Available online at www.sciencedirect.com

ScienceDirect

IFAC-PapersOnLine 49-1 (2016) 492–497

Exploiting Sensor Response Times to Design Sensor Networks for Monitoring Water Distribution Networks

Venkata Reddy Palleti * Shankar Narasimhan *
Raghunathan Rengasamy *

* *Indian Institue of Technology-Madras, Chennai-600036 (e-mail: naras@iitm.ac.in)*

Abstract: Water Distribution Networks (WDNs) are an integral part of society. Deliberate introduction of chemical or biological agents through accessible sites of a WDN can spread through the entire system and cause widespread damage to public health. In order to protect against such deliberate attacks on a WDN, an effective and efficient online monitoring system through sensors is needed. It is clear that sensors located at different nodes respond at different times depending on which vulnerable node is attacked. In the present study, we design sensor networks for contamination detection and identification which exploit the differences in sensor response times as additional information. A hydraulic analysis of the network is first carried out for a given loading condition to determine the flow directions and flow velocities in different pipes. Directed paths between vulnerable nodes and potential sensor nodes are used to construct a bipartite graph, and the sensor network design problem is formulated as a minimum set cover problem. Algorithms based on greedy heuristics are used to solve the set cover problem and obtain the corresponding sensor network. The proposed method is applied on two WDNs, and the use of sensor response times to obtain a design with reduced number of sensors is demonstrated.

Keywords: Water Distribution Networks, sensor network design, sensor response times, Greedy algorithms.

1. INTRODUCTION

A water distribution network is considered to be a critical infrastructure of any city. The primary purpose of any WDN is to efficiently deliver water to the consumers for different uses such as domestic, industrial and commercial. WDNs are complex networks consisting of water storage facilities such as reservoirs, water tanks, water treatment plants, pipes and hydraulic elements such as valves, pumping stations etc. Water distribution systems are inherently vulnerable to various types of threats such as physical attacks or contamination of the water supply. Physical attacks on WDNs include destruction of pipes, pumping stations, water tanks and other facilities, which can result in inconvenience to consumers and economic loss. On the other hand, contamination of the water supply can have more severe consequences. The contamination may be accidental or intentional. Accidental contamination occurs due to the breakage of pipeline or malfunctioning of water treatment plants. An intentional contamination can occur due to deliberate acts of terrorism (Haimes et al., 1998). In such attacks, the water is contaminated by the introduction of chemical or biological agents through accessible points in a WDN. The potential consequences of the intentional contamination include public health crisis such as sickness, death and long-lasting psychological effects. Hence, it is necessary to detect the contaminants quickly in order to mitigate the effects of such deliberate attacks. Monitoring systems through sensors need to be installed to detect the contaminants. Budget constraints and maintenance reasons make it infeasible to locate sensors at every potentially vulnerable site in a network. Therefore, we need to optimize the sensor placement such that intrusions can be detected and identified quickly.

The problem of intrusion detection in a WDN has been addressed by various researchers and several optimization models and algorithms have been proposed. Single objective models were used to solve the sensor placement problem (Ostfeld and Salomons, 2004; Kessler et al., 1998; Propato, 2006; Berry et al., 2005). Later, the sensor placement problem was solved by incorporating multiple objectives such as minimizing the time of detection, minimizing the population or amount of water consumption prior to detection, and maximization of the detection likelihood (Aral et al., 2010; Ostfeld et al., 2008). Recently, Palleti et al. (2014) have proposed sensor network design algorithms for intrusion detection and identification in WDN which ensure observability and identifiability conditions. These are essential properties that need to be satisfies by all sensor network designs. In their approach the sensor network design only uses the information regarding whether sensors located at different nodes respond or not. It is evident that sensors located at different nodes respond at different times depending on which vulnerable node is attacked. However, this information is not exploited in the previously (Palleti et al., 2014) proposed design approach. In the present study, we propose sensor network

10.1016/j.ifacol.2016.03.102

design methods which exploit the differences in sensor response times as additional information, leading to a possible reduction in number of sensors. We compare the present sensor network designs with the designs proposed by Palleti et al. (2014).

