Jiří Jaromír Klemeš, Petar Sabev Varbanov and Peng Yen Liew (Editors)
Proceedings of the 24th European Symposium on Computer Aided Process Engineering – ESCAPE 24
June 15-18, 2014, Budapest, Hungary. Copyright © 2014 Elsevier B.V. All rights reserved.

Optimal Sensor Placement for Contamination Detection and Identification in Water Distribution Networks

Venkata Reddy Palleti*, Shankar Narasimhan, Raghunathan Rengaswamy

Department of Chemical Engineering, Indian Institute of Technology, Madras, Chennai-600036, India
venki.pec@gmail.com

Abstract

Water Distribution Networks (WDN) is often exposed to either intentional or accidental contamination. In order to protect against such intrusions, an effective and efficient online monitoring system through sensors is needed. Detection of contaminants in WDN is challenging and it is not possible to place sensors at each and every potential point of intrusion, due to cost and maintenance reasons. Instead, as few sensors as possible, should be located optimally such that intrusions can be detected quickly. This is known as sensor network design problem for intrusion detection in WDNs. Several optimization models and algorithms have been proposed for intrusion detection in a WDN. In this study, we design sensor networks which satisfy the two important properties of observability and identifiability. Observability denotes the ability of the sensor network to detect the occurrence of the intrusion, whereas identifiability refers to the ability to unambiguously deduce the point (or source) of intrusion from the set of sensors affected. A hydraulic analysis of the network is first carried out for a given loading condition to determine the flow directions. The concept of a directed path is then used to construct a bipartite graph, and map the sensor network design problem to that of a minimum vertex set cover problem. Algorithms based on greedy heuristics are used to solve the set covering problem and obtain the corresponding sensor network. The proposed method is illustrated using a fairly large scale urban WDN.

Keywords: Water distribution network, Sensor network, Observability, Identifiability

1. Introduction

Water Distribution Network (WDN) is a vital part of any city. Water distribution networks consist of a network of pipes, reservoirs, pumps, valves, storage tanks and sumps. Due to its complex structure, WDNs are inherently vulnerable to either intentional or accidental contamination. Accidental contamination occurs due to the intrusion of contaminant from the ground or sewage lines through cracks in the pipelines. Accidental intrusion of chemical or biological contaminants can cause an outbreak of health related problems, which can reach epidemic proportions. Intentional contamination can be due to acts of terrorism or mischief which can pose a far greater threat to human life. A monitoring system through online sensors is needed to protect against such intrusions. Locating sensors at each and every point of intrusion is the trivial solution for protecting WDN. Obviously, such a solution is not economically feasible and maintenance of these sensors may also be difficult. Hence, it is necessary to optimally locate a limited number of sensors.

Several optimization models and algorithms are developed for identifying the strategic location of sensors in WDN. Lee and Deininger (1992) were the first to address the optimal location of monitoring stations in a water distribution system. An integer programming model was developed to maximize sensor coverage and its model is based on the steady state simulation and network connectivity. Kumar et al. (1997) and Kessler et al. (1998) improved the method proposed by Lee and Deininger (1992). Ostfeld and Solomons (2004) extended the same model by constructing the randomized pollution matrix and solved using a genetic algorithm. Watson et al. (2004) were the first to introduce a multi-objective formulation to the optimal sensor placement. The Battle of Water Sensor Network (BWSN) challenge (Ostfeld et al., 2008) highlighted the performance of the sensor network design by considering multiple objectives. Mustafa et al. (2010) developed a single-objective optimization model that incorporates the four criteria adopted in BWSN and solved it using a progressive genetic algorithm (PGA). Perelman and Ostfeld (2013) proposed a methodology using Bayesian network statistics to estimate likelihood of the injection location of a contaminant and its propagation in the water distribution system.

2. Problem Description

Despite several optimization models and algorithms, some of the basic questions are yet to be addressed for contamination detection for a given water distribution network. In this study, the two basic questions addressed are (1) what is the minimum number of sensors required to detect an intrusion for a given set of potential intrusion locations in a water distribution network? (Observability problem), and (2) what is the minimum number of sensors required to detect and identify the source of intrusion for a given set of potential intrusion locations in a water distribution network? (Identifiability or resolution problem).

