Ehsani, N. and Afshar, A. (2011) Optimization of Contaminant Sensor Placement in Water Distribution Networks: Multi-Objective Approach. Water Distribution Systems Analysis 2010: pp. 338-346.
http://dx.doi.org/10.1061/41203(425)32

OPTIMIZATION OF CONTAMINANT SENSOR PLACEMENT IN WATER DISTRIBUTION NETWORKS; MULTI-OBJECTIVE APPROACH

Nima Ehsani [1], Abbas Afshar [2]

[1] Civil Engineering Department, Iran University of Science & Technology
Tehran, Tehran, Iran
nima_ehsani@civileng.iust.ac.ir

[2] Civil Engineering Department, Iran University of Science & Technology

Tehran, Tehran, Iran
a_afshar@iust.ac.ir

Abstract

The most complex issues threatening water distribution networks are deliberate injection of chemical or biological contaminants into distribution water networks. The main problem in using contaminant sensors in drinking water distribution systems is the high purchase, installation and maintenance prices for this kind of sensors. In addition, we need to provide society with maximum health protection; contamination must be detected as fast as possible and all possible threats need to be covered. In this study we took into consideration two main objectives: (1) Maximizing Detection Likelihood and (2) Minimizing Expected Time to Detection. A new formulation that combines these two objectives into a single formulation is introduced and then we use Genetic Algorithm for solving the problem. As second approach, NSGA II is used for solving the sensor problem considering both objectives separately and at the same time. An example network is used to demonstrate the functionality of the proposed methods.

Keywords

Sensor, Contaminant, Water Networks, Genetic Algorithm, NSGA II, Optimization

1. INTRODUCTION

Among different issues threatening water distribution networks, the most complex ones are deliberate injection of chemical or biological contaminants into distribution water networks as there is uncertainty about contaminant type, starting location of contamination in network, starting and ending time of contamination event, and results of attacks. In recent years researches resulted in significant improvements in water quality sensors. These sensors work better and faster in contrast with older models and they are more reliable. The main problem in using these new sensors is the high purchasing, installation, and maintenance expenses. In addition to financial limitations there are other design requirements that must be taken into account. Maximum water security must be provided to the public. Contamination events must be detected as fast as possible and all possible threats need to be covered. We therefore need algorithmic methods to find the optimum layout of sensors network in water distribution systems.

For this purpose different models are presented and each tries to cover the mentioned problems. Lee and

Deininger (1992) developed a method based on Demand Coverage concepts. Their solution was established on the base that water quality reduces as time passes and water gets farther from its source. By using a water fraction matrix they changed the problem to an Integer Programming problem and specified locations for monitoring stations in network in a way that a few sensors could control the water quality of a large part of network. Kumar et al (1997) also used Coverage Matrix which was developed by Lee and Deininger by offering some changes in the methodology.

Kessler et al (1998) introduced a new term in their studies, Level of Service. Level of Service is the maximum volume of water which is allowed to pass through a specific node according to detection time before water contamination is detected by a sensor. After hydraulic simulation of water flow in the network, a directed graph will be constructed which will be used instead of the real network in solving the problem. All edges are weighted according to average flow velocity in the pipe. Using this graph and knowing length of the pipes, minimum traveling time between each two nodes will be calculated. By using this information, effects of an attack may be determined.

Berry et al (2003) presented a model for optimizing the placement of sensors in municipal water networks in which they considered a sensor placement formulation for which optimal sensor configurations minimize the expected fraction of the population that is at risk for an attack. They weighted each node by the number of people potentially consuming water at that point and used a mixed-integer programming model to exactly solve it. In another study, Berry et al (2004) introduced a new Integer Programming based model for solving the problem. Their model requires a risk profile, which is a probability distribution that weights the likelihood of an attack at a specific node at a specific time. This model minimizes the number of people exposed to a dangerous level of contamination.

Al-Zahrani and Moied (2003) developed the Lee and Deininger's model. They solved the problem under multiple flow scenarios considering that flow demands in a water distribution network vary during the day. They solved the optimization problem using Genetic Algorithms.

Ostfeld and Salomons (2004) presented a methodology for finding the optimal layout of an early warning detection system. They considered extended period unsteady hydraulics and water quality conditions for solving the optimization problem. This model is capable of solving the sensor problem under multiple injections (up to three injections) per pollution event. A given defensive level of service to the public was defined as a maximum volume of contaminated water exposure at a concentration above a minimum level. They used Genetic Algorithms for solving the optimization problem.

