Received June 15, 2019, accepted July 9, 2019, date of publication August 5, 2019, date of current version August 19, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2933271

Low-Delay and High-Coverage Water Distribution Networks Monitoring Using Mobile Sensors

JUNBIN LIANG[1], XUEYUE XUE[1], ZONGJIAN HE[2], (Member, IEEE), AND VICTOR C. M. LEUNG[3], (Fellow, IEEE)

[1]Guangxi Key Laboratory of Multimedia Communications and Network Technology, School of Computer and Electronics Information, Guangxi University, Nanning 530004, China
[2]Centre for eResearch, Faculty of Science, University of Auckland, 1010 Auckland, New Zealand
[3]Department of Electrical and Computer Engineering, The University of British Columbia, Vancouver, BC V6T 1Z4, Canada

Corresponding author: Junbin Liang (liangjb@gxu.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 61562005, in part by the Natural Science Foundation of Guangxi Province under Grant 2018GXNSFBA281169 and Grant 2018JJA170215, and in part by the Thousands of Young and Middle-aged Backbone Teachers Training Program for Guangxi Higher Education [Education Department of Guangxi (2017) No. 49].

ABSTRACT Urban water distribution networks (WDNs) are usually threatened by leakage, reflux, infiltration and internal pollution. To ensure the safety of water supply, it is essential to properly monitor the WDNs. However, it is not easy to design a fast, fine-grained and comprehensive monitoring solution under limited cost, due to the large scale and high complexity of WDNs. In recent years, mobile sensor networks have been widely utilized to monitor WDNs with data uploaded by static nodes or receiver nodes. However, existing solutions haven't addressed the impact of duplicated uploading caused by the topology of WDNs. Hence, the coverage rate is unsatisfactory. In addition, the mobile sensor nodes could be potentially far away from the receiver nodes, which might introduce long data acquisition delay. In this work, we propose a low delay and high coverage WDNs monitoring solution, with the constraint of limited number of mobile and receiver nodes. The proposed solution can find the optimal deployment strategy of mobile and receiver nodes by constructing a probability dependent model between mobile and receiver nodes, and coordinating the trade-off between low-delay and high-coverage under multi-objective optimization model. We also analyzed the theoretical lower bound of the algorithm, by utilizing the submodular function characteristics of the objective. Extensive simulation proved that the proposed algorithm can obtain higher coverage with lower delay.

INDEX TERMS Mobile sensor networks, water distribution networks, node deployment, multi-objective optimization, submodular set functions.

I. INTRODUCTION

Monitoring urban Water Distribution Networks (WDNs) is vital to people's daily lives. For example in 2017, 98.4% of the Chinese urban population consumed 58 billion square meters of fresh water from water pipe as long as 757,000 kilometers [1]. However, most of the WDNs are facing serious problems like aging, corroding and internal pollutions [2]. In order to ensure the safety of water supply, it is essential to maintain the health of the WDNs and prevent potential damages in advance. Therefore, effective monitoring of the water networks is critical for both damage prevention and decision making [3], [4].

Unfortunately, due to the large scale and complicated topology of the WDNs, it is not easy to monitor the pipes and locate those with problems. Conventionally, researchers have proposed to utilize static sensors for pipeline monitoring. But static sensors have several drawbacks. E.g., the high cost (the price of a single Hach sensor is approximately USD 3,000 - 5,000 [5]), limited monitoring range, complicated installation, and highly environmental requirements (acoustics based sensors require a low noise environment [6]). Therefore, static sensors can only be applied to perform small-scaled monitoring for key positions, and is infeasible to be applied to the whole networks.

The associate editor coordinating the review of this manuscript and approving it for publication was Zehua Guo.

In recent years, the advances of computing, sensing and communication technologies enable the possibility of utilizing mobile sensors to monitor the WDNs [7], [8]. Mobile sensors can be deployed at certain water pipe junctions to obtained accurate monitoring data while moving with the water flow, and then upload data when passing through the receiver. By carrying different sensor components, mobile sensors can be used to acquire multiple types of data. For example, a mobile sensor equipped with gyroscope and acceleration sensor components can be used to acquire WDNs structure data [9], with water pressure or acoustic sensor components can be used to locate small breakages [10], with water sensing components can be used to detect the intrusion of contaminants [11].

However, the existing solutions suffer from low monitoring coverage and long delay in data collection. Since there is a dependency between the probability that mobile sensor arrives at different receivers, it is not conducive to the improvement of monitoring coverage. For example, if the mobile sensor arriving at position v_a will surely arrive at v_b, deploying receivers together at position v_a and position v_b will not achieve higher monitoring coverage than deploying a receiver only at v_b. In addition, there is a contradiction between increasing monitoring coverage and reducing delay. For example, the closer mobile sensor input position and receiver deployment position, the smaller area can be monitoring, but the faster data can be obtained. Therefore, there is a contradiction between monitoring coverage and delay. The neglect of this has led to a long delay in the existing solutions [12].

To solve these issues, this paper proposes a low delay and high coverage WDNs monitoring approach with mobile sensors. The proposed solution formulates WDNs monitoring problem with forbearing stratified multi-objective optimization problem [13]. It balances the trade-off between high coverage and low delay. The primary objective (level 1) of the proposed model is to maximize the coverage rate. By considering the arrival probability of mobile sensor nodes to the receivers, we precisely formulate the problem as a monitoring coverage maximization (MCM) problem. I.e., to maximize the ratio of monitored pipes and total pipes in WDNs. Then, by utilizing the submodular characteristics [14], we solve the problem by combining both forward and backward greedy algorithms. After obtaining the approximated optimal solution of the primary objective, we tackle the secondary objective (level 2): expected monitoring delay minimization (EMDM). I.e., to minimize the expected delay between the deployment of mobile sensor nodes and their initial data upload. Then, based on the optimal solution of the primary objective, we relax the coverage rate requirements gradually, and form a set of feasible forbearing solutions. After that, we use the interchange heuristic algorithm [14] to find out the solution for the secondary objective, which is also the optimal solution for the whole forbearing stratified multi-objective optimization. Finally, the objective of high coverage and low delay monitoring under limited cost can be achieved.

To summaries, the contribution of this paper are as follows:

1) We formulate the mobile sensor and receiver nodes deployment problem as monitoring coverage maximization and expected monitoring delay minimization problems, and combine them together into a forbearing stratified multi-objective optimization problem, in which MCM is the primary problem and EMDM is the secondary.
2) We prove that the MCM problem is NP-hard.
3) By analyzing the submodular characteristics of MCM problem, we solve the problem by using both forward and backward greedy algorithm, and analyzes the lower bound. Then, we use the interchange heuristic algorithm to solve the EMDM problem.
4) By iteratively eliminating some non-optimal solutions in the solution space, the performance of our algorithm are greatly improved.
5) We conduct extensive simulation and analysis. The results validate that the proposed approach can obtain better coverage rate while keeping lower monitoring delay.

