Battle of the Water Sensor Networks (BWSN): A Design Challenge for Engineers and Algorithms

Water Distribution Systems Analysis 2006
Cincinnati, Ohio U.S.A.
August 27-29, 2006
www.eng.uc.edu/wdsa2006/

Detailed Problem Description and Rules

November 17, 2005

Avi Ostfeld
Technion - I.I.T
Dept. Civil & Envr. Engr.
Haifa 32000
ISRAEL
Email: ostfeld@tx.technion.ac.il
PH +972-4-8292782
FAX +972-4-8228898

James G. Uber
University of Cincinnati
Dept. Civil & Envr. Engr.
765 Baldwin Hall; PO Box 210071
Cincinnati, OH 45221-0071
Email: jim.uber@uc.edu
PH 513-556-3643
FAX 513-556-2599

Elad Salomons
OptiWater
6 Amikam Israel St.
Haifa 34385
ISRAEL
Email: selad@optiwater.com
PH +972-54-2002050
FAX +972-15154-2002050

1. Introduction

There has been increased interest in development of sensor networks to detect accidental and deliberate contamination events in water distribution systems and thereby facilitate corrective action or public notification. Optimization models and solution algorithms have been developed using various metrics to locate a limited number of sensors that minimize the impact of such contamination events. These optimization models have made simplifying assumptions about design objectives, network contaminant transport, sensor response, event detection, emergency response, and installation and maintenance costs. Therefore it is not known how well the designs will perform in the field under real-life conditions. Solutions can be difficult to obtain for large networks. Finally, little is known about how these design algorithms compare to the efforts of human designers, and thus what advantages they offer for practical design of sensor networks.

2. Approach and Schedule

The Battle of the Water Sensor Networks (BWSN) will objectively compare the performance of contributed sensor network designs, consisting of a set of sensor locations. Independent research teams and practicing engineers will contribute their designs for two different water distribution networks. Each team will be asked to develop designs according to a precise set of rules (see below) to facilitate design comparisons. These rules specify the design performance metrics, the characteristics of contamination events, and the detection technology used to raise an alarm.

While teams will be encouraged to contribute designs for each network, partial contributions will be accepted. A comparison of the various solutions will be presented at a special session during the 8th Annual Water Distribution Systems Analysis Symposium, to be held in Cincinnati, Ohio, U.S.A., August 27-29, 2006 (www.eng.uc.edu/wdsa2006). Some contributors may be invited to present a synopsis of their solution approach during this special session. If results warrant, a jointly authored journal publication will be prepared to archive the design challenge and contributed solutions. The schedule of events for the *Battle of the Water Sensor Networks* is outlined in Table 1.

Table 1: Schedule of events

Date	**Event**
July 13, 2005	Initial announcement
November 17, 2005	Publication of problem details and competition rules
May 24-31, 2006	Submission of sensor network designs by participants
June 7, 2006	Distribution of contributed sensor designs to participants
August 1, 2006	Distribution of preliminary comparison to participants
August 27-29, 2006	Public presentation of results at WDSA06
October 1, 2006	Development of jointly authored journal manuscript (as warranted)

3. How to Participate

Each participating team must submit an on-line abstract for the WDSA2006 conference (www.eng.uc.edu/wdsa2006) that discusses briefly the proposed design approach (e.g., trial and error with simulation, integer programming, heuristics, etc.). When submitting the abstract, ***the topic area must be identified as "Battle of the Water Sensor Networks"*** – this will identify your team of authors as a participant in the competition. Each team must summarize their final design results in a brief conference paper; these must be uploaded to the same web site no earlier than May 24 and no later than May 31, 2006, to assure independence of contributed solutions. All conforming designs will be included in the public presentation of summary results at the conference and will be published as part of the conference proceedings. Submission of an abstract and paper for the *Battle of the Water Sensor Networks* does not, however, imply or require an oral presentation at the WDSA2006 conference.

Submitted papers describing the final designs from each team should be brief and to the point. It is not necessary to describe the competition, as that will be included in the summary comparison paper. Each paper should briefly state that it is part of the *Battle of the Water Sensor Networks* design competition, and reference the summary comparison paper also published as part of the same proceedings: A. Ostfeld, J. Uber, and E. Salomons "Battle of the Water Sensor Networks: A Summary of Contributed Designs," *Proceedings of the Conference on Water Distribution Systems Analysis*, Cincinnati, Ohio, U.S.A., 2006.