Propato (2006) introduced Mixed-Integer Linear Models for Sensor Location Design. Afshar et al (2006) developed a model using Ant Colony Optimization (ACO) algorithm for the first time. Their model was based on Al-Zahrani and Moied (2003) and Lee and Deininger (1992) researches results. In The Battle of the Water Sensor Networks (BWSN) (Ostfeld et al, 2008) a comparison was made between different models efficiency and reliability. Aral Et al (2010) used an optimization algorithm based on a simulation-optimization and a single-objective function approach which incorporates multiple factors used in the design of the system. They also introduced a reliability constraint concept into the optimization model such that the minimum number of sensors and their optimal placement can be identified in order to satisfy a pre specified reliability criterion for the network.

In this study we took into consideration two main objectives: (1) Maximizing Detection Likelihood and (2) Minimizing Expected Time to Detection. We assume that when a contamination event is detected, a general warning system will raise an alarm and water consumption will stop immediately. Also we assume that sensors are ideal and do not fail to detect a contamination event when polluted water is passing by. With regards to the mentioned objectives in solving the sensor problem, the probability of detection of a contamination event will be maximized and general warning will ban water consumption as fast as possible so number of casualties and amount of damages following a probable attack will be minimized.

In this study we introduce a new formulation which combines Maximizing Detection Likelihood and Minimizing Expected Time to Detection into one formulation and then by using genetic algorithms we solve the optimization problem. In the second approach we use NSGA II (Deb et al, 2002) considering both objectives separately and at the same time in the form of a multi objective optimization problem so

Ehsani, N. and Afshar, A. (2011) Optimization of Contaminant Sensor Placement in Water Distribution Networks: Multi-Objective Approach. Water Distribution Systems Analysis 2010: pp. 338-346.
http://dx.doi.org/10.1061/41203(425)32

we can have an optimum Pareto solution instead of a single solution.

2. PROPOSED OPTIMIZATION TECHNIQUE

In this study we use the following definitions for the proposed design objectives (Ostfeld et al, 2008):

2.1 Detection Likelihood; Z_1:

In summary, we can define Detection Likelihood in a sensor network as the probability of detection of a contamination event in water distribution system by at least one sensor. In other words, with regards to this objective, best solution may be obtained when proposed sensor layout can detect every attack on network.

Assuming a specific water distribution system and a defined layout for sensor network, the detection likelihood is estimated by:

$$Z_1 = \frac{1}{A}\sum_{r}^{A} d_r \qquad (1)$$

Where d_r equals 1 if contamination event r is detected by sensors network, and zero if contamination event r is not detected. A is the total number of contamination events considered. Z_1 should be maximized.

2.2 Expected Time to Detection; Z_2:

For a particular contamination scenario, the detection time (t_i) for a specific sensor placed in location i, is the time interval between the start of the contamination event and the first detection of contaminant presence in water distribution system by that sensor. The detection time for a sensor network (T_d) is estimated by:

$$\mathbf{T_d} = \begin{cases} 0 & \text{contamination event is not detected} \\ \min t_i & \text{contamination event is detected} \end{cases} \qquad (2)$$

As previously mentioned, contamination detection means to ensure that at least one of the sensors will be activated while contaminated water is passing through the water network. Thus for detection of a contamination event, firstly at least one sensor is needed to be present in contaminated water path through network and secondly the concentration of contaminant must be above the minimum concentration that sensors are able to detect. The objective function to be minimized which is the Expected Time to Detection is estimated by:

$$Z_2 = \frac{1}{A_d} \sum_{i}^{A_d} T_d \qquad (3)$$

A_d is the number of contamination events which are detected by the sensor network. In this formulation we ignore the effects of undetected contamination events.

http://dx.doi.org/10.1061/41203(425)32

2.3 Detection Likelihood + Expected Time to Detection; Z_3:

In this section we developed a new formulation combining the two objectives mentioned previously into a single objective. General form of this function is similar to Z_2 (Expected Time to Detection). The only difference is a penalty value which we call t_{max}. In Z_2 function whenever a contamination event is not detected, T_d equals zero. In other word the effects of undetected contamination events are ignored so by considering t_{max} instead of zero, effects of undetected contamination events are considered. The detection time for a sensor network (T_d) is estimated by:

$$T_d = \begin{cases} t_{max} & \text{contamination event is not detected} \\ \min t_i & \text{contamination event is detected} \end{cases} \qquad (4)$$

The value of t_{max} is considered equal to hydraulic and quality simulation time. The objective function to be minimized here is estimated by:

$$Z_3 = \frac{1}{A} \sum_{i}^{A} T_d \qquad (5)$$

A is the total number of contamination events (detected plus undetected). By using t_{max} , while minimizing Z_3 , optimization algorithm will arrange sensors in a way that minimizes penalty value. In other words it minimizes the number of undetected contamination events and maximizes the detection likelihood.
Moreover a constraint is considered in this problem which is total number of sensors equals *S* and *S* is a constant value. In this way placing more than one sensor in a single node is impossible.
Application of these methods is illustrated by solving the problem for an example water distribution network.