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. Sources such as reservoirs or tanks, from where water is supplied and demand points where water is consumed are represented as nodes. Hydraulic elements such as pipes, valves and pumps are represented as edges in the graph. A real life WDN can consist of several hundred nodes and pipes. In general, large sections of the pipeline system are buried underground and some of the components are laid aboveground. Hence, it is a difficult task to inject the contaminants at every node of a WDN. The components of a WDN which are above ground such as sources, pumping stations, water treatment plants and fire hydrants are easily accessible for intentional contamination. Therefore, in this work we consider the access points for injecting contaminants into a WDN as reservoirs, tanks, water treatment plants, deep wells, pumping stations and fire hydrants.

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 at any point of time. Due to flow of the contaminated water from a vulnerable node that is attacked, other parts of the network are also affected. The nodes which are contaminated by vulnerable nodes are called as affected nodes. A steady state hydraulic simulation can be performed to determine the flow directions in all pipes for a specified loading condition. Based on the simulation, the set of affected nodes corresponding to each vulnerable node attack can be determined. Clearly, if a sensor is located at any affected node corresponding to a vulnerable node, then it is possible to detect whether the corresponding vulnerable node has been attacked. It is assumed that sufficient quantities of contaminant is 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 deployed. It is also assumed that the sensors can detect a wide range of contaminants and are not prone to failure. In this work, for simplicity, we assume that at most one vulnerable node is attacked at a time, although it can be extended to deal with simultaneous attack on multiple nodes also. The problem is to determine the nodes where sensors have to be located for monitoring a WDN. This is also referred to as the sensor network design problem.

2. METHODOLOGY

2.1 Algorithm for observability

The two basic properties that any sensor network design for monitoring WDNs should satisfy are observability and identifiability. Observability is defined as the ability of the sensor network to detect the presence of a contaminant in a WDN regardles of which vulnerable node is attacked. If there is an intrusion at any of the vulnerable nodes, observability condition ensures that the intrusion would be detected by at least one sensor. Identifiability refers to the ability of a sensor network to identify the exact vulnerable node that is attacked, from the responses of the located sensors. In addition to these, it is also important to detect the an attack as quickly as possible and to minimize the size of population affected before detection. These objectives are also important and can also be included to select the best design from among all observable/identifiable sensor network designs in a lexicographic manner. However, in this work we only address the problem of designing observable and identifiable sensor network design.

Palleti et al. (2014) proposed algorithms for sensor network design that satisfy observability condition. In this work, we extend these algorithms to design observable sensor networks that exploit the differences in response times of sensors located at different nodes. In the first phase of the algorithm we construct a bipartite graph between vulnerable nodes and their corresponding affected nodes identified through a hydraulic simulation as follows.

Step 1 : For a specified loading condition, hydraulic analysis of the WDN is carried out by considering every vulnerable node in turn as the attacked node, and the flow directions in all pipes are obtained. A directed graph of the WDN is constructed based on the flow directions.

Step 2 : Contaminant propagation time in each pipe is calculated assuming that contaminant transportation takes place with a velocity equal to the flow velocity of water. There can be multiple paths with different propagation times from a vulnerable node to the corresponding affected nodes. The path which takes minimum time for propagation of contaminant is chosen and the corresponding time is noted. In this way, all shortest paths are constructed from every vulnerable node to all the corresponding affected nodes using Floyd's algorithm described by Deo (1974).

Step 3 : The affected nodes and their corresponding response times are calculated for each vulnerable node. Here, we construct a set U_i consisting of ordered pairs (S_k, t_k), where S_k are the sensors responding and t_k are the corresponding minimum response times when vulnerable node i is attacked. Likewise, we construct the sets for all N vulnerable nodes. Construct a bipartite graph $\mathbf{B}$ by drawing edges from every vulnerable node to the corresponding affected nodes (S_k).

