5b}$$

Substituting (2.3) and (2.2a) in (2.5b) and performing some manipulations, we obtain the more compact form:

$$\frac{\partial p(\tau,x)}{\partial \tau}=-\frac{c^2}{A}\frac{\partial f(\tau,x)}{\partial x} \tag{2.6a}$$

$$\frac{1}{A}\frac{\partial f(\tau,x)}{\partial \tau}=-\frac{\partial p(\tau,x)}{\partial x}-\frac{8\,\lambda\,c^2}{\pi^2\,D^5}\frac{f(\tau,x)|f(\tau,x)|}{p(\tau,x)}. \tag{2.6b}$$

For numerical purposes, we define scaled flows $f(\tau,x)\leftarrow\alpha_f f(\tau,x)$ and pressures $p(\tau,x)\leftarrow\alpha_p p(\tau,x)$, where α_f and α_p are scaling factors. Scaling (2.6) and rearranging, we obtain the final form:

$$\frac{\partial p(\tau,x)}{\partial \tau}=-c_1\frac{\partial f(\tau,x)}{\partial x},\quad \tau\in\mathcal{T},\quad x\in\mathcal{X} \tag{2.7a}$$