3. EXAMPLE NETWORK

The methodology is presented using EPANET 2.0 example 3 shown in Figure 1. The reason this network is chosen is that it demonstrates the medium complexity for solving the optimization of contaminant sensor placement in urban water distribution networks problem. EPANET 2.0 example 3 consists of 97 nodes. Each node is a possible location for sensor placement and also an attack is probable to occur in every node. In this study 1164 attacks (*A*=1164) are modeled. Each attack consists of a single contaminant injection in only one node at the mass rate of 200 gr/min and for a 2 hours interval. For each network's node, 12 contamination events are modeled. Injection starting time is at 0, 2, 4, …, 20, 22. For this example the total number of sensors, *S*, is equal to 10.
In this study EPANET Programmer's Toolkit Version 2.00.07 (January 2001) is used for analyzing the water quality behavior of water distribution system. All analyses are performed for the system's extended period simulation timeframe, 24 hours. As a result t_{max} in Z_3 equals 24 hours or 1440 minutes.
After analyzing the water quality behavior of network, it is time for constructing the contamination matrix using the obtained results. For this purpose we used contamination matrix introduced by Ostfeld and Salomons (2004). We assumed that sensors are perfect and able to detect any contamination if its concentration in sensor's location is above 1 *mg/Lit*. A contamination matrix is an $N \times A$ matrix of 0–1 coefficients, where *N* is the number of nodes and *A* is the number of contamination events considered. The rows of the matrix list all contamination events, while the columns list all nodes. As mentioned before, in this example $N = 97$ and $A = 1164$.

Ehsani, N. and Afshar, A. (2011) Optimization of Contaminant Sensor Placement in Water Distribution Networks: Multi-Objective Approach. Water Distribution Systems Analysis 2010: pp. 338-346.
http://dx.doi.org/10.1061/41203(425)32