II. RELATED WORKS

We can classify existing WDNs monitoring researches into two categories according to their data collection approaches: manual data collection and automatic data collection.

A. MANUAL DATA COLLECTION

This category of monitoring normally relies on human interference to capture the mobile sensor nodes and collect data. Therefore, they are typically applied to delay insensitive scenarios. In addition, the mobile sensor nodes are generally isolated and standalone. That is, they would neither communicate with each others, nor send data to the receiver node.

In 2010, Lai *et al.* [15] proposed a single mobile sensor based solution to probe the hidden pipelines. Their approach can construct a 3D model of the monitored pipeline by recording the accelerometer readings of mobile sensor nodes. Data are only collected after the mobile node leaves the pipeline. The solution is more applicable for small-scaled monitoring as there is only one single mobile node for data collection. In 2010, Kim *et al.* [16] proposed SPAMMS, a framework based on the fluid transported by pipelines. RFID tags are used to provide event and location related information. However, the proposed solution did not consider the collaboration among multiple mobile nodes, and did not model the mobility of mobile sensors in pipeline.

Suresh *et al.* [17], [18] proved that optimal event detection with minimum node and lowest false negative rate is an NP-hard problem. Based on that, they designed a solution called FBSM, which would deploy multiple mobile nodes and utilize beacon nodes for event detection. In [19], they further formulated the maximization of minimum pipeline monitor probability problem as an integer linear programming problem, and the maximization of mean pipeline monitor probability as an integer non-linear programming one. Solutions are also proposed accordingly.

In 2016, Oliker *et al.* [20] proposed a water pipeline pollution monitoring solution by leveraging both static and mobile sensor nodes. The solution tried to maximized the monitoring coverage by using binary integer linear programming to make the deployment decision of both mobile and static sensor nodes. It divided the water pipeline into multiple clusters, and assumed a cluster is monitored as long as there is at least one sensor in it, which can potentially cause blind monitoring areas. Meanwhile, this approach didn't mathematically model the mobility of sensors, but only offered mobile trajectory simulation, which may also hinder the monitoring results.

B. AUTOMATIC DATA COLLECTION

In automatic data collection, mobile sensor nodes can automatically upload their sensory data, which can reduce the delay of data uploading and simplify the entire pipeline monitoring process. In this category, mobile nodes can not only communicate with data center, but also collaborate with each other by exchanging information. According to the existence of dedicated receiver node in the system, we can further classify them into direct data uploading and receiver based data uploading.

1) DIRECT DATA UPLOADING

This type of solutions normally requires mobile sensors to be able to autonomously control their own mobility and communicate with data center directly, which can increase the monitoring cost significantly. In addition, the autonomous mobility of nodes entails further requirements on pipeline size and water flow speed. The direct communication with data center also poses new requirements to the depth and material of pipeline. Hence, these solutions usually have a considerable amount of limitations.

In 2011, Chatzigeorgiou *et al.* [21] proposed a pipeline leakage detection solution with autonomous mobile sensor nodes, and evaluated the impact of water flow speed, mobile arms and other factors. The proposed solution mainly focused on the design of mobile sensor nodes, and did not cover much about the applications of them into real monitoring scenarios. In 2012, Li *et al.* [22] proposed a 3D water pipeline monitoring solution. Their approach would firstly determine the minimum monitoring spots that can cover the whole pipeline by integer linear programming, and then monitor those spots by controlling a mobile sensor node to visit them. Similarly, their approach only involved in one single mobile node and didn't have multi-node collaboration.

In 2012, Lai *et al.* [23] presented an in-pipe monitoring solution with mobile sensor node networks. To ensure the network connectivity, existing mobile nodes would deploy another node within a certain distance while moving in the pipeline. Therefore, mobile nodes should stay static in most of the time. In addition, to prevent congestion in the pipeline, when the battery of a node run out, all nodes between the pipeline exit and that node should be replaced. These limitations would possibly increase the overall monitoring cost and hinder its real-world deployment.

2) RECEIVER-BASED DATA UPLOADING

In this type of solution, receiver nodes are deployed into certain positions of the water pipeline. Sensory data are only collected when the sensor nodes are in communication range of the receivers. Therefore, sensor nodes are not necessarily autonomous and can be carried by water flows. In addition, with the help of receiver nodes, the data collection delay can also be reduced, compared with the manual data collection approaches.

Trinchero *et al.* [24], [25] offered a pipeline monitoring solution with both in-pipe mobile sensor nodes and above-ground base stations. They mainly designed the hardware of mobile nodes and tested the communication between mobile nodes and base stations. Base stations have to be deployed on ground surface at a certain distance to ensure the sensory data to be collected. However, the large-scaled deployment of base stations would incur high cost and difficulty in monitoring.

In 2015, Suresh *et al.* presented OFBS [26], a mobile sensor networks based solution. The solution defined the sensing model of nodes, communication protocols among nodes, and physical layer details. Based on this, mobile sensors can pass monitored information to beacon and then to other nodes. Unfortunately, OFBS didn't answer the question that how many mobile sensors are needed to accurately detect abnormal events. Meanwhile, the deployment of beacon nodes and mobile nodes are decided without a proper mobility model. This will also bring negative effects for locating abnormal events.

In 2016, Sankary and Ostfeld [27] reported an on-line water pipeline pollution monitoring solution, which is composed of mobile sensors, static sensors and receiver nodes. In their solution, the receivers are only deployed at pipeline junctions to reduce the deployment density. However, this approach also didn't utilize any mathematical mobility model other than simulated trajectories. Meanwhile, the solution didn't explore the problem of where to deploy the receivers, but only discussed the impact of the number of receivers and pipeline junctions. These weakness could be negative for cost control.

In 2016, Du et al. [28] presented a novel solution. They modelled the problem as a mixed integer programming question, and analyzed the characteristics of it using submodular set theory [29]. Then, a greedy algorithm based solution as well as its lower bound are given. Unfortunately, this solution ignored the dependency relationship of sensors when reaching multiple receivers, which would affect the overall coverage of monitoring.

III. PROBLEM FORMULATION

In order to model the monitoring coverage maximization and expected monitoring delay minimization problem with cost constraint, we will firstly provide rigorous definitions and annotations used in this paper. Then, we will formally

TABLE 1. Summary of notations.

