Multi-objective optimization for conjunctive placement of hydraulic and water quality sensors in water distribution systems

Ami Preis, Andrew Whittle and Avi Ostfeld

ABSTRACT

Near real-time continuous monitoring systems have been proposed as a promising approach for enhancing drinking water utilities detect and respond efficiently to threats on water distribution systems. Water quality sensors are aimed at revealing contamination intrusions, while hydraulic pressure and flow sensors are utilized for estimating the hydraulic system state. To date optimization models for placing sensors in water distribution systems are targeting separately water quality and hydraulic sensor network goals. Deploying two independent sensor networks within one distribution system is expensive to install and maintain. It might thus be beneficial to consider mutual sensor locations having dual hydraulic and water quality monitoring capabilities (i.e. sensor nodes which collect both hydraulic and water quality data at the same locations). In this study a multi-objective sensor network placement model for conjunctive monitoring of hydraulic and water quality data is developed and demonstrated using the multi-objective non-dominated sorted genetic algorithm NSGA II methodology. Two water distribution systems of increasing complexity are explored showing tradeoffs between hydraulic and water quality sensor location objectives. The proposed method provides a new tool for sensor placements.

Key words | genetic algorithms, multi-objective, optimization, sensors, water distribution systems

INTRODUCTION

In recent years there has been a growing interest in the development of sensor networks to detect threats related to the operation of water distribution systems. A water quality sensor network is the main constituent of an early warning setup against a deliberate contamination intrusion (ASCE 2004) while hydraulic sensors (pressure and flow) are vital both for estimating the hydraulic state of the system and potentially also for detecting leaks or burst events.

Clearly, if all system nodes could be monitored then the maximum level of alertness would have been accomplished. This is obviously not the case thus various methodologies have been proposed for optimizing the placement of hydraulic and water quality monitoring stations.

Literature review

Hydraulic sensors typically monitor pressure or flow rates. This work focuses on the placement of pressure monitoring sensors as they are used more frequently than flow rate sensors as it is cheaper and easier to collect pressure data, and pressure transducers give instantaneous readings whereas most flow meters do not react instantaneously to changes in flow (de Schaetzen et al. 2000). Flow rates are usually measured at all entry points to the network; on main pipes at the entrance into sub-networks (i.e. inflow to bounded demand zones); and/or at the outlet of elevated tanks and pumping stations. Thus, the selection of flow rate measurement locations is straightforward and limited to specific locations.

Quite a few methodologies have been developed for the optimal placement of pressure monitoring sensors (Bush & Uber 1998; Vitkovsky et al. 2003). The common approach is to apply optimization methods (e.g. mixed integer programming, genetic algorithms) to find sensor locations that maximize a sensitivity function which gives a measure of the sensitivity of pressure at selected locations to predefined variations in unknown parameters of interest such as nodal demands and pipe roughness coefficients.
A multi-objective framework for sensor layout design was suggested by Kapelan et al. (2003) in which the optimal locations are determined with the goal of collecting data that will be used later in the calibration of the analyzed water system hydraulic model. The problem is formulated as a two-objective optimization problem involving maximization of the calibrated model accuracy by minimization of the relevant uncertainties, and total costs.


Watson et al. (2004) were the first to introduce a multi-objective formulation to sensor placement. Preis & Ostfeld (2008a) developed a multiobjective optimization evolutionary model for enhancing the response against a deliberate contamination intrusion. Preis & Ostfeld (2008b) used NSGA II (Deb et al. 2000) for optimizing the tradeoffs between the sensor detection likelihood, redundancy; and expected detection time.

Weickgenannt et al. (2010) used also NSGA II to trade-off sensor placements and risk. Dorini et al. (2010) developed single and multi-objective methodologies casted in a framework entitled SLOTS (sensors local optimal transformation system).

Recently, Hart & Murray (2010) reviewed the literature on contamination warning systems (CWSs) with emphasis on optimization-based sensor placement strategies, summarizing the state of the art in this field.

METHODS

So far optimization algorithms for placement of sensors in water networks were implemented separately for water quality or hydraulic purposes. Deploying two independent sensor networks within one distribution system is generally not a cost-effective approach since an in-line sensor tapping point is expensive to install and maintain and is bounded by physical constraints such as limited access to underground pipelines and/or power sources. Thus, there is a need to consider a sensor network with dual hydraulic and water quality monitoring capabilities (i.e. sensor nodes which collect both hydraulic and water quality data at the same tapping points). In this study a multi-objective sensor network placement model for integrated monitoring of hydraulic and water quality parameters using the non-dominated sorted genetic algorithm-NSGA-II method is presented and demonstrated. The two design objectives utilized follow.