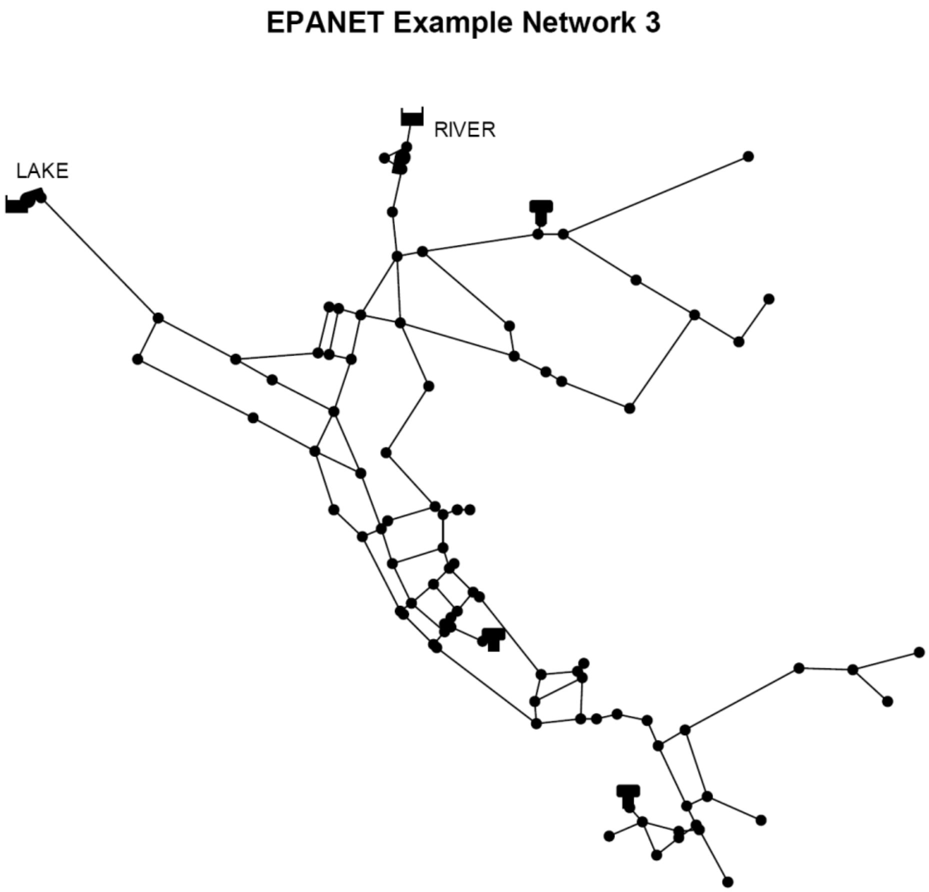

Figure 1. EPANET 2.0 example 3

In each attack during simulation time (24 hours), if contaminant concentration in a node rises above the minimum level of 1 mg/L, that node is considered contaminated. It means that the correspondent element of that node and attack in contamination matrix equals 1. Otherwise the node is considered uncontaminated and the correspondent element in contamination matrix equals 0. Table 1 illustrates **a part** of contamination matrix which is used for solving the example problem.

Table 1. Part of the contamination matrix which is used for solving the example problem.

detection matrix	**node 1**	**node 2**	**node 3**	**node 4**	**node 5**	**node 6**	**node 7**	**node 8**	**node 9**	**node 10**	**node 11**
attack 1	0	0	0	0	1	0	0	0	0	0	0
attack 2	0	0	0	0	0	1	0	0	0	0	0
attack 3	0	0	1	0	0	0	0	0	0	0	0
attack 4	0	0	0	0	0	0	0	0	0	0	0
attack 5	0	0	0	0	0	0	0	0	0	0	0
attack 6	1	0	0	0	0	0	0	0	0	0	0
attack 7	0	1	1	0	0	0	1	0	1	0	0
attack 8	0	0	1	0	0	0	1	0	1	0	0
attack 9	1	0	0	1	1	0	1	0	1	0	0
attack 10	1	0	0	0	1	0	1	0	1	1	1
attack 11	1	0	0	0	0	1	1	0	1	1	1

Ehsani, N. and Afshar, A. (2011) Optimization of Contaminant Sensor Placement in Water Distribution Networks: Multi-Objective Approach. Water Distribution Systems Analysis 2010: pp. 338-346.
http://dx.doi.org/10.1061/41203(425)32

Using contamination matrix for a specific attack scenario and sensor layout, if at least one sensor is presented at a contaminated node, that event will be detected. Otherwise the contamination event will remain undetected. Using this information and Z_1 formula, detection likelihood for sensor layout can be determined.

If the contamination event is detected, by using EPANET's water quality analysis results and sensor network layout, the detection time (t_i) for a specific sensor placed in location *i* may be determined and as a result min t_i and T_d may be determined.

As mentioned before, if using Z_2, for an undetected contamination event T_d equals zero and while using Z_3, for an undetected contamination event T_d equals t_{max} which equals 1140 minutes in this example. Table 2 illustrates **a part** of the time matrix which is used here. This table shows how many minutes after starting of an attack, contaminant concentrations in a specific node reaches 1 *mg/L*.

Table 2. Part of the time matrix which is used for solving the example problem (minutes).

detection matrix	**node 1**	**node 2**	**node 3**	**node 4**	**node 5**	**node 6**	**node 7**	**node 8**	**node 9**	**node 10**	**node 11**
attack 1	0	0	0	0	85	0	0	0	0	0	0
attack 2	0	0	0	0	0	305	0	0	0	0	0
attack 3	0	0	110	0	0	0	0	0	0	0	0
attack 4	0	0	0	0	0	0	0	0	0	0	0
attack 5	0	0	0	0	0	0	0	0	0	0	0
attack 6	70	0	0	0	0	0	0	0	0	0	0
attack 7	0	10	1350	0	0	0	295	0	290	0	0
attack 8	0	0	10	0	0	0	175	0	170	0	0
attack 9	320	0	0	10	970	0	225	0	220	0	0
attack 10	235	0	0	0	10	0	240	0	235	170	170
attack 11	415	0	0	0	0	10	550	0	540	335	345

For solving the problem using Genetic Algorithms we need to determine chromosomes shape. It is possible to represent each chromosome as a bit string. In this method a chromosome consists of *N* (number of nodes=97) genes each can take a value of 0 or 1. Whenever a gene takes 1, it means that a sensor is placed in the corresponding node. The total number of genes that take one must be equal to *S* (total number of sensors=10).