Step 4 : Construct pseudo-nodes which represent pair-wise affected nodes and also calculate the difference between their sensor response times for each vulnerable node. Here, we generate the ordered triplets $(S_k, S_l, \Delta t_{kl})$ $\forall (S_k, t_k)$, $(S_l, t_l) \in U_i$ and $k < l$. Where $\Delta t_{kl} = t_k - t_l$. For simplicity, the triplet can be represented as an ordered pair $(X_{kl}, \Delta t_{kl})$. Where X_{kl} are pseudo-nodes representing nodes S_k and S_l together. Further, the ordered pairs $(S_k, t_k) \in U_i$ are replaced by $(S_k, 0)$ for consistency in the set operations. Therefore, the set U_i contains $(S_k, 0)$, $(S_l, 0)$, $(X_{kl}, \Delta t_{kl})$ for every affected node S_k and S_l which responds to an attack on vulnerable node i. Draw edges in B from every vulnerable node to the corresponding pseudo-nodes (X_{kl}). Thus, the bipartite graph $\mathbf{B}$ has edges from every vulnerable node to the corresponding single affected nodes as well as pseudo-nodes.

Once the bipartite graph is constructed, now the second phase of the algorithm is to choose the minimum number of affected nodes on which to locate sensors such that every vulnerable would be observed by at least one chosen affected node. Even though the problem can be formulated as an integer programming problem, due to the high complexity associated with such a formulation, we propose a greedy heuristic algorithm to obtain the optimal sensor network design. A modified version of the greedy algorithm proposed by Palleti et al. (2014) is explained below.

Step 5 : Initially, make a copy **B'** of the bipartite graph **B**.

Step 6 : Choose a affected node in **B'** which has maximum number of arcs incident on it. From the bipartite matrix **B** list the vulnerable nodes which are connected to the chosen affected node. The following steps are performed.

a) If the chosen affected node is associated with a single node, mark this node and add all vulnerable nodes connected to this node to the list of covered vulnerable nodes.

b) If the chosen affected node is associated a pseudo-node, mark this node and add all vulnerable nodes connected to this node to the list of covered vulnerable nodes. In addition, the respective single-nodes which are part of the chosen pseudo-node have to be marked. Vulnerable nodes which are connected to the single nodes also added to the list of covered vulnerable nodes.

Step 7 : If one or more vulnerable nodes in the present selection is also covered in a previous selection, then remove these vulnerable nodes from latter.

Step 8 : Delete all edges from covered vulnerable nodes (obtained from *Step* 6) that are incident on the unmarked affected nodes of **B'**. If all vulnerable nodes are covered then go to *Step* 9 else go to *Step* 6 .

Step 9 : Check the list of covered vulnerable nodes of all marked affected nodes. If list is empty, discard the corresponding affected node from the list of marked affected nodes.

Step 10 : If all vulnerable nodes which are covered by a marked pseudo-node can be covered with only one of the nodes in this pair, then the corresponding single-node should be used to replace the pseudo-node.

Step 11 : The sensor network design is to locate sensors on all marked affected nodes.

2.2 Algorithm for identifiability

Assuming that at most one node is attacked, the algorithm for designing identifiable sensor networks is formulated and solved by constructing an expanded bipartite graph as follows. *Step* 1 : Initially construct the bipartite graph by performing *Steps* 1 to 3 mentioned in section 2.1.

Step 2 : Construct the sets $T_{ij} = U_i \cup U_j$ - $U_i \cap U_j$. ${}^N\mathrm{C}_2$ such sets can be generated.

Step 3 : Consider each set T_{ij} as an artificial vulnerable node and draw edges from T_{ij} to $S_k \ \forall \ (S_k, 0) \in T_{ij}$ and to $X_{kl} \ \forall \ (X_{kl}, \Delta t_{kl}) \in T_{ij}$. An expanded bipartite graph is generated by augmenting artificial vulnerable nodes and their corresponding edges to the bipartite graph **B**. Therefore, the expanded bipartite graph **B** consists of $N+{}^N\mathrm{C}_2$ vulnerable nodes and edges are drawn to their corresponding single affected nodes and pseudo-nodes.

Step 4 : Once the expanded bipartite matrix is constructed now we solve for the minimum number sensor nodes by performing *Steps* 6 to 11 on the expanded bipartite graph **B**.