Maximize $f_1$ – contaminant events sensor
detection likelihood

Given a sensor network (i.e. number and locations) the detection likelihood (i.e. the probability of detection) is estimated by:

$$f_1 = \frac{1}{S} \sum_{i=1}^{S} d$$

where $d$ is the a binary variable receiving the value of 1 if a contamination event is detected and 0 otherwise, and $S$ the total number of the contamination events sample.

To evaluate the fitness of the sensor network detection likelihood ($f_1$), a contamination matrix (Ostfeld & Salomons 2004) is constructed and simulated using the EPANET (USEPA 2002) water quantity and quality network solver.

Since contamination injections can occur at any node at any time, the theoretical number of possible injection events is high, and grows substantially with system size. To cope with this difficulty and to sample a small yet representative portion of the entire set of pollution events, a heuristic sampling procedure was developed. Firstly, a sample of contaminant injection locations is chosen out of the entire water network nodes. The selection of the contaminant injection locations is based on the network configuration and hydraulics where all the geographical parts of the distribution system are represented by the selected sample of the system's nodes. Secondly, a time interval of 30 min was used to represent all possible injection starting times (i.e. a contaminant injection can start every 30 min). Thereafter, the contaminant injection mass rate and duration are represented by three levels: low, medium, and high of 50, 250, and 500 g/min, and 60, 180, and 360 min, respectively.

Maximize $f_2$ – sensor hydraulic sensitivity to variations
in nodal demands

The hydraulic sensitivity function is computed using a sensitivity matrix (Bush & Uber 1998) of size $n \times m$ where $n$ is the...
number of sensor nodes and $m$ is the number of nodes at
which the demand is altered according to a pre-defined pro-
cedure: the base demand ($D_{\text{base}}$) of each node $j$ of the vector
of nodes $m$ is multiplied in 0.2 intervals by a set of demand
multiplication factors (DMF) ranging from DMF$_1$ = 0.5 to
DMF$_w$ = 2.9, where the actual water consumption at
each node $j$ equals to $D_{jk} = D_{j\text{ (base)}} \times$ DMF$_{k=1}$ ... $w$, and $w$
is the number of demand multiplication factors (i.e. 13
multipliers).

Each element $S_{ij}$ in the sensitivity matrix is defined as:

$$S_{ij} = \sum_{k=1}^{w} \left( \frac{P_i(D_{jk}) - P_i(D_{j\text{ (base)}})}{D_{jk} - D_{j\text{ (base)}}} \right)$$

and the objective function to maximize:

$$f_i = \sum_{i=1}^{n} \sum_{j=1}^{m} S_{ij}$$

where $P_i(D_{jk})$ is the simulated pressure at sensor node $i$ for a
predefined demand $D_{jk}$ at node $j$, and $P_i(D_{j\text{ (base)}})$ the simu-
lated pressure (using EPANET) at sensor node $i$ for the
base demand $D_{j\text{ (base)}}$ at node $j$.

**Multi-objective optimization**

Multiobjective optimization (http://en.wikipedia.org/wiki/
Multi-objective_optimization) deals with finding the vector
of decision variables which satisfies a set of constraints
and optimizes a vector function whose elements represent
the objective functions.

In this study the evolutionary multi-objective NSGA-II
(Deb et al. 2000) algorithm was employed following its gen-
eral implementation success (Deb et al. 2002), and in
particular its accomplishments for multiobjective water dis-
tribution systems management problems (e.g. Prasad &
Park 2004; Prasad et al. 2004; Farmani et al. 2005; Vamv-
akeriou-Lyroudia et al. 2005; Preis & Ostfeld 2008b;
Weickgenannt et al. 2010). The algorithm employs a non-
dominated sorting approach using a selection operator
which creates a mating pool by combining the parent and
offspring populations, and by selecting solutions with
respect to fitness and spread. Generations are populated
starting with the best non-dominated front and succeeding
until the specified population size is reached. If at the final
stage there are more individuals in the non-dominated
front than in the available space, the crowded distance-
based niching strategy is invoked to choose which
individuals of that front will enter into the next population.
The algorithm was implemented in Visual Basic.

**RESULTS AND DISCUSSION**

The methodology is demonstrated on two example applica-
tions of increasing complexity: Example 1 [Anytown,
USA (Walski et al. 1987)] and Example 2 [Network 1 of
the Battle of the Water Sensor Networks (BWSN1, Ostfeld
et al. 2008)]. The maximum number of the NSGA-II gener-
ations for both examples was set to 100 with each
generation having a population of 48 strings. The NSGA-II
algorithm ended if either the upper bound of 100 gener-
ations was attained or if at 10 successive generations no
new non-dominated solutions were found. For Anytown,
USA, the objective functions evaluations were utilized
with a contamination matrix of 500 intrusion events and
sensitivity matrix parameters of $n = 2$ sensors, $m = 16$
demand nodes, and $w = 13$ demand multiplication factors,
where for BWSN1: 1,000 intrusion events, $n = 4$, $m = 110,$
and $w = 13$. Sensors assumed to have a lower level detection
sensitivity of 0.01 mg/L, and instantaneous detection capa-
bilities. The average computational time on a DELL PC
(2.66 GHz, 3.0 GB of RAM) for a single NSGA-II iteration
was about 10 min for Anytown, USA, and about 40 min for
BWSN1.