The other approach involves using arrays of integer numbers instead of bit strings to represent chromosomes. In this method a chromosome consists of *S* (total number of sensors=10) genes each can take a value between 1 and *N* (number of nodes=97). Whatever values a gene takes means that a sensor is placed in the corresponding node. As total number of sensors equals *S* (10), genes cannot take similar values. In this study we used the second approach for representing a chromosome. Total population for both multi objective and single objective approaches is equal to 500 and stopping criteria is reaching 100^{th} generation.

As mentioned before, Z_3 function combines Z_1 and Z_2 functions into one formulation. Considering the numerical value of t_{max} instead of zero for T_d for undetected events, acts like a penalty function. This penalty value leads the algorithm in a manner such that for minimizing Z_3 arranges sensors in a way that sensor network detects maximum number of contamination events. In other words detection likelihood will be maximized. Table 3 shows 5 superior solutions for Z_3 obtained from solving the problem using single objective Genetic Algorithm. Table 4 shows values of Z_1 and Z_2 functions for these 5 solutions.

Ehsani, N. and Afshar, A. (2011) Optimization of Contaminant Sensor Placement in Water Distribution Networks: Multi-Objective Approach. Water Distribution Systems Analysis 2010: pp. 338-346.
http://dx.doi.org/10.1061/41203(425)32

Table 3. 5 superior solutions for Z_3.

GA Solution	Sensor 1	Sensor 2	Sensor 3	Sensor 4	Sensor 5	Sensor 6	Sensor 7	Sensor 8	Sensor 9	Sensor 10	Z3 (Minute)
Solution 1	9	64	90	7	86	37	81	33	16	54	355
Solution 2	9	64	89	7	86	37	80	34	16	55	356
Solution 3	9	64	90	7	86	38	80	33	17	54	362
Solution 4	9	64	89	7	86	38	81	33	17	54	362
Solution 5	9	64	90	7	86	38	81	33	17	54	362

Table 4. Values of Z_1 and Z_2 functions for 10 superior solutions for Z_3.

GA Solution	Sensor 1	Sensor 2	Sensor 3	Sensor 4	Sensor 5	Sensor 6	Sensor 7	Sensor 8	Sensor 9	Sensor 10	Z_1 (%)	Z_2 (Minute)
Solution 1	9	64	90	7	86	37	81	33	16	54	81.01	100
Solution 2	9	64	89	7	86	37	80	34	16	55	81.27	106
Solution 3	9	64	90	7	86	38	80	33	17	54	80.50	101
Solution 4	9	64	89	7	86	38	81	33	17	54	80.41	99
Solution 5	9	64	90	7	86	38	81	33	17	54	80.41	100

Figure 2 illustrates optimum Pareto solutions obtained by considering Z_1 and Z_2 functions using NSGA II and the best solution for Z_3 function (Table 4).