Symbols	Definitions
$V = \{v_1, \cdots, v_{\|V\|}\}$	A collection of pipeline junctions. v_i denotes the i-th junction.
$E = \{e_1, \cdots, e_{l_{jk}}, \cdots, e_{\|E\|}\}$	A collection of pipelines. $e_{l_{jk}}$ denotes the l-th pipeline, from v_j to v_k.
$G(V, E)$	The WDN, defined as a directed acyclic graph.
$T\_P = \{t\_p_{ij}\}$	Probability transition matrix. t_p_{ij} denotes the probability of sensors moving from v_i to v_j.
$H = \{h_1, \cdots, h_{\|H\|}\}$	Mobile sensor deployment decision set. $h_i = v_j$, $h_i \in H$ denotes mobile sensors are deployed at v_j.
V_r	The collection of junctions where receivers are deployed.
$S = \{s_1, \cdots, s_{\|S\|}\}$	Uploadable junction set. $s_i = v_j$, $s_i \in S$ denotes a receiver or a static node is deployed at v_j.
N	Maximum allowed number of mobile sensor nodes.
$T\_T = \{t\_t_{i,j}\}$	Time cost matrix. t_t_{ij} denotes the time cost for mobile sensors travel from v_i to v_j.
$M\_T = \{m\_t_{ij}(S)\}$	Mutually exclusive time matrix. $m\_t_{ij}(S)$ is the expected travel time of a node from v_i to v_j without uploading under decision S.
$Ue_{l_{jk}}(S, H)$	The probability of pipeline $e_{l_{jk}}$ is monitored under decision S, H.
$A\_T = \{a\_t_{i,j}\}$	Adjacent time matrix. a_t_{ij} denotes the travel time from v_i to v_j via pipeline $e_{l_{ij}}$.
V_s	Junction set in which static nodes are deployed.
$\Omega_{C_i} = \{C_{i,1}, \cdots, C_{i\|V\|}\}$	An event collection. Event C_{ij} denotes a mobile sensor travels from v_i to v_j.
$\Omega_{D_i} = \{D_{ij_1}, \cdots, D_{ij_{\|S\|}}\}$	An event collection. Event D_{ij_k} denotes the first uploadable junction that mobile sensors at v_i can travel is v_{j_k}.
$F\_V = \{f\_v_{ij}\}$	Water speed matrix. f_v_{ij} denotes the water speed from v_i to v_j via pipe $e_{l_{ij}}$.
$u\_p_i(S)$	$u\_p_i(S)$ denotes the probability of a mobile sensor can upload data successfully at junction v_i under decision S.
M	Maximum allowed number of receiver nodes.
$A\_P = \{a\_p_{ij}\}$	Adjacent probability matrix. a_p_{ij} denotes the probability of a node at v_i arrive at v_j via pipe $e_{l_{ij}}$.
$M\_P = \{m\_p_{i,j}(S)\}$	Mutually exclusive probability matrix. $m\_p_{ij}(S)$ is the probability of a node from v_i to v_j without uploading under decision S.

describe the pipeline monitoring problem and offer the model for monitoring coverage maximization and expected monitoring delay minimization. Finally, we will unify the whole monitoring problem with a forbearing stratified multi-objective optimization model.

A. PRELIMINARIES

Definition 1 (Monitoring Coverage (MC)):

We say a pipe is monitored as long as one mobile sensor node has passed through it and uploaded its sensory data at a certain receiver. The overall monitoring coverage rate is defined as the ratio of monitored pipes to total pipes.

Definition 2 (Expected Monitoring Delay (EMD)):

We define the expected monitoring delay of a single mobile sensor as the time interval from it is deployed into the pipeline to the first time it uploads data. The mean of expected monitoring delay of all sensor nodes is defined as overall expected monitoring delay.

The notations and symbols used in this paper are summarized in Table 1.

B. FORMAL PROBLEM DESCRIPTION

In this paper, we use a directed acyclic graph $G(V, E)$ to denote the water distribution network, in which $V = \{v_1, \cdots, v_{|V|}\}$ is the junction set of pipelines. There is no water flow from v_i to v_j if $i > j$. Set $E = \{e_1, \cdots, e_{l_{j,k}}, \cdots, e_{|E|}\}$ denotes the pipes, in which $l_{j,k} = 1 \cdots |E|$ denotes the identifier of the pipe from v_j to v_k. In addition, $|V|$ and $|E|$ are the number of junctions and pipes, respectively.

Normally, the WDNs are monitored by static sensors deployed at junctions. These static nodes are denoted by V_s, in which $v_i \in V_s$ means a static node is deployed at v_i. Since the deployment strategy of static sensors has been well studied, we treat V_s as an arbitrary constant in this research. When an exception is detected by static sensors, mobile sensors will be deployed to obtain a more accurate monitoring result. We use multi-set H to denote the deployment of mobile sensors. $h_i = v_j, v_j \in H$ means one mobile sensor is deployed at v_j. If n mobile sensors are deployed at v_j, there will be n instances of v_j in multi-set H.

Mobile sensors will accumulate sensory data while travelling in the WDNs. When they pass by a static or a receiver node, the accumulated data will be uploaded to the data center via them. We use V_r to represent the deployment of receiver nodes, $v_i \in V_r$ means a receiver node is deployed at v_i. Notice that it is unnecessary to deploy both receiver and static nodes at the same junction. I.e., $V_r \cap V_s = \phi$. To simplify the description, we use uploadable junction set S to denote the junctions that have either a receiver or a static node deployed. I.e., $S = V_r \cup V_s$. Considering the budget limit in real world deployment, we use N to denote the maximum allowed mobile sensor node and M to denote the maximum allowed receiver node.

The mobile sensors will travel along the pipes after deployed. When they arrive at a junction, they will choose a random adjacent pipe to enter. We use the same model as in literature [28]. The probability a_p_{ij} that a sensor at v_i would choose pipe $e_{l_{ij}}$ to reach v_j is determined by the water speed matrix $F\_V = \{f\_v_{ij}\}$. That is,

$$a\_p_{ij} = \frac{f\_v_{ij}}{\sum_{v_k \in V} f\_v_{ik}}, e_{l_{ij}} \in E \tag{1}$$

f_v_{ij} is the water speed in pipe $e_{l_{ij}}$ from v_i to v_j. f_v_{ij} is treated as a constant here, which means that the flow in the pipeline will not change during the movement of the mobile sensor in the pipeline. This can be achieved when the user's water demand is stable, or when the WDN is deliberately adjusted [30] to control the water flow. If such a pipe does not exist or is not opened, we set $f\_v_{ij} = 0$ and $a\_p_{ij} = 0$. Formula 1 shows that a_p_{ij} is independent from time.

In addition, let l_e_{ij} donate the length of pipe $e_{l_{ij}}$ from v_i to v_j. Since the water speed is constant, we can obtain the adjacent time matrix $A\_T = \{a\_t_{ij}\}$,

$$a\_t_{i,j} = \begin{cases} \dfrac{l\_e_{ij}}{f\_v_{ij}}, & e_{l_{ij}} \in E \wedge f\_v_{ij} \neq 0 \\ 0, & e_{l_{ij}} \notin E \vee f\_v_{ij} = 0 \end{cases} \tag{2}$$

C. PROBLEM MODELING

In order to find out a low delay and high coverage monitoring solution, we build a forbearing stratified multi-objective model, with monitoring coverage rate maximization as the primary objective, and expected monitoring delay minimization as the secondary objective. In preparation for it, we need to model both monitoring coverage rate and expected monitoring delay under decision set S and H first.