3. CASE STUDY

We illustrate the proposed algorithms on two WDNs. The obtained sensor network designs are compared with the designs when response times are not incorporated, as proposed by Palleti et al. (2014). The water distribution network considered for the first case study is shown in Figure 1. The details of the network are given in the appendix. This distribution system consists of 16 nodes, 3 storage tanks, 2 reservoirs, 25 pipes and 2 pumping stations. We consider 3 storage tanks, 2 reservoirs represented by T1, T2, T3, R1, R2 respectively as vulnerable nodes for contamination. Hydraulic analysis of the network is carried out for a given loading condition using EPANET 2.0 software (Rossman, 2000). Hydraulic analysis gives us the pressures at all nodes and flow velocities in all pipes for a specified loading condition. This information is used to construct the directed graph. Here, direction of edges refer to the direction of flow between the nodes.

The sensor locations for observability problem are presented in Table 1. Table 1 shows that sensor located at 7 is sufficient to detect the contamination for both the scenarios. It is to be noted that the observable sensor network is unaffected by the additional information, i.e. response times of sensors.

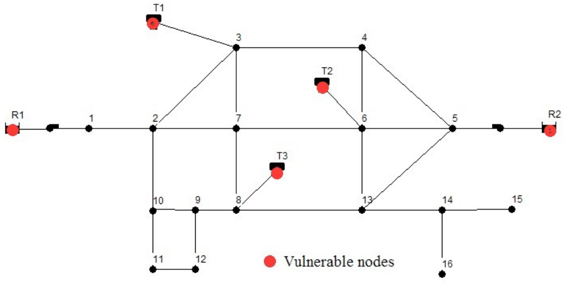

Fig. 1. Layout of WDN

Table 1. Sensor locations that satisfy Observability condition for Figure 1

Observability	Sensor locations
When sensor response times are not exploited	7
When sensor response times are exploited	7

Table 2. Sensor locations that satisfy identifiability condition for Figure 1

Identifiability	Sensor locations
When sensor response times are not exploited	3, 8, 4
When sensor response times are exploited	4, 7

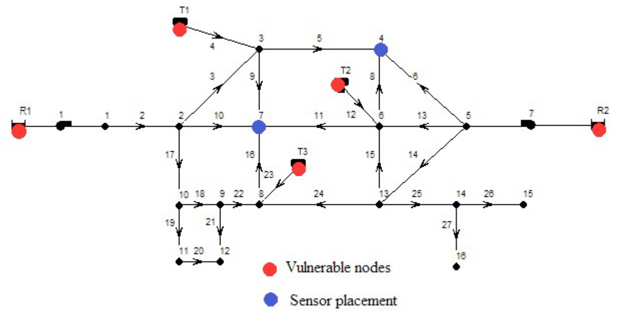

Fig. 2. Identifiable sensor network for the Figure 1 when sensor response times are exploited

The sensor locations for identifiability problem is explained in Table 2 under the assumption that at most one node is attacked. It is observed that 3 sensors (located at 3, 8, 4) are required for identifiability when sensor response times are not considered. In contrast to this, exploiting the sensor response times reduce the required number of sensors to 2. Therefore, using the sensor response times, sensor placement at the locations 4 and 7 is sufficient to identify the exact location of intrusion. The corresponding sensor network design is tabulated in Table 3.