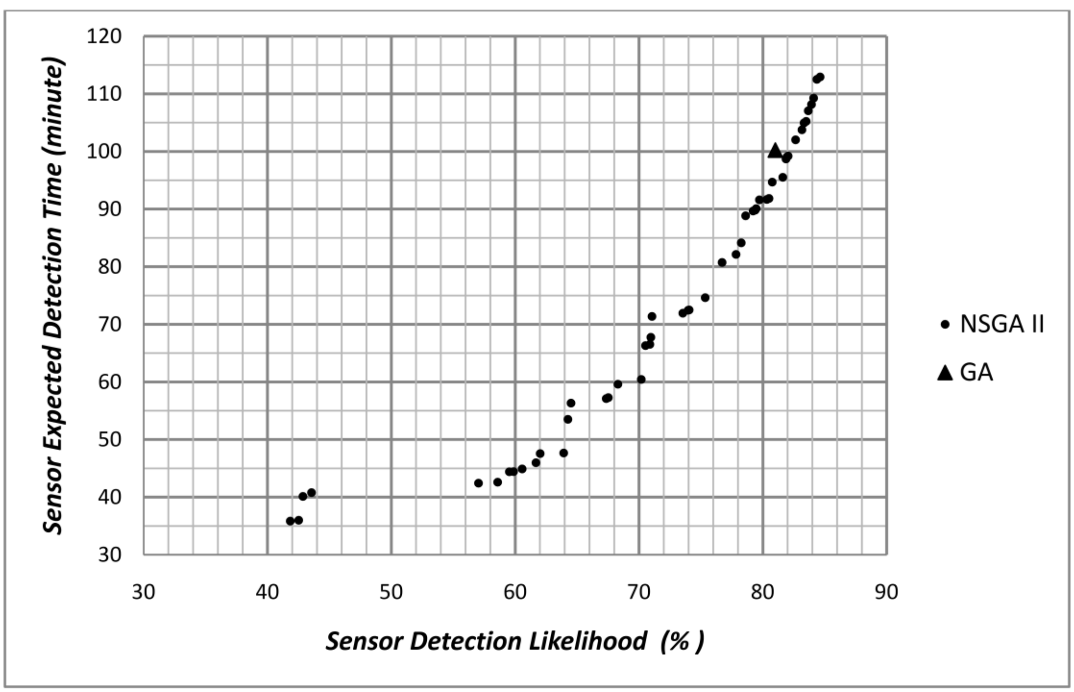

Figure 2. Results obtained from both methods

Ehsani, N. and Afshar, A. (2011) Optimization of Contaminant Sensor Placement in Water Distribution Networks: Multi-Objective Approach. Water Distribution Systems Analysis 2010: pp. 338-346.
http://dx.doi.org/10.1061/41203(425)32

4. CLUSION

In this study we used two different approaches for solving Optimization of Contaminant Sensor Placement in Urban Water Distribution Networks problem. Considering different objectives that may be used for solving the problem, we chose (1) Detection Likelihood and (2) Expected Time to Detection as the most important parameters in solving the problem. In the first approach, using a new formulation, we combined two mentioned objectives into one formulation and then we solved the problem as a single objective optimization problem using Genetic Algorithms. In the second approach, we used NSGA II for solving the sensor problem considering both objectives separately and at the same time in the form of a multi objective optimization problem. Both approaches showed a good performance in finding the solutions. First approach as a single objective method provided us with one solution for sensor layout. However it is faster in contrast with the second approach. Moreover the multi objective approach is computationally more expensive but it provides an optimum Pareto front that gives the possibility of choosing the proper answer from a set of non-dominated solutions.

References

Afshar. A, R. Maknoon, A. Afshar. (2006). "Optimum Layout for Water Quality Stations through Ant Colony Algorithm", Water and Wastewater, vol.59, pp.2-11, 2006.

AL-Zahrani M A, Moied K. "Optimizing water quality monitoring stations using genetic algorithms." The Arabian Journal for Science and Engineering, 2003, 28(1B): 57-75.

Aral, M. M., Guan, J. and Maslia, M. L. "Optimal Design of Sensor Placement in Water Distribution Networks,"ASCE Journal of Water Resources Planning and Management. Volume 136, Issue 1, pp. 5-18 (January/February 2010).

Berry, L. Fleischer, W.E. Hart, and C.A. Phillips. "Sensor placement in municipal water networks". In Proceedings of the World Water and Environmental Resources Conference, 2003.

Berry, W.E. Hart, C.A. Phillips, and J. Uber. "A general integer-programming-based framework for sensor placement in municipal water networks". In Proceedings of the World Water and Environmental Resources Conference, 2004.

Deb K, Pratap A, Agarwal S, Meyarivan T. A fast and elitist multiobjective genetic algorithm: NSGA-II. IEEE Trans Evol Comput 2002; 6(2):182–97.

Kesseler, Ostfeld, and G. Sineri, "Detecting Accidental Contaminations in Municipal Water Networks", J. Water Resour. Planning and Management, ASCE, 124(4) (1998), pp. 192–198.

Kumar, M.L., Kansal, and G. Arora, "Identification of Monitoring Stations in Water Distribution System", J. Envir. Engrg., ASCE, 123 (8) (1997). pp. 746–752.

Lee, and R.A. Deininger, "Optimal Iocations of Monitoring Stations in Water Distribution system", J. Envir. Engrg., ASCE, 118(1) (1992), pp. 4–16.

Ostfeld and E. Salomons. Optimal layout of early warning detection stations for water distribution systems security. Journal of Water Resources Planning and Management, 130(5):377–385, 2004.

Ostfeld, A., Uber, J.G., Salomons, E., Berry, J.W., Hart, and …. (2008), "The Battle of the Water Sensor Networks (BWSN): A Design Challenge for Engineers and Algorithms", ASCE Journal of Water Resources Planning and Management, 134(6), 556-568.

Propato Marco. "Contamination Warning in Water Networks: General Mixed-Integer Linear Models for Sensor Location Design." Journal of water resources planning and management © ASCE / July/August 2006.