To calculate monitoring coverage rate, we need to analyze the coverage of a single mobile sensor, which includes three phases. (1) from deployment till entering the monitored pipe. (2) data collection in the monitored pipe (3) from exiting the monitored pipe till arriving an uploadable junction. The probability of entering the monitored pipe has been given by Equation 1. Then, we can obtain the monitoring coverage rate by accumulating the coverage rate of multiple mobile sensors.

Similarly, to calculate the overall expected monitoring delay, we have to get the expected monitor delay of a single mobile sensor first. To this end, we need to get the expected travel time of a single mobile sensor among junctions, and calculate the expected delay till data uploading, that is also the expected delay that the node first arrives at an uploadable junction. Finally, we can calculate the expected monitoring delay of multiple mobile sensors. The problem modelling process can be depicted in Figure 1.

1) THE PROBABILITY OF MOVING BETWEEN ADJACENT JUNCTIONS

In this paper, we use similar mobility probability model with literature [14], [15], [19], [23]. If we sort the probability that mobile sensors choose the next pipe a_p_{ij} according to their subscripts, we can get an adjacent probability matrix $A\_P = \{a\_p_{ij}\}$. Since $G(V, E)$ is acyclic, we know that A_P is an upper triangular matrix with a zero main diagonal. Therefore,

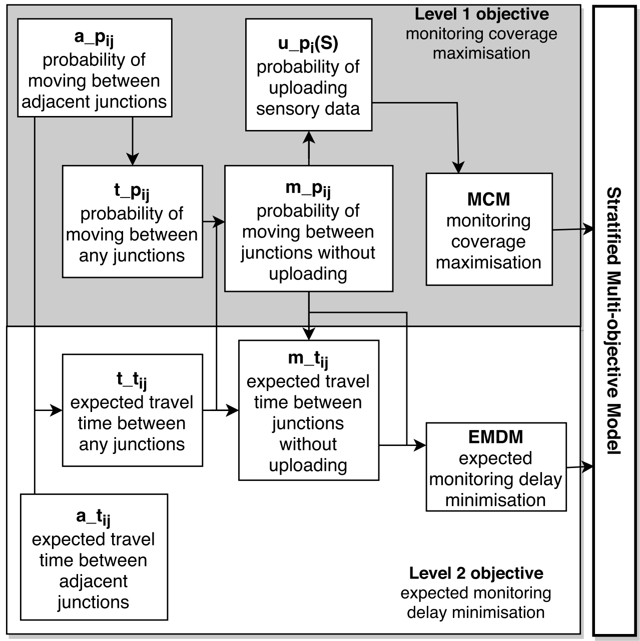


FIGURE 1. **Problem modelling.**

$A\_P^k \equiv 0, \forall k \geq |V|$. In this paper, we define a matrix T_P such that every element t_p_{ij} shows the probability of a sensor moving from v_i to v_j, through arbitrary intermediate pipelines. If I is identity matrix, then

$$T\_P = \sum_{k=1}^{|V|-1} A\_P^k + I \tag{3}$$

For convenience's sake, after obtaining the arrival probability of a single mobile sensor, we now analyze the relationships that a mobile sensor arrives at different junctions. We use C_{ij} to denote the event that a mobile sensor travels from junction v_i to v_j. Then we can obtain an event collection $\Omega_{C_i} = \{C_{ij}|, 1 \leq i, j \leq |V|\}$, and $P(C_{ij}) = t\_p_{ij}$. It is possible that a mobile sensor can arrive at multiple consecutive junctions successively, and the current position can have impact on the following ones. Therefore, the events in Ω_{C_i} are neither independent nor mutually exclusive.

2) THE PROBABILITY OF MOVING BETWEEN ANY JUNCTIONS

A single mobile sensor can upload its sensory data at any uploadable junction. Therefore, combining with uploadable junction set S, we can obtain event collection Ω_{D_i} by making Ω_{C_i} mutually exclusive.

We use D_{ij} to characterize the event that the first uploadable junction a sensor at v_i will visit is v_j. Then, we can obtain the event collection $\Omega_{D_i} = \{D_{ij_k} | v_{j_k} \in S, j_a < j_b \Leftrightarrow a < b, 1 \leq i, j_k, j_a, j_b \leq |V|, 1 < k, a, b < |S|\}$. If we have a mutually exclusive probability matrix $M\_P = \{m\_p_{ij}(S)\}$ such that $m\_p_{ij}(S) = P(D_{ij}, S)$, then we can obtain the probability of

D_{ij} as follows:

$$D_{ij} = C_{ij} - \bigcup_{v_g \in S, g<j} D_{ig} C_{gj} \tag{4}$$

$$m\_p_{ij}(S) = t\_p_{ij} - \sum_{v_g \in S, g<j} m\_p_{ig}(S) t\_p_{gj} \tag{5}$$

For probability $m\_p_{ij}(S)$, we have the following properties:

Theorem 1: $0 \leqslant m\_p_{ij}(S) \leqslant t\_p_{ij}$

Proof: It is obvious from Equation 5 above. □

Theorem 2: Events in Ω_{D_i} are pairwise mutually exclusive.

Proof: $\forall D_{ij}, D_{ik} \subseteq \Omega_{D_i}, i < j < k$, we have $D_{ij} \cap D_{ik} = D_{ij} \cap (C_{ik} - D_{i,j} C_{jk} - \cup_{v_g \in S, g<k, g \neq j} D_{ig} C_{gk}) = \phi$ □

Combining with Theorem 2, we can calculate the probability that a single mobile sensor will upload its data successfully, as:

$$u\_p_i(S) = \sum_{v_j \in S} m\_p_{ij}(S) \tag{6}$$

3) MONITORING COVERAGE MAXIMIZATION MODELLING

To simplify the description, we use $n\_p_{ijk}(S)$ to represent the probability that a node deployed at v_i failed to monitor the pipe $e_{l_{jk}}$. Obviously, we have:

$$n\_p_{ijk}(S) = 1 - t\_p_{ij} a\_p_{jk} u\_p_k(S) \tag{7}$$

The probabilities of different sensors can monitor a pipe are independent. According to sum rule of independent events, we know that pipe $e_{l_{jk}}$ can be monitored by multiple sensors with probability $Ue_{l_{jk}}$, and

$$Ue_{l_{jk}}(S, H) = 1 - \prod_{v_i \in H} n\_p_{ijk}(S) \tag{8}$$

Considering the cost limitation that only N mobile sensors and M receivers are allowed, we can formulate the monitoring coverage maximization problem $MCM(S, H)$ as model I:

$$\max_{S,H} \sum_{e_{l_{jk}} \in E} \frac{Ue_{l_{jk}}(S, H)}{|E|} \tag{9a}$$

$$\text{Subject to: } |H| \leqslant N \tag{9b}$$

$$|S| \leqslant M + |V_s| \tag{9c}$$

$$V_s \subseteq S \tag{9d}$$

Equation 9a is the overall monitoring coverage rate. Equation 9a and 9c are the constraints of mobile sensors N and receivers M. Equation 9d shows static sensors all belong to uploadable junctions.