The above questions have been answered in the context of locating sensors for fault diagnosis in chemical plants (Raghuraj et al., 1999). These concepts and related algorithms have been suitably adapted for designing sensor networks using a minimum number of sensors which ensures observability and identifiability of intrusions in a WDN.

A water distribution network can be represented as a graph, G = (V, E), where, E represents the edges, and V represents the vertices or nodes. Nodes are used to represent sources, such as reservoirs or tanks, from where water is supplied, as well as demand points where water is consumed. Nodes are also used to represent fire hydrants. The point where two or more pipes or a pipe divides into several branches meet is also represented as a node. Pipes, valves and pumps are represented as edges in the graph. A real life WDN can consist of several hundred nodes and pipes. Typically, nodes representing sources or fire hydrants are potentially vulnerable sites for intentional contamination, whereas unintended chemical or biological contamination can occur at any point in the WDN. In the current work, we consider only sources of WDN such as main or intermediate reservoirs, tanks, and deep wells as potential sites of intrusion, water treatment plants and pumping stations, as well as fire hydrants are also considered as potential intrusion sites. The above sites in a WDN are the most vulnerable and can be accessed relatively easily and contaminated by deliberate acts of terrorism. The nodes which are potential sites of intrusion are termed as vulnerable nodes. It is assumed that a contaminant can be introduced at any one of the vulnerable nodes of the

WDN. The contaminant is then transported along with the flow direction of water. It is also assumed that sufficient quantities of contaminant are introduced at a vulnerable node such that the concentration level of the contaminant in any pipe is above the minimum detectable level of the sensors being considered. Given a WDN and its vulnerable nodes, the problems are (a) to determine the minimum number of sensors and their location which ensures observability of the contaminant, irrespective of which vulnerable node is affected, and (b) determine the minimum number of sensors and their locations which ensures that the contaminant is observable and the contaminated vulnerable node is also identified from the sensors response.

3. Methodology

Greedy algorithm for observability:

Observability refers to the ability to detect the intrusion by at least one sensor. For a given set of sensors located in the WDN, observability condition ensures that a contaminant introduced at any vulnerable node would be observed by at least one sensor. The algorithm proposed for solving the observability problem combines graph theoretic concepts with a greedy optimization algorithm. The first phase of the algorithm is to construct a bipartite graph as follows.

Step 1: A hydraulic analysis of the WDN is carried out for a specified loading condition by considering every vulnerable node in turn as the attacked node, and the flow directions in the pipes are obtained. Based on this a directed graph of the WDN is constructed. Step 2: All directed paths from each vulnerable node to all demand points are constructed. This can be done using an efficient depth first search algorithm as described by (Deo, 1974). From the directed paths, the unions of all nodes that are present in all the directed paths corresponding to each vulnerable node are identified. These nodes are denoted as affected nodes corresponding to the vulnerable node. Step 3: The nodes are divided into two sets, one consisting of only the vulnerable nodes and the other consisting of all nodes of the graph. A bipartite graph is constructed between the two sets by adding edges from a vulnerable node to all affected nodes corresponding to the vulnerable node.

It may be noted that in Step 1 a hydraulic analysis has to be performed for each vulnerable node being considered as the attacked node. When performing a hydraulic analysis for a vulnerable node considered as the attacked node, only a single loading condition can be specified. However, when performing a hydraulic analysis for a different vulnerable node being considered as the attacked node, a different loading condition may be specified. Thus, for example, when a reservoir or other sources of water, pumping station, or treatment plant is considered as the attacked vulnerable node, the fire hydrant node can be treated as a demand node for which a demand flow rate can be specified, whereas when a fire hydrant node is considered as the attacked node, it is treated as a source node and the pressure at this node can be specified for performing the hydraulic analysis. Thus, the flow directions need not be the same for all the hydraulic analyses corresponding to different vulnerable nodes.

The observability problem is to choose the minimum number of nodes from the affected nodes on which sensors have to be located, such that there exists at least one edge from each vulnerable node to the chosen set of affected nodes. This is known as the node or set cover problem in graph theory and is known to be an NP-hard problem. A greedy algorithm for solving this problem has been developed by Raghuraj et al. (1999). The algorithm starts with an empty set of covered vulnerable nodes and repeats the following two steps.