Table 3 explains how these sensor networks are able to distinguish between the vulnerable nodes for both the cases. In the case of sensor network design when sensor response times are not included three sensors are required. For example, if three sensors located at 3, 8, 4 respond, we can conclude that node R1 is attacked. If only sensors located at 8, 4 respond then vulnerable node R2 is attacked. Thus, we can distinguish between the attacked vulnerable nodes by locating sensors at 3, 8, 4. Only two sensors are required for the identifiable sensor network when response times of sensors are included. Column 3 of Table 3 shows the sensor locations and their minimum response times (expressed in brackets) when the corresponding vulnerable node is attacked. Clearly, from Table 3 it is observed that the set of sensors which respond for an intrusion in R1, R2, T1 and T2 are the same. Though the same set of sensors respond, we can determine which vulnerable node has been attacked by using the response times of sensors. For example, the time elapsed between sensor responses for an attack on R2 is 15 minutes whereas it is 3 minutes in an event of attack on T2. If vulnerable nodes R1 or T1 is attacked, then the order in which the sensors respond is sensor at node 7 followed by sensor at node 4, whereas they respond in reverse order if vulnerable nodes R2 and T2 are attacked. This enables us to distinguish between R1, T1 and R2, T2. We can exploit the differences in sensor response times for further resolution. For example, if sensor at node 7 first responds followed by sensor at node 4, and the difference between their response times is 31 minutes, then we can conclude R1 is attacked, whereas if the difference between their response times is 10 minutes we can conclude that T1 is attacked. It may be noted that we do not need to know the time at which a vulnerable node has been attacked, because the difference in sensor response times is not dependent on the time of attack, under a single loading condition assumption.

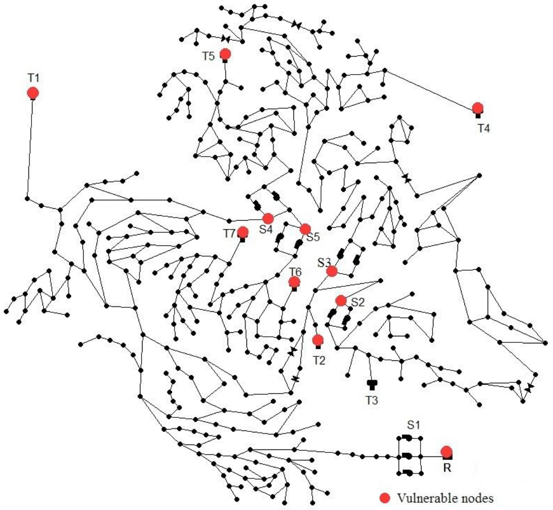

Fig. 3. Layout of D-Town Water distribution network

Table 4. Sensor locations that satisfy observability condition for Figure 3

Observability	Sensor locations
When sensor response times are not exploited	16, 56, 54, 72, 82
Sensor response times are exploited	16, 58, 57, 102, 132

Table 5. Sensor locations that satisfy identifiability condition for Figure 3

Identifiability	Sensor locations
Without sensor response times	54, 128, 231, 318, 439, 511, 1169
Sensor response times are exploited	318, 212, 56, 355, 487, 345, 128

For the second case study, D-town water distribution network is considered an urban WDN taken from Marchi et al. (2014) and is shown in Figure 3. The distribution system consists of 399 nodes, 443 pipes, 7 storage tanks, 11 pumps, 5 valves and a single reservoir. Hydraulic analysis of the network is carried out for a given loading condition using EPANET 2.0 software (Rossman, 2000) to obtain the directed graph.

We consider total 11 vulnerable nodes that includes one reservoir, four pumps, six tanks and these nodes are marked in Figure 3. Vulnerable nodes corresponding to the nodes T1, T2, T4, T5, T6 and T7 represent tanks, nodes S2, S3, S4 and S5 represent pumping stations and The node R represents reservoir.

The sensor placement for observability condition is shown in Table 4. From the table 4, it shows that the required number of sensors are same in both the cases but the sensor placement is different. Clearly, observable sensor network is not affected by the sensor response times.

Under the assumption that at most one vulnerable node is attacked, the sensor placement for identifiability is tabulated in Table 5. It is observed that 7 sensors are needed to exactly identify the attacked node for both the cases. The detailed sensor network design is explained in Table 6.