To allow efficient and fair assessment of contributed designs, papers submitted to the *Battle of the Water Sensor Networks* are asked to include the following sections: Abstract; Introduction (brief); Methodology (whether qualitative or quantitative); Summary of sensor designs for networks 1 and 2, and cases A, B, C, and D (see below); Discussion of Results; Conclusions; and References. The Discussion of Results section should include method-specific information about designs and design effort (e.g., person-hours, computer type and execution time/memory requirements, contamination scenarios used for design purposes, time step used to compute the objective functions, or modifications made to the original EPANET inp file). Designs submitted with incomplete information may be excluded from the comparison.

Once an abstract has been submitted to the *Battle of the Water Sensor Networks* topic area as described above, forward a copy of the abstract to both Avi Ostfeld and James Uber (see above contact information). A package of problem data and network models will be forwarded to all authors submitting abstracts that describe a conscientious design approach. Authors are asked to not forward the problem data and network models to non-participants, so that the organizers can maintain a reliable participant list.

4. Contributed Designs

Contributed sensor network designs will be evaluated using four quantitative design objectives: expected time of detection (Z_1), expected population affected prior to detection (Z_2), expected contaminated water demand prior to detection (Z_3), and expected likelihood of detection (Z_4). Calculation procedures for each of the four design objectives are described below. Teams are free to develop their designs however they wish. Partial designs are also accepted. Teams are not required to calculate and submit these objective measures for their various designs; the organizers will independently perform these calculations using identical procedures for every submitted design.

Expected Time of Detection (Z_1)

For a particular contamination scenario, the time of detection by a sensor is the elapsed time from the start of the contamination event to the first presence of a non-zero contaminant concentration at that sensor. Denote this time of detection t_i where the subscript *i* refers to the *i*th sensor location. The time of detection for the sensor *network* and this particular contamination event, t_d, is the minimum among all sensors present in the design, $t_d = \min_i t_i$. The objective function to be minimized is the expected value computed over the assumed probability distribution of contamination events,

$$Z_1 = E(t_d)\ ,$$

where $E()$ denotes the mathematical expectation, and will be approximated by Monte-Carlo simulation for purposes of comparing contributed designs. The assumed probability distribution of contamination events is discussed below in conjunction with other design assumptions.

Expected Population Affected Prior to Detection (Z_2)

For a particular contamination scenario, the population affected is a function of the ingested contaminant mass. The ingested contaminant mass in turn depends on the time of detection for the sensor network, as described above; a key assumption is that no mass is ingested after detection. For a particular contamination scenario, the mass ingested – prior to detection – by any individual at network node *i* is calculated,

$$M_i = \phi \Delta t \sum_{k=1}^{N} c_{ik} \rho_{ik}\ ,$$

where ϕ is the mean volumetric ingestion rate (Liters/day), Δt is the evaluation time step (days), c_{ik} is the contaminant concentration for node *i* and time step *k* (mg/Liter), ρ_{ik} is a "dose rate multiplier" for node *i* and time step *k* (unitless), and *N* is the number of evaluation time steps prior to detection, *i.e.* the largest integer such that $N\Delta t \le t_d$. The series ρ_{ik}, $k=1,\ldots$ has a mean of 1 (so ϕ is truly the mean volumetric ingestion rate) and is intended to model the variation in ingestion rate throughout the day. We assume the ingestion rate varies with the water demand rate at the respective node, thus $\rho_{ik} = q_{ik} / \overline{q}_i$, where q_{ik} is the water demand for time step k and node i, and $\overline{q}_i$ is the average water demand at node i.

A typical dose-response model is used to express the probability that any person ingesting mass M_i will be affected (*i.e.* become infected or symptomatic),

$$R_i = \Phi[\beta \log_{10}((M_i / W) / D_{50})]$$

where R_i is the probability [0, 1] that a person who ingests contaminant mass M_i will become infected or symptomatic, β is the so-called Probit slope parameter (unitless), W is the assumed body weight (kg), D_{50} is the dose that would result in a 0.5 probability of becoming infected or symptomatic (mg/kg), and Φ is the Standard Normal Cumulative Distribution Function.

The population affected for a particular contamination scenario is calculated, $P_a = \sum_{i=1}^{m} R_i P_i$, where P_i is the population assigned to node i and m is the total number of nodes. The objective function to be minimized is the expected value of P_a computed over the assumed probability distribution of contamination events,

$$Z_2 = E(P_a),$$

where $E()$ denotes the mathematical expectation, and will be approximated by Monte-Carlo simulation for purposes of comparing contributed designs. The assumed probability distribution of contamination events is discussed below in conjunction with other design assumptions.