Step 1: Choose an affected node which has maximum number of arcs incident on it. Mark this node, and add all vulnerable nodes connected to this node to the list of covered vulnerable nodes. Stop, if all vulnerable nodes are covered, else go to Step 2.
Step 2: Delete all edges from covered vulnerable nodes that are incident on the unmarked affected node and go to Step 1.

By locating sensor on all marked affected nodes, we can cover all the vulnerable nodes, which implies that if any of the vulnerable node is affected, at least one sensor with measure the response. This solves the observability problem. It should be noted that the greedy algorithm will not lead to the minimum number of sensors but is expected to give good solutions which are close to the minimum. An improved version of this algorithm is also presented in Raghuraj et al. (1999).

Greedy algorithm for identifiability

Identifiability refers to the ability to observe and identify the vulnerable node that is attacked, from the responses of the located sensors. Assuming that at most one vulnerable node is attacked, the problem of determining the minimum number of sensors for identification of the vulnerable node is formulated using an expanded bipartite graph. Let T_i be the set of affected nodes for a particular vulnerable node i. Let n be the number of vulnerable nodes. An expanded bipartite graph consisting of $n + n(n-1)/2$ vulnerable nodes is constructed as follows:

1. Define the sets $U_{ij} = U_{ji} = T_i \cup T_j - T_i \cap T_j$. The number of such sets generated is $n(n-1)/2$. Corresponding to each such set, an artificial node is added to the set of vulnerable nodes in the original bipartite graph. Thus, there are $n + n(n-1)/2$ vulnerable nodes in the expanded bipartite graph.
2. Edges are drawn from each artificial vulnerable node corresponding to the set U_{ij} to the affected nodes in the corresponding set.

The greedy algorithm described in section 3.1 is used to solve the observability problem of the expanded bipartite graph. By locating sensors on all affected nodes determined by the greedy algorithm, it is possible to observe and identify which vulnerable node is attacked. The set of sensors that covers U_{ij} is a set of sensors which respond only if vulnerable nodes T_i or T_j is affected, but not if both nodes are affected. This makes it possible to identify whether T_i or T_j is attacked. Only in the case when U_{ij} is a null set, it is not possible to distinguish whether T_i and T_j is attacked.

4. Case Study

An urban WDN taken from Mohankumar et al. (2008) is shown Figure 1. The distribution system consists of 116 pipes, 29 control valves, 3 on-off valves and 7 pumps. It has 3 source nodes each at an elevation of 100m. Hydraulic analysis of the network is carried out using EPANET 2.0 software to obtain the directed graph.

