Automatica 72 (2016) 166–176

Contents lists available at ScienceDirect

Automatica

journal homepage: www.elsevier.com/locate/automatica

Sensor placement for fault location identification in water networks: A minimum test cover approach☆

Lina Sela Perelman[a], Waseem Abbas[b], Xenofon Koutsoukos[b], Saurabh Amin[a]

[a] *Massachusetts Institute of Technology, United States*

[b] *Vanderbilt University, United States*

ARTICLE INFO

Article history:
Received 12 May 2015
Received in revised form
21 March 2016
Accepted 22 May 2016

Keywords:
Fault identification
Minimum test cover
Water networks

ABSTRACT

This paper focuses on the optimal sensor placement problem for the identification of pipe failure locations in large-scale urban water systems. The problem involves selecting the minimum number of sensors such that every pipe failure can be uniquely localized. This problem can be viewed as a minimum test cover (MTC) problem, which is NP-hard. We consider two approaches to obtain approximate solutions to this problem. In the first approach, we transform the MTC problem to a minimum set cover (MSC) problem and use the greedy algorithm that exploits the submodularity property of the MSC problem to compute the solution to the MTC problem. In the second approach, we develop a new *augmented greedy* algorithm for solving the MTC problem. This approach does not require the transformation of the MTC to MSC. Our augmented greedy algorithm provides in a significant computational improvement while guaranteeing the same approximation ratio as the first approach. We propose several metrics to evaluate the performance of the sensor placement designs. Finally, we present detailed computational experiments for a number of real water distribution networks.

1. Introduction

Infrastructure deterioration, demand-supply uncertainty, and risk of disruptions pose new challenges in maintaining modern infrastructures. Resilient urban infrastructures including water distribution systems, transportation networks, and electric grids are crucial for societal well-being. *Smart* infrastructure operation driven by sensing and actuation technologies have been identified as one of the primary solutions towards resilient urban systems (Pandharipande, Calabrese, Lim, & Rajagopal, 2014; Zheng & Kleiner, 2014). Through a network of sensors, an individual fault or correlated failures in a system component can be detected and localized, and restorative actions can be executed in response to these faults. Whereas network observability for a given sensing capability has been widely studied in the context of fault detection, sensor placement for fault isolability, i.e. the ability to distinguish between faults, has not been a commonly studied problem, especially in the context of pipe bursts in water distribution networks.

The goal of this work is to *design a sensor placement configuration for identification of pipe failure locations by using the minimum number of sensors.* The underlying idea behind our approach is to ensure that the sensor placement results in a collective output that is *unique* for each failure event. Specifically, our main contributions are as follows, we:

- Define the *localization* of pipe bursts as the design objective of a sensor network configuration, and using ideas from combinatorial optimization, we formulate the fault location identification problem as a *minimum test cover* (MTC) problem;
- Develop a computationally efficient *augmented greedy* algorithm to solve the minimum test cover problem (resp. identification problem), which is significantly faster in comparison to the previous approaches and therefore, scalable to large-scale networks; and
- Test and evaluate our sensor placement approach on a batch of real-networks of various sizes and parameters using practically relevant performance measures.

☆ This work was supported in part by FORCES (Foundations Of Resilient Cyber-Physical Systems), which receives support from the National Science Foundation (NSF award numbers CNS-1238959, CNS-1238962, CNS-1239054,CNS-1239166), the AFRL LABLET-Science of Secure and Resilient Cyber-Physical Systems (Contract ID: FA8750-14-2-0180, SUB 2784-018400), and the NSF CAREER award #1453126. The material in this paper was not presented at any conference. This paper was recommended for publication in revised form by Associate Editor Michele Basseville under the direction of Editor Torsten Söderström.

E-mail addresses: linasela@mit.edu (L. Sela Perelman), waseem.abbas@vanderbilt.edu (W. Abbas), Xenofon.Koutsoukos@vanderbilt.edu (X. Koutsoukos), amins@mit.edu (S. Amin).

http://dx.doi.org/10.1016/j.automatica.2016.06.005
0005-1098/

Nomenclature

C_i^t	Set of pair-wise link failures detected by the sensor i
C_i	Set of link failures detected by the sensor i
$\mathcal{C}$	Collection of all C_i's
$\mathcal{C}^t$	Collection of all C_i^t's
f_D	Detection function
f_I	Identification function
h	Hydraulic head
I_D	Normalized detection score
I_I	Normalized identification score
I_L	Normalized localization score
I_W	Number of elements in the largest localization set
k	Maximum number of link failures detected by any sensor
ℓ_j	jth (failure) event
ℓ_{ij}^t	Unordered pair of (failure) events ℓ_i and ℓ_j
$\mathcal{L}$	Set of all (failure) events
$\mathcal{L}^t$	Set of all pair-wise (failure) events
L	Localization set
m	Total number of sensors
$\mathcal{M}$	Influence matrix
$\mathcal{M}^t$	Transformed influence matrix
n	Total number of events
p	Pressure
q	Flow
S_i	The location of the ith sensor
$\mathcal{S}$	Set of all sensors
$\mathbf{y}_{\mathcal{S}}$	Outputs of sensors in the set $\mathcal{S}$

Our paper is motivated by the need to consider localization of pipe bursts in the deployment phase of new sensing technologies, since this consideration can significantly reduce the response time and overall costs of fault localization to the distribution utilities. We base our work on the use of low-cost, high-rate online sensors measuring water pressure for remote detection of pipe burst using data mining techniques. Real-world examples are the PIPENET in Boston, MA, US (Stoianov, Nachman, Madden, & Tokmouline, 2007) and the WaterWise in Singapore (Allen et al., 2011). The sensor placement problem is not unique to the water sector and can be found in many engineering applications for system operation. We discuss some of the related work in Section 7.