Till now, we have finished the formulation of overall monitoring coverage rate maximization problem. Now, we will model the expected monitoring delay.

4) EXPECTED TRAVEL TIME BETWEEN ANY JUNCTIONS

Suppose we have time cost matrix $T\_T = \{t\_t_{ij}\}$, where t_t_{ij} is the expected time cost of mobile nodes moving from v_i to v_j. Let $t\_t_{ij}^{(b)}$ represents the expected time cost of mobile nodes moving from v_i to v_j via junctions b. Hence, $t\_t_{ij}^{(1)}$ shows the time cost of moving directly from v_i to v_j via pipe $e_{l_{ij}}$. We use $a\_p_{ij}^{(b)}$ to indicate the element in the i-th row and j-th column of matrix A_P to the b-th power. Then, $a\_p_{ij}^{(b)}$ is the probability of traveling from v_i to v_j via b junctions. Therefore, we can calculate t_t_{ij} and $t\_t_{ij}^{(b)}$ with:

$$t\_t_{ij}^{(b)} = \begin{cases} a\_t_{ij}, & b = 1 \\ \dfrac{\sum_{k=i+1}^{j=1} (t\_t_{ik}^{(b-1)} + a\_t_{kj}) a\_p_{ik}^{(b-1)} a\_p_{kj}}{a\_p_{ij}^{(b)}}, & \begin{cases} a\_p_{ij}^{(b)} \neq 0 \\ \wedge 2 \leqslant b \leqslant |V| - 1 \end{cases} \\ 0, & a\_p_{ij}^{(b)} = 0 \vee b \geqslant |V| \end{cases} \tag{10a}$$

$$t\_t_{ij} = \sum_{b=1}^{|V|-1} t\_t_{ij}^{(b)} \tag{10b}$$

5) EXPECTED TRAVEL TIME BETWEEN JUNCTIONS WITHOUT UPLOADING

Let mutually exclusive time matrix $M\_T = \{m\_t_{ij}(S)\}$, where $m\_t_{ij}(S)$ is the expected time cost of a single node from v_i to v_j without passing any uploadable junctions.

$$m\_t_{ij}(S) = \frac{t\_t_{ij} t\_p_{ij} - \sum_{v_k \in S, i<k<j} (m\_t_{ik}(S) + t\_t_{kj}) m\_p_{ik}(S) t\_p_{kj}}{m\_p_{ij}(S)} \tag{11}$$

Then, we can obtain the expected monitoring delay of a single mobile node, as:

$$\sum_{v_j \in S} m\_t_{ij}(S) m\_p_{ij}(S) \tag{12}$$

6) EXPECTED MONITORING DELAY MODELLING

Now, we are ready to calculate the overall expected monitoring delay $EMDM(S, H)$, under decision set S and H, as:

$$\min_{(S,H)} \sum_{v_i \in S} \sum_{v_j \in S} \frac{m\_t_{ij}(S) m\_p_{ij}(S)}{N} \tag{13}$$

7) FORBEARING STRATIFIED MULTI-OBJECTIVE OPTIMIZATION

Now, we can unify the two models into a forbearing stratified multi-objective optimization problem.

$$\max_{S,H} (\text{L1 } MCM(S, H), \text{L2 } -EMDM(S, H)) \tag{14a}$$

$$TO = \{(S, H) | MCM^1 - EMDM(S, H) \leq \epsilon\} \tag{14b}$$

$$\text{Subject to: } |H| \leqslant N \tag{14c}$$

$$|S| \leqslant M + |V_s| \tag{14d}$$

$$V_s \subseteq S \tag{14e}$$

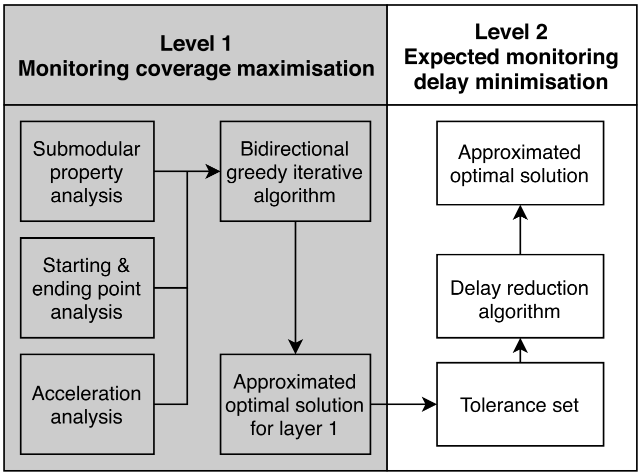


FIGURE 2. Solution design.

$MCM(S, H)$ is the monitoring coverage problem, and $-EMDM(S, H)$ is the opposite of expected monitoring delay problem. L1 $MCM(S, H)$ means the first level optimization objective is monitoring coverage maximization, and the solution is marked as MC^1. L2 $-EMDM(S, H)$ means the second level optimization objective is the oppsoite value of expected monitoring delay maximization. The feasibility solution space of L2 is TO determined by MC^1 and ϵ, which is the allowed monitor coverage decrease.

IV. SOLUTION

In this section, we firstly present the key idea of the solution, and the mathematical proof. After that, the detailed implementation of algorithms are given.

A. KEY IDEA OF SOLUTION DESIGN

Obviously, in the problem of this paper, both the mobile sensor input positions and the receiver deployment positions can be calculated before the start of monitoring and not dynamically adjusted during the monitoring process. Therefore, this paper will use a centralized algorithm to find the optimal solution. In addition, since the mobile sensor does not have the autonomous mobility capability and does not involved in the dynamic adjustment problem, the algorithm of this paper will not consider the coordination cost of communication and control.

To find out the optimal solution, we will solve the formulated problem level by level, as depicted in Figure 2. We firstly solve level one problem, monitoring coverage rate maximization by exploring the submodular property of it. Based on this, we will design a bidirectional greedy algorithm based solution.

To coordinate the execution between two greedy algorithms, we need a starting solution set for backward greedy algorithm that can gradually reduce from. Similarly an ending set is also required for forward greedy algorithm. In another word, starting / ending analysis is required. Because of the high computational complexity of solving MCM, we need to speed up the computation. To do so, we can transfer the original model into an iterative one, and also compress the feasible solution space.

After having obtained the first level solution, we will solve the second level expected monitor delay minimization problem. The first level solution is just an intermediate one, but it can offer a feasible solution space for the second level problem. Then, by adjusting the deployment of mobile sensors, we can minimise the expected monitoring delay within the tolerance set to obtain the sub-optimal solution of the second level problem, which is also the solution of the entire system.