<u>*Expected Demand of Contaminated Water Prior to Detection (Z_3)*</u>
Z_3 is the expected volume of consumed contaminated water prior to detection,

$$Z_3 = E\ (V_d)$$

where V_d denotes the total volumetric water demand that exceeds a predefined hazard concentration, and $E()$ is the mathematical expectation, which will be approximated by Monte-Carlo simulation for purposes of comparing contributed designs. As for the expected population affected, a key assumption is that no water is delivered after detection. The assumed probability distribution of contamination events is discussed below in conjunction with other design assumptions. Z_3 (as Z_2 and Z_1) is to be minimized.

<u>*Detection likelihood*</u> (Z_4)
Given a sensor network design (i.e., number and locations) the detection likelihood (Z_4 – the probability of detection) is defined:

$$Z_4 = \frac{1}{N} \sum_{i=1}^{N} d_i$$

where: $d_i = 1$ if contamination scenario i is detected, and zero otherwise, and N is the number of the total contamination scenarios considered. Z_4 is to be maximized.

5. Design Assumptions and Cases

A summary of the design cases is given in Table 2 below. EPANET version 2.00.10 will be the standard model used to simulate transport. For each of two networks (see next section), participants are asked to provide designs for locating 5 sensors and 20 sensors for a base case (A) and three derivative cases (B, C, D). To avoid confusion, participants are asked to use the following notation when identifying particular sensor network designs in their written submission: N[1-2][A-D][5,20], where the first bracketed list represents the network, followed by the design case, and finally followed by the sensor network size (5 or 20 sensors). For example, the 5 sensor base case (A) design for network 1 would be denoted N1A5. Each team is asked to provide a summary table in their report, with one row per design case, listing the case identifier (e.g., N1A5), and the EPANET node IDs selected for that case.

Base Case (A)

a. All quantities affecting network model water quality predictions are assumed to be known and deterministic unless explicitly stated otherwise. Sensor network designs will be challenged by an ensemble of contamination scenarios sampled from a statistical distribution; the probability distribution of contamination events is described as follows. Contaminant intrusions occur at network nodes, with injection flow rate of 125 Liter/hr, contaminant concentration of 230,000 mg/L, and injection duration of 2 hrs. The contaminant is stable after injection in finished drinking water. Each contamination scenario involves a single injection location, which may occur at any network node and begin at any time with equal probability. For purposes of design evaluation, contaminant concentrations will be predicted with a 5 minute time step. The sizes of the Monte Carlo ensembles used to approximate each design objective will be reported, and sensitivity analysis will be used to assess stability of the design objective estimates. As outlined above an independent calculation using a random set of Monte Carlo simulations will be used to assess the design of each team, using identical procedures for every submitted design.

b. For purposes of calculating the expected population affected prior to detection (Z_2): ϕ = 2 Liters/day, β = 0.34, D_{50} = 41 mg/kg, W = 70 kg. For purposes of estimating node population, the total per capita water consumption rate is assumed to be 300 Liters/day.

c. For purposes of calculating the expected demand of contaminated water prior to detection (Z_3) the hazard concentration threshold is C = 0.3 (mg/Liter).

d. Sensors instantly detect any non-zero contaminant concentration and action is taken to eliminate further exposure without delay.

Derivative Case (B)

Identical to base case (A) except that the injection duration is increased to 10 hrs.

Derivative Case (C)

Identical to base case (A) except that the response delay is 3 hour, *i.e.*, it takes 3 hr after detection for emergency response to limit contaminant exposure.

Derivative Case (D)

Identical to base case (A) except that any contamination scenario involves two injection locations, which may occur at any two distinct nodes with equal probability. The contamination scenario may begin at any time with equal probability, but both injections are synchronized to begin at the same time.

6. Network Models

Network 1 and 2 model descriptions will be provided in EPANET input format. Network 1 is a small example (approximately 100 nodes) while network 2 is a large example (approximately 12000 nodes).

Table 2: Summary of Design Assumptions and Cases

Base case (A)			Derivative case		
Component	**Characteristics**	**Description**	**B**	**C**	**D**
System	Possible injection locations	Nodes			
	Existing sensors	None			
	Demands	Deterministic			
Injection	Nodes injection probabilities	Even			
	Number of injection events	One (i.e., a single random attack)			Two
	Duration	2 (hr)	10 (hr)		
	Flow	125 (Liter/hr)			
	Constituent concentration	230000 (mg/Liter)			
	Constituent type	Conservative (i.e., no decay, no interaction)			
Sensors	Detection delay	Real time (i.e., zero delay)		3 (hr)	
	Detection sensitivity	Ideal (i.e., above zero)			
Design	Number of sensors	5 ; 20			

Comment: Contaminant concentrations are assumed to be predicted with a five minute time step.

7. Additional Information

Contact Avi Ostfeld (ostfeld@techunix.technion.ac.il) or Jim Uber (jim.uber@uc.edu).