In Section 2, we present the network and the sensing models and formulate the detection and identification problems as the minimum set cover (MSC) and minimum test cover (MTC) problems, respectively. A key aspect of the MTC problem formulation is the choice of the objective function, which is to select the minimum number of *tests* from a collection of tests such that every event can be uniquely classified in one of the given *categories* based on selected tests' outcomes (Moret & Shapiro, 1985). In our setup, the set of outcomes of tests comprises of the output vector from sensors, events are pipe failures, and classification categories are the possible locations of the failed pipes. In Section 3, we present a solution approach as in Halldórsson, Halldórsson, and Ravi (2001) and Svärd, Nyberg, and Frisk (2013), in which the MTC is first transformed to the MSC and then solved using the greedy approximation (Minoux, 1978).

In Section 4 we present an *augmented greedy* algorithm for solving the MTC that does not require the complete transformation of the MTC to the equivalent MSC, and directly computes the objective function in a greedy fashion. This algorithm is much faster than the standard greedy approach and considerably improves the scalability of our approach. In Sections 5 and 6, we demonstrate our approach using a benchmark and a batch of twelve real water distribution networks of various sizes and specifications. We suggest four metrics to evaluate the performance of the design including detection, identification, and localization scores. Although we demonstrate our results in the context of water networks, our algorithm provides an improved solution to the generic test cover problem. Section 8 summarizes our work and proposes future extensions.

2. Problem formulation

Consider the problem of placing online sensors measuring hydraulic pressures with high frequency such that the identification of pipe failure locations is maximized. Based on the number of pipes where link failures (i.e., pipe bursts) can happen, we consider n link failures as a set of failure events, denoted by $\mathcal{L} = \{\ell_1, \ldots, \ell_n\}$. For the ease of presentation and without the loss of generality, let ℓ_j denote the failure event at the jth pipe. Moreover, we define a set of sensors that can be placed at m nodes of the network as $\mathcal{S} = \{S_1, \ldots, S_m\}$. Here, S_i denotes the location of the ith sensor. The outputs from sensors, which are based on the change in pressure induced by the failure event, are denoted by $\mathbf{y}_{\mathcal{S}}$.

2.1. Network dynamics and sensing model

A water distribution network can be represented by a graph comprising nodes (supply and demand) connected by links (pipes, valves, and pumps). Physical failures of the infrastructure, such as pipe bursts, cause a disturbance in the flow, which moves through the system as a pressure wave known as *water hammer*, or *surge* with very high velocity, varying typically in the range of 600–1500 ($\frac{\mathrm{m}}{\mathrm{s}}$) (Misiunas, 2005). This implies that the steady state analysis employed by traditional methods such as supervisory control and data acquisition (SCADA) systems are inadequate and that the transient system dynamics between the initial and the final steady state conditions need to be considered.

The transient system state can be typically described by mass and momentum partial differential equations (Wylie, Streeter, & Suo, 1993). The method of characteristics (MOC) is a numerical technique typically used to approximate the solution of the hydraulic transients. The MOC transforms the partial differential equations into ordinary differential equations that evolve along specific characteristic lines of the numerical grid, which are solved explicitly to compute the head and flow, $h_{i,t+1}$, $q_{i,t+1}$, at new point in time and space. Here, t and i indicate the discrete points of the numerical grid. For a given pipe, the two characteristic equations describing the hydraulic transients are formulated as Misiunas (2005):

$$h_{i,t+1} = \frac{1}{2}\Big[h_{i-1,t} + h_{i+1,t} + b\left(q_{i-1,t} - q_{i+1,t}\right) + r\left(q_{i+1,t}|q_{i+1,t}| - q_{i-1,t}|q_{i-1,t}|\right)\Big] \quad (1)$$

$$q_{i,t+1} = \frac{1}{b}\Big[h_{i,t+1} - h_{i+1,t} + q_{i+1,t} - r|q_{i+1,t}|\Big], \quad (2)$$

where r is the resistance coefficient associated with the steady state, and b is the impedance coefficient associated with the transient state. For $b = 0$ the set of Eqs. (1), (2) is reduced to the steady state, where the head loss along a pipe occurs only due to friction (Todini & Rossman, 2013). Additional information describing transient dynamics can be found in the supporting information (SI) (Sela Perelman, Abbas, Koutsoukos, & Amin, 2015).

The effect of a pipe burst at location i can be translated into boundary conditions using the orifice head-flow relation (Wylie et al., 1993). Before the burst occurs, the cross-section area of the orifice is equal to zero and it increases during a burst, hence we can expect a sudden change in the hydraulic head. The relationship between the head and the pressure, measured by the sensors at