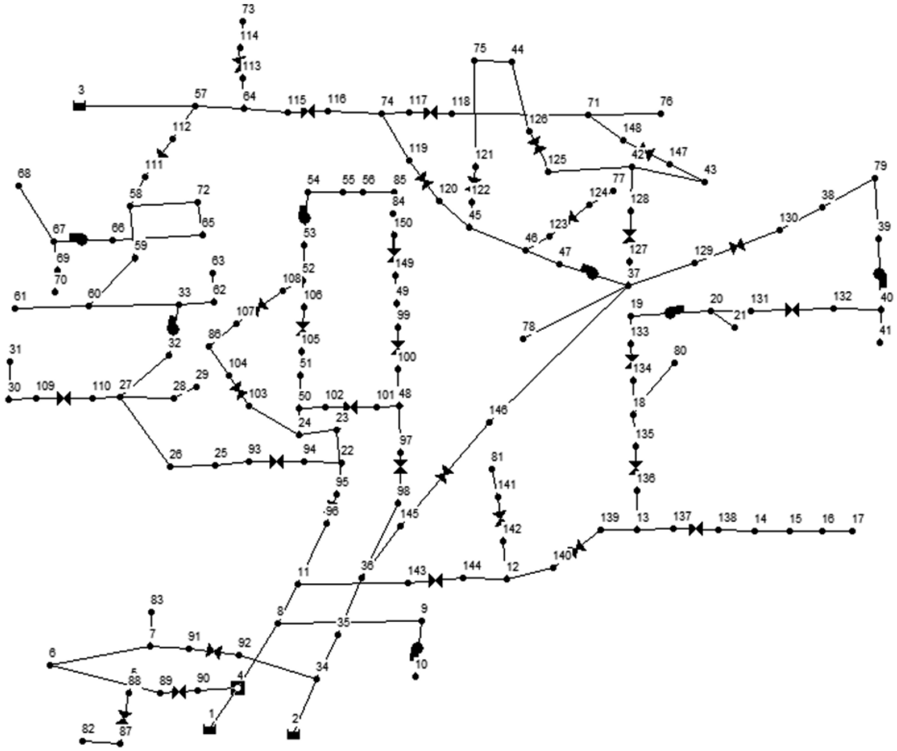

Figure 1. Layout of the water distribution network

There are 9 vulnerable nodes corresponding to nodes 1, 2, 3, 19, 32, 37, 39, 53, and 66. Nodes 1, 2, and 3 correspond to reservoirs while nodes 19, 32, 37, 39, 53, and 66 are pumping stations, and therefore these are considered as vulnerable. For simplicity, the same loadings conditions are imposed for every vulnerable node. Since there is no flow through one of the pumps (edge 155) which connects source node 10 to 9, it is not considered as a vulnerable node.

Results of the algorithm for satisfying the observability and identifiability conditions under the assumption that only one vulnerable node is attacked are shown in Table 1. The results show that for detecting whether any vulnerable node is attacked (observability condition) it is necessary to locate only three sensors at nodes 40, 54, and 67. On the other hand if it is required to identify which vulnerable node has been attacked, it is necessary to locate three additional sensors at nodes 20, 33 and 71.

Table 2 shows how the response of the sensors can be used to detect and identify the attacked vulnerable node, corresponding to the observable and identifiable sensor network designs, respectively. Row one of this table indicates for the observable sensor network design that at least one of three sensors detects a contaminant when any one of the vulnerable nodes is attacked. Second row of the table shows that different combinations of sensors detect a contaminant when different vulnerable nodes are attacked. This indicates that the attacked vulnerable node can be identified from the set of sensors that detect a contaminant.

Table 1. Results for sensor locations that satisfy observability and identifiability conditions

Criteria	Optimal sensor locations
Observability	40, 54, 67
Identifiability	40, 54, 67, 20, 33, 71

Table 2. Sensors that detect a contaminant when one of the vulnerable nodes is attacked

Attacked vulnerable node	1	2	3	19	32	37	39	53	66
Observable sensor network	40, 54, 67	40, 54	67	40	67	40	40	54	67
Identifiable sensor network	54, 20, 33	40, 71, 54	71, 67	20	33	40, 71	40	54	67

5. Conclusion

Optimal sensor network design for given water distribution network is studied. The concepts of observability and identifiability are used in the problem formulation. Greedy heuristic algorithms are proposed for the sensor network design problem using graph theory concepts. Hydraulic simulation of large scale urban water distribution system is carried out using EPANET 2.0 under various loading conditions.

References

M. A. Mustafa, J. Guan, L. M. Morris, 2010, Optimal design of sensor placement in water distribution networks, Journal of Water Resources Planning and Management, 136, 5-18.

L. Perelman, A. Ostfeld, 2013, Bayesian networks for Source intrusion detection, Journal of Water Resources Planning and Management, 139, 426-432.

A. Ostfeld, E. Salomons, 2005, Optimal layout of early warning detection stations for water distribution system security, Journal of Water Resources Planning and Management, 130, 377-385.

A. Ostfeld, J. G. Uber, 2008, The Battle of the Water Sensor Networks (BWSN): A Design Challenge for Engineers and Algorithms, Journal of Water Resources Planning and Management, 134, 556-568.

R. Raghuraj, M. Bhusan, R. Rengaswamy, 1999, Locating sensors in complex chemical plants based on fault diagnostic observability criteria, AIChE Journal , 45, 310-321.

B. Lee, R. Deininger, 1992. Optimal locations of monitoring stations in water distribution system, Journal of Environmental Engineering, 118, 4-16.

A. Kumar, A. Kansal, G. Arora, 1997. Identification of monitoring stations in water distribution system, Journal of Environmental Engineering, 123, 746-752.

A. Kessler, A. Ostfeld, G. Sinai, 1998, Detecting accidental contaminations in municipal water networks, Journal of Water Resources Planning and Management, 124, 192-198.

M. Kumar, S. Narasimhan, S. M. Bhallamudi, 2008. State estimation in water distribution network using graph-theoretic reduction strategy, Journal of Water Resources Planning and Management, 134, 395-403.

N. Deo, 1974, Graph theory with application to engineering and computer science. Prentice-Hall, Englewood Cliffs, N.J.