1) SUBMODULAR FUNCTION ANALYSIS

Submodular optimization [14] is a special combinatorial optimization. It can be defined as:

Definition 3 (Submodular function):

U is a finite non-empty set and $A \subseteq B \subseteq U$ and $a \in U \setminus B$. We say set function f is submodular if $f : 2^U \rightarrow \mathbb{R}$ satisfy $f(A \cup \{a\}) - f(A) \geq f(B \cup \{a\}) - f(B)$

In addition, we define monotone of submodular function as:

Definition 4 (Monotone):

We say f is monotone if $\forall A \subseteq B$, we have $f(A) \leqslant f(B)$.

With the definition of submodular function and monotone, we will analyze the properties of MCM model.

We denote Equation 9a as $f_S(H) = g_H(S) = \sum_{e_{ljk} \in E} Ue_{ljk}(S, H)/|E|$, then we will have the following theorems:

Theorem 3: If $H = D\_H$ is a fixed value, then $U_{e_{ljk}}(S, D\_H)$ is a monotone submodular function.

Proof: For all $A \subseteq B \subseteq V$ and $v_\alpha \in V \setminus B$.

(1) Obviously, the new uploadable junctions will not reduce the monitoring coverage. So, $U_{e_{ljk}}(S, D\_H)$ is a monotone function.

(2) $U_{e_{ljk}}(B \cup \{v_\alpha\}, D\_H) - U_{e_{ljk}}(B, D\_H)$ indicates the probability that mobile sensors deployment according to B_H can reach v_α without passing B. Obviously,

$$U_{e_{ljk}}(A \cup \{v_\alpha\}, D\_H) - U_{e_{ljk}}(A, D\_H) \geqslant U_{e_{ljk}}(B \cup \{v_\alpha\}, D\_H) - U_{e_{ljk}}(B, D\_H)$$

Thus, $U_{e_{ljk}}(S, D\_H)$ is a monotone submodular function. □

Theorem 4: If $H = D\_H$ is a fixed value, then $g_{D\_H}(S)$ is a monotone submodular function.

Proof: For all $A \subseteq B \subseteq V$ and $v_\alpha \in V \setminus B$.

(1) According to the Theorem 3,

$$\begin{aligned} g_{D\_H}(A \cup \{v_\alpha\}) - g_{D\_H}(A) &= \sum_{e_{ljk}} (U_{e_{ljk}}(A \cup \{v_\alpha\}, D\_H) - U_{e_{ljk}}(A, D\_H))/|E| \\ &\geqslant 0 \end{aligned}$$

So, $g_{D\_H}(S)$ is a monotone and set function.

(2) According to the Theorem 3,

$$\begin{aligned} &g_{D\_H}(A \cup \{v_\alpha\}) - g_{D\_H}(A) \\ &\quad = \sum_{e_{ljk}} (U_{e_{ljk}}(A \cup \{v_\alpha\}, D\_H) - U_{e_{ljk}}(A, D\_H))/|E| \\ &\quad \geqslant \sum_{e_{ljk}} (U_{e_{ljk}}(B \cup \{v_\alpha\}, D\_H) - U_{e_{ljk}}(B, D\_H))/|E| \\ &\quad = g_{D\_H}(B \cup \{v_\alpha\}) - g_{D\_H}(B) \end{aligned}$$

Thus, $g_{D\_H}(S)$ is a monotone submodular function. □

Theorem 5: If $S = D\_S$ is a fixed value, then $U_{e_{ljk}}(D\_S, H)$ is a monotone submodular function.

Proof: For all $A \subseteq B \subseteq V$ and $v_\alpha \in V \setminus B$.

(1)

$$\begin{aligned} &U_{e_{ljk}}(D\_S, A \cup \{v_\alpha\}) - U_{e_{ljk}}(D\_S, A) \\ &\qquad = \prod_{v_i \in A} n\_p_{ijk}(D\_S) * (1 - n\_p_{\alpha jk}(D\_S)) \geqslant 0 \end{aligned}$$

So, $U_{e_{l j,k}}(D\_S, H)$ is a monotone set function.

(2)

$$\begin{aligned} &U_{e_{ljk}}(D\_S, B \cup \{v_\alpha\}) - U_{e_{ljk}}(D\_S, B) \\ &\quad = (U_{e_{ljk}}(D\_S, A \cup \{v_\alpha\}) - U_{e_{ljk}}(D\_S, A)) \\ &\qquad * \prod_{v_i \in B \setminus A} n\_p_{ijk}(D\_S) \\ &\quad \leqslant U_{e_{ljk}}(D\_S, A \cup \{v_\alpha\}) - U_{e_{ljk}}(D\_S, A) \end{aligned}$$

Thus, $U_{e_{l j,k}}(D\_S, H)$ is a monotone submodular function. □

Theorem 6: If $S = D\_S$ is a fixed value, then $f_{D\_S}(H)$ is a monotone submodular function.

Proof: For all $A \subseteq B \subseteq V$ and $v_\alpha \in V \setminus B$.

(1)

$$\begin{aligned} &f_{D\_S}(A \cup \{v_\alpha\}) - f_{D\_S}(A) \\ &\quad = \sum_{e_{ljk} \in E} (\prod_{v_i \in A} n\_p_{ijk}(D\_S) * (1 - n\_p_{\alpha jk}(D\_S)))/|E| \\ &\quad \geqslant 0 \end{aligned}$$

So, in combination with Definition 4, $f_{D\_S}(H)$ is a monotone set function.

(2) In combination with Theorem 5,

$$\begin{aligned} &f_{D\_S}(A \cup \{v_\alpha\}) - f_{D\_S}(A) \\ &\quad = \sum_{e_{ljk} \in E} (U_{e_{ljk}}(D\_S, A \cup \{v_\alpha\}) - U_{e_{ljk}}(D\_S, A))/|E| \\ &\quad \geqslant \sum_{e_{ljk} \in E} (U_{e_{ljk}}(D\_S, B \cup \{v_\alpha\}) - U_{e_{ljk}}(D\_S, B))/|E| \\ &\quad = f_{D\_S}(B \cup \alpha) - f_{D\_S}(B) \end{aligned}$$

Thus, $f_{D\_S}(H)$ is a monotone submodular function. □

Based on Theorem 4 and 6, we know that the monitoring coverage model I can only be maximized when both Equation 9a and 9c reach their upper bound. Hence, model I is equal to the model II below.

$$\max_{S,H} \sum_{e_{l j,k} \in E} \frac{Ue_{l j,k}(S, H)}{|E|} \tag{15a}$$

$$\text{Subject to: } |H| = N \tag{15b}$$

$$|S| = M + |V_s| \tag{15c}$$

$$V_s \subseteq S \tag{15d}$$

Theorem 7: The monitoring coverage rate maximization problem I and II are NP-Hard.

Proof: Considering a special case of coverage rate maximization *II*: given optimal solution S^*, solve for optimal solution H^*. This is reduced to a typical submodular optimization problem, which is known to be NP-hard [8]. Therefore, coverage rate maximization *II* is also NP-hard. And its equivalent problem coverage rate maximization *I* is also NP-hard. □

2) START AND EXIT CONDITIONS ANALYSIS

According to Theorem 7, the problem we are solving is NP-hard, which means there is no polynomial-time optimal solution, unless $P = NP$. To reduce the complexity, we solve the problem with two steps: (1) given an uploadable junction set S, solve the optimal mobile sensor deployment set H, $g'(S) = \max_H f_S(H)$. (2) solve the optimal uploadable junction set S.