**Example 1**

Anytown USA (Figure 1) is comprised of 35 pipes, 16 con-
sumer nodes, two elevated storage tanks, one pumping station,
and one well. The pipes, nodes, tanks, pumping station
characteristics, and the 24 h demand flow pattern the

---

**Figure 1 | Example 1 layout and selected solutions.**
system is imposed to, are as in Walski et al. (1987), thus not repeated herein.

Figure 2 describes the Pareto optimal front of the detection likelihood \((f_1)\) versus the hydraulic sensitivity \((f_2)\), emphasizing solution 1.1 which provides the best detection likelihood, and solution 1.2 which provides the highest hydraulic sensitivity. Figure 1 outlines the sensor locations for solutions 1.1 and 1.2, respectively.

Observing Figure 1 it can be argued that the optimal locations found for the hydraulic sensors (i.e. the most ‘sensitive’ locations) differ from the locations one would intuitively select using straightforward engineering judgment. For example, following a simplified but logical approach, node 17 would probably be considered to be a better monitoring location (i.e. ‘more sensitive’) than nodes 11 and 12 as it has a higher demand and it is farther away from the source and tanks. The model results show, however, that node 17 is less sensitive than nodes 11 and 12.

Example 2

The second example application incorporates a more complicated hydraulic regime and is aimed at testing the model performance on a more realistic case study. The system layout (Ostfeld et al. 2008) is shown in Figure 3. It consists of one constant head source, two elevated storage tanks, 170 pipes, 129 nodes (out of which 110 are demand nodes, the rest are internal with zero consumption), and two pumping stations.

Figure 4 presents the Pareto optimal front of the detection likelihood \((f_1)\) versus the hydraulic sensitivity \((f_2)\), emphasizing solution 2.1 which provides the best detection likelihood, and solution 2.2 which provides the highest.
hydraulic sensitivity. Figure 3 shows the sensor locations for solutions 2.1 and 2.2, respectively.

As for example 1, there is a clear tradeoff between \( f_1 \) and \( f_2 \) (Figure 4). It can also be seen (Figure 3) that most \( f_1 \) and \( f_2 \) optimal sensor locations (i.e. nodes 70, 83, 84, 100, 101) are at the densest part of the system near the majority of the water consumption nodes, which for this case might coincide with an engineering intuition as where sensors should be concentrated.

CONCLUSIONS

A multi-objective model for the placement of sensor network with a dual hydraulic and water quality monitoring capability has been presented and demonstrated using two example applications of increasing complexity. Two competing objective functions were considered: (1) maximizing the sensor network contamination events detection likelihood, and (2) maximizing the sensor network hydraulic sensitivity. Pareto fronts were plotted and sensor locations were analyzed. The analysis provided explanatory results, and thus confirmed the methodology ability to provide a multidimensional tool for sensor placement decision making.

It was observed that the optimal solutions for \( f_1 \) (i.e. contaminant detection likelihood) and \( f_2 \) (i.e. hydraulic sensitivity) provide a range of possible sensor locations from which a decision maker would need to select a compromised solution which takes into consideration both objectives or gives priority to one of the objective functions. For example, the water utility might prefer to perform a gradual sensor network deployment, giving more priority to design objective \( f_2 \) at the initial stages (to get a reliable calibration of the hydraulic model of the system) and at subsequent stages, for design objective \( f_1 \).

Ongoing research efforts are conducted for implementation of the methodology on larger water distribution system; for reducing the computational time required for solving the problem; and for incorporating other objective functions such as minimizing costs and maximizing reliability (i.e. quantifying the uncertainty of the sensors measurements thus as to better screen out false alarms). Another issue which needs further consideration is the number and type of contamination events which should be selected such that any other contamination instances which will be imposed on the system will result in similar or less harmful outcomes. A partial solution to this is provided in Perelman & Ostfeld (2010).

ACKNOWLEDGEMENT

This work has been supported by the National Research Foundation of Singapore (NRF) and the Singapore – MIT Alliance for Research and Technology (SMART) through the Center for Environmental Modeling and Sensing.

REFERENCES


First received 8 November 2010; accepted in revised form 11 February 2011