Table 3. Identifiable sensor network design for Figure 1

Vulnerable node	Identifiable sensor network	
	Without sensor response times	Exploiting sensor response times
		Sensor location {Minimum response time (in *Minutes*) }
R1	3, 8, 4	7 {*12*}, 4 {*43*}
R2	8, 4	4 {*6*}, 7 {*21*}
T1	3	7 {*13*}, 4 {*23*}
T2	4	4 {*15*}, 7 {*18*}
T3	8	7{*16*}

Table 6. Identifiable sensor network design for Figure 3

Vulnerable Node	Identifiable sensor network	
	Without sensor response times	Exploiting sensor response times
T1	128, 54, 231	318, 212, 56, 355
T2	128, 1169	128, 345, 487
T4	511, 128	128, 487
T5	54	56
T6	231	355
T7	318	318
R	318, 54, 128, 1169, 231, 439	318, 128, 56, 345, 355, 487
S2	318, 231	345
S3	318, 54	128, 487
S4	318, 54, 128, 439, 231	318, 56, 355
S5	511, 318, 54, 128, 231	318, 355

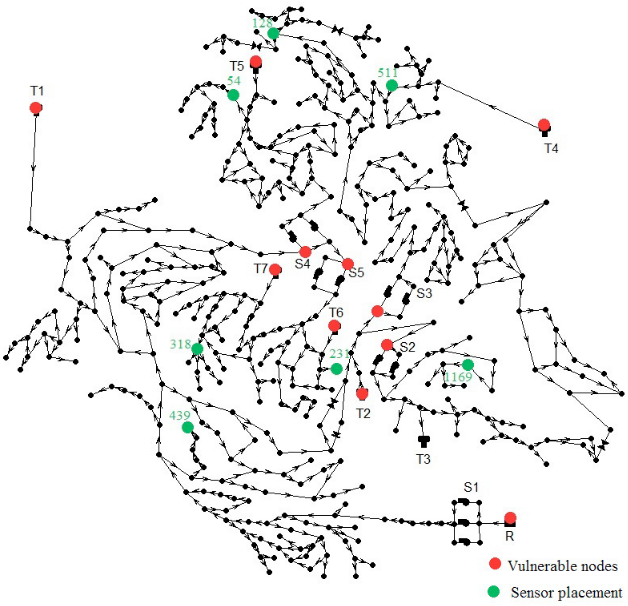

Fig. 4. Sensor placement for D-town WDN when sensor response times are not considered

From Table 6, it is possible to identify the attacked nodes exactly based on sensor responses. The sensor locations when response times of sensors are not exploited is shown in Figure 4. For example, consider the sensor network when response times are not included, if the sensors located at 318, 54 respond then we conclude that vulnerable node S3 is attacked. Likewise, we can distinguish between all vulnerable nodes. In the case of sensor network when response times are exploited, the set of sensors located at nodes 128 and 487 both respond when vulnerable nodes T4 or S3 is attacked. When node T4 is attacked the precedence order (in time) of sensors response is 128, 487 and 487, 128 when node S3 is attacked. Therefore, the information of difference between the sensor response times is useful in identifying which node is attacked. It is to be noted that the present methodology using sensor response times does not guarantee a reduction in the number of sensors (as observed in the second case study) for all WDNs in comparison to the previous methodology studied by Palleti et al. (2014). Further, it is also clear from the two case studies that the number of sensors required when response times of sensors are included is less than or equal to the number of sensors required when information of response times are not included.

4. CONCLUSION

In this work, we have proposed algorithms for designing observable and identifiable sensor networks for detecting intentional contamination of water distribution networks. The proposed algorithms exploit differences in sensor response times to reduce the number of sensors used. The proposed algorithms are illustrated using two case studies. These methods can be extended to deal with other important objectives such as minimizing time of detection or minimizing the population exposed to contamination by using a lexicographic optimization strategy.

APPENDIX

Table 7. Node data

Node No	Elevation (*feet*)	Demand (*Gallons/minute*)
1	90	0
2	110	694
3	95	694
4	105	2083
6	103	2428
7	97	2083
5	100	694
8	103	1044
13	110	0
10	112	0
11	115	350
12	112	350
9	107	0
14	120	0
15	135	175
16	130	175

Table 8. Tank data

Tank No	Elevation (feet)	Initial Level (*feet*)	Minimum Level (*feet*)	Maximum Level (*feet*)	Diameter (*feet*)
T1	220	30	0	40	50
T2	220	30	0	40	50
T3	220	30	0	40	50

Table 9. Pipe Network data

Pipe No	From Node	To Node	Length (*feet*)	Diameter (*inches*)	D-W Coefficient ($feet \times 10^{-3}$)
4	T1	3	700	6	0.025
2	1	2	800	12	0.025
3	3	2	5000	6	0.025
10	2	7	5500	10	0.025
9	3	7	3100	4	0.025
5	3	4	3700	3	0.025
8	4	6	2500	3	0.025
11	7	6	3700	4	0.025
16	7	8	2700	4	0.025
24	8	13	3100	4	0.025
15	6	13	2500	4	0.025
13	6	5	2900	10	0.025
6	5	4	3900	10	0.025
14	5	13	4500	8	0.025
25	13	14	1600	6	0.025
26	14	15	1750	6	0.025
27	14	16	1500	6	0.025
17	2	10	3100	8	0.025
19	10	11	1600	6	0.025
20	12	11	1500	6	0.025
21	9	12	1650	4	0.025
22	9	8	2900	4	0.025
18	9	10	1900	6	0.025
12	T2	6	900	6	0.025
23	T3	8	1900	8	0.025

The characteristic equation of the pump curves used in the present study is given by

- Pump curve P1 is h_p =200−3.12 × 10^{-8} × $(Q)^{2.32}$
- Pump curve P1 is h_p =180−3.12 × 10^{-8} × $(Q)^{2.32}$

Where, h_p is pump head in feet, h_0 is pump shut off head feet, Q is pump discharge (Gallons per minute).

- Reservoir R1 is elevated at 100 feet.
- Reservoir R2 is elevated at 120 feet.

Table 10. Pump Data

Pump No	From Node	To Node	Pump curve
1	R1	1	P1
7	R2	5	P2

REFERENCES

Aral, M., Guan, J., and Maslia, M. (2010). Optimal design of sensor placement in water distribution networks. *Journal of Water Resources Planning and Management*, 136(1), 5–18.

Berry, J., Fleischer, L., Hart, W., Phillips, C., and Watson, J. (2005). Sensor placement in municipal water networks. *Journal of Water Resources Planning and Management*, 131(3), 237–243.

Deo, N. (1974). *Graph theory with applications to engineering and computer science*. Prentice-Hall series in automatic computation. Englewood Cliffs, N.J. Prentice-Hall.

Haimes, Y., M., N., L., J., Jackson, B., and Fellows, J. (1998). Reducing vulnerability of water supply systems to attack. *Journal of Infrastructure Systems*, 4(4), 164–177.

Kessler, A., Ostfeld, A., and Sinai, G. (1998). Detecting accidental contaminations in municipal water networks. *Journal of Water Resources Planning and Management*, 124(4), 192–198.

Marchi, A., Salomons, E., Ostfeld, A., and Kapelan, Z. (2014). Battle of the water networks ii. *Journal of Water Resources Planning and Management*, 140(7), 04014009.

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

Ostfeld, A., Uber, J., Salomons, E., Berry, J., Hart, W., Phillips, C., Watson, J., Dorini, G., Jonkergouw, P., Kapelan, Z., di Pierro, F., Khu, S., Savic, D., Eliades, D., Polycarpou, M., Ghimire, S., Barkdoll, B., Gueli, R., Huang, J., McBean, E., James, W., Krause, A., Leskovec, J., Isovitsch, S., Xu, J., Guestrin, C., VanBriesen, J., Small, M., Fischbeck, P., Preis, A., Propato, M., Piller, O., Trachtman, G., Wu, Z., and Walski, T. (2008). The battle of the water sensor networks (bwsn): A design challenge for engineers and algorithms. *Journal of Water Resources Planning and Management*, 134(6), 556–568.

Palleti, V.R., Narasimhan, S., and Rengaswamy, R. (2014). Optimal sensor placement for contamination detection and identification in water distribution networks. In *24th European Symposium on Computer Aided Process Engineering*, volume 33, 1447 – 1452.

Propato, M. (2006). Contamination warning in water networks: General mixed-integer linear models for sensor location design. *Journal of Water Resources Planning and Management*, 132(4), 225–233.

Rossman, L.A. (2000). *EPANET 2 user's manual*. U. S. Environmental protective Agency, Cincinnati.