We use greedy algorithm for both steps, as greedy algorithms perform well for monotone submodular optimizations [23]. The approximation ratio is close to $1 - (1 - K)^K$ under forward greedy algorithm, where K is the number of iterations in greedy algorithm. Correspondingly, the approximation ratio is close to $1 - (1 - \frac{1}{Z-K})^{Z-K}$ under backward greedy algorithm, where Z is the total number of elements in universal set, $Z - K$ is the number of iterations in the algorithm.

For problem (1), according to Theorem 6, $f_{D\_S}(H)$ is a monotone submodular function over multi-set H. Given D_S, we can find the approximated optimal solution $f_{D\_S}(H)$ and $g'(D\_S)$ with forward greedy algorithm. To start off, we can set $H = \phi$ and the algorithm ends when $|H| = N$.

Similarly for problem (2), according to Theorem 4, $g_{D\_H}(S)$ is also a monotone submodular function over set S, and hence greedy algorithm can also be applied. Differently, if the number of receivers $|V_r|$ is unconstrained, universal set $V - V_s$ would be the optimal deployment solution. This motivates us of applying backward greedy algorithm. I.e., starting the solution set from V and removing receivers gradually in every iteration.

Furthermore, during the execution of backward greedy algorithm, an intermediate solution *ISink* can be obtained, such that if receivers are deployed at every node in *ISink*, the upload probability is maximized, i.e., $u\_p_i(S) = 1$. Otherwise, if any of the node in *ISink* has no receiver deployed, the upload probability would be $u\_p_i(S) < 1$. Such *ISink* can be given by:

Proposition 1: Minimum Complete Upload Set

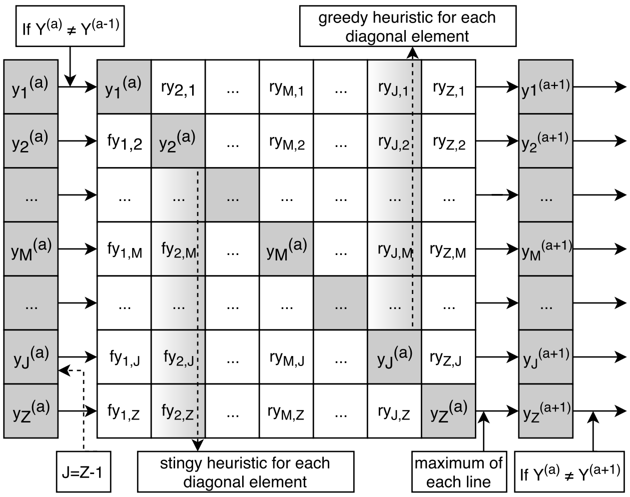


FIGURE 3. Greedy algorithm.

For the set $ISink = \{v_i | v_i \in V, \sum_{j=1}^{|V|} t\_p_{ij} = 0\}$, *the following properties are satisfied:*

(1) $1 \leqslant i \leqslant |V|, u\_p_i(ISink) = 1$

(2) $\forall S \subset ISink, \exists 1 \leqslant i \leqslant |V|, u\_p_i(S) < 1$

Proof:

(1) For the first property of *ISink*, it is proved by the counter-evidence method.

Suppose $\exists 1 \leqslant i \leqslant |V|$, then we have $u\_p_i(ISink) \neq 1$, in other words $\exists u\_p_i(ISink) < 1$.

That is, there is a mobile sensor in v_i, without passing any receiver, leaving the network from the terminal junction v_j.

That is, $\exists v_j \in V, \sum_{j=1}^{|V|} t\_p_{ij} = 0$ and $v_j \notin ISink$. This contradicts the definition of *ISink*.

So, $\forall 1 \leqslant i \leqslant |V|, u\_p_i(ISink) = 1$

(2) For the second property of *ISink*, assuming $v_j \in ISink$, then obviously the mobile sensor put in v_j will leave the WDN directly.

So,

$$u\_p_j(S) = 0 \neq 1$$

So, $\forall S \subset ISink, \exists 1 \leqslant i \leqslant |V|, u\_p_i(S) < u\_p_i(ISink) = 1$ □

So for the problem (2), the starting point of the forward greedy algorithm is $\varnothing$, and the end point is set to $|S| = |ISink|$, The starting point of the backward greedy algorithm is *ISink*, and the end point is $\varnothing$.

3) BIDIRECTIONAL GREEDY ALGORITHM

After analyzing the start and exit conditions of the greedy algorithms, we demonstrate how to use both greedy algorithms to find out receiver deployment set in Figure 3.

Take the complete upload set *ISink*, $|ISink| = Z$. Obviously, the number of receivers is usually insufficient to ensure that the entire WDN is completely covered, that is, there are $M < Z$. Greedy algorithm specific cooperation process is as follows:

(1) Initialization: Let a=0, perform initial backward greedy, let $x_Z^{(a)'} = ISink$. Then, let $x_i^{(a)'} = x_{i+1}^{(a)'} \backslash \{argmax_{v_\alpha \in x_{i+1}^{(a)'}} g'(x_{i+1}^{(a)'} \backslash \{v_\alpha\} \cup V_s)\}$, until $i = 1$. Let $x_o = \varnothing$. Then perform initial forward greedy, let $x_i^{(a)} = x_{i-1}^{(a)} \cup \{argmax_{v_\alpha \in V \backslash (x_{i-1}^{(a)} \cup V_s)} g'(x_{i-1}^{(a)} \cup \{v_\alpha\} \cup V_s)\}$, until $i = Z$. Let $Y^{(a)} = \{y_1^{(a)}, \ldots, y_Z^{(a)}\}$, where $y_i^{(a)} = max\{x_i^{(a)'}, x_i^{(a)}\}$.

(2) Iteration: Forward greedy for each element $y_i^{(a)}$ in $Y^{(a)}$, let $fy_{i,i}^{(a)} = y_i^{(a)}$, then each step, let $fy_{i,b}^{(a)} = fy_{i,b-1}^{(a)} \cup \{argmax_{v_\alpha \in V \backslash (fy_{i,b-1}^{(a)} \cup V_s)} g'(fy_{i,b-1}^{(a)} \cup \{v_\alpha\} \cup V_s)\}$ until $b = Z$. Backward greedy for each element $y_i^{(a)}$ in $Y^{(a)}$, let $ry_{i,i} = y_i^{(a)}$, then every step, let $ry_{i,b}^{(a)} = ry_{t,b+1}^{(a)} \backslash \{argmax_{v_\alpha \in ry_{i,b+1}^{(a)}} g'(ry_{i,b+1}^{(a)} \backslash \{v_\alpha\} \cup V_s)\}$ until $b = 1$.

(3) Judgment: Let $Y^{(a+1)} = \{y_1^{(a+1)}, \ldots, y_Z^{(a+1)}\}$. where $y_i^{(a+1)} = max\{ry_{i,1}^{(a)}, \ldots, ry_{i,i-1}^{(a)}, y_i^{(a)}, fy_{i,i+1}^{(a)}, \ldots, fy_{i,Z}^{(a)}\}$. if $Y^{(a+1)} = Y^{(a)}$, complete the iteration, let $a + 1 = d$, go to (4); otherwise go to (2).

(4) End: we get the algorithm solution: $MCM^1 = y_M^{(d)}$.

4) ALGORITHM ACCELERATION

The computational complexity of model II is considerably high with a large amount of feasible solutions. Two approaches can be utilized to speed up the computation. The first is to transform the coverage rate maximization model into an iterative one; The second is to eliminate feasible solutions to be searched without demoting the final solution of maximum coverage rate.

We firstly introduce how to convert model II into an iterative model III. Let $\overline{U_{e_{ljk}}}(S, H) = 1 - U_{e_{ljk}}(S, H)$, $H(n)$ is a set H with n elements, $\overline{U_{e_{ljk}}}(S, H)$ denotes the probability of mobile sensors not covering v_j with given S and H. If $H(n+1) = H(n) \cup \{v_\alpha\}$, then we can obtain the following accelerated iterative coverage rate maximization model III.

$$MCM(S, H) = 1 - OBJECTIVE(S, H)/|E| \tag{16a}$$

$$OBJECTIVE(S, H) = \min_{S,H} \sum_{e_{ljk} \in E} \overline{U_{e_{ljk}}}(S, H) \tag{16b}$$

$$\begin{cases} \overline{U_{e_{ljk}}}(S, \{v_i\}) = n\_p_{ijk}(S) \\ \overline{U_{e_{ljk}}}(S, H(n+1)) = \overline{U_{e_{ljk}}}(S, H(n))\overline{U_{e_{ljk}}}(S, \{v_\alpha\}) \end{cases} \tag{16c}$$

Subject to:

$$|H| = N \tag{16d}$$

$$|S| = M + |V_s| \tag{16e}$$

$$V_s \subseteq S \tag{16f}$$

Equation 16a and 16b are the objective functions, and Equation 16c shows the iterative process. Equation 16d, 16e, and 16f are the constraints.

Proposition 2: Some insert placement of mobile sensors that can be excluded in H. $\forall i_1 \forall i_2 \forall S \forall H$, *we have*

$$t\_p_{i_1 i_2} = 1 \rightarrow OBJECTIVE(S, H \cup \{v_{i_1}\}) \leqslant OBJECTIVE(S, H \cup \{v_{i_2}\})$$

Proof: Since $t\_p_{i_1\ i_2} = 1$,

$$t\_p_{i_1\ j} \geqslant t\_p_{i_1\ i_2} t\_p_{i_2\ j} = t\_p_{i_2\ j}$$

Then,

$$n\_p_{i_1\ jk}(S) \leqslant n\_p_{i_2\ jk}(S)$$

So,

$$\begin{aligned} OBJECTIVE(S, H \cup \{v_{i_1}\}) &= \sum_{e_{l_{jk}} \in E} \overline{U_{e_{l_{jk}}}}(S, H \cup \{v_{i_1}\}) \\ &= \sum_{e_{l_{jk}} \in E} \overline{U_{e_{l_{jk}}}}(S, H) n\_p_{i_1\ jk}(S) \\ &\leqslant \sum_{e_{l_{jk}} \in E} \overline{U_{e_{l_{jk}}}}(S, H) n\_p_{i_2\ jk}(S) \\ &= OBJECTIVE(S, H \cup \{v_{i_2}\}) \end{aligned}$$

□

Proposition 3: Some deployment placement of receivers that can be excluded in S. $\forall i_1 \forall i_2 \forall S \forall H, v_{i_1} \notin S, v_{i_2} \notin S$, we have

$$\begin{aligned} t\_p_{i_1\ i_2} = 1 \rightarrow\ & OBJECTIVE(S \cup \{v_{i_1}\}, H) \\ & \geqslant OBJECTIVE(S \cup \{v_{i_2}\}, H) \end{aligned}$$

Proof: (1) First, use the counter-evidence method to prove $t\_p_{i_1\ i_2} = 1 \rightarrow u\_p_j(S \cup \{v_{i_1}\}) \leqslant u\_p_j(S \cup \{v_{i_2}\})$.

Suppose $\exists i_1 \exists i_2, t\_p_{i_1\ i_2} = 1 \wedge u\_p_j(S \cup \{v_{i_1}\}) > u\_p_j(S \cup \{v_{i_2}\})$, then there must be a case where a mobile sensor pass by a receiver deployed in accordance with $S \cup \{v_{i_1}\}$ from the junction v_j, but is not received by a receiver deployed in accordance with $S \cup \{v_{i_2}\}$.

Considering that the two placement schemes differ only in the junctions v_{i_1} and v_{i_2}. Further, there must be some case in which the mobile sensor can be received by the receiver placed at the junction v_{i_1}, but it is not received by the receiver placed at the junction v_{i_2}.

However, since $tp_{i_1\ i_2} = 1$, any mobile sensor passing through the junction v_{i_1} must pass through the junction v_{i_2}. Therefore, $t\_p_{i_1\ i_2} = 1 \rightarrow u\_p_j(S \cup \{v_{i_1}\}) \leqslant u\_p_j(S \cup \{v_{i_2}\})$.

(2) Then, we get:

$$\begin{aligned} OBJECTIVE(S \cup \{v_{i_1}\}, H) &= \sum_{e_{l_{jk}} \in E} \overline{U_{e_{l_{jk}}}}(S \cup \{v_{i_1}), H\} \\ &= \sum_{e_{l_{jk}} \in E} \prod_{v_i \in H} n\_p_{ijk}(S \cup \{v_{i_1}\}) \\ &\geqslant \sum_{e_{l_{jk}} \in E} \prod_{v_i \in H} n\_p_{ijk}(S \cup \{v_{i_2}\}) \\ &= OBJECTIVE(S \cup \{v_{i_2}\}, H) \end{aligned}$$

□

Proposition 4: Some deployment placement of receivers that can be excluded through computing in S. $\forall i_1 \forall S \forall H, v_{i_1} \notin S$, we have

$$\begin{aligned} u\_p_i(S) = 1 \rightarrow\ & OBJECTIVE(S \cup \{v_{i_1}\}, H) \\ & = OBJECTIVE(S, H) \end{aligned}$$

Proof: Since $u\_p_i(S) = 1$,

$$\begin{aligned} u\_p_i(S \cup \{v_{i_1}\}) &= 1 = u\_p_i(S) \\ n\_p_{ijk}(S \cup \{v_{i_1}\}) &= n\_p_{ijk}(S) \end{aligned}$$

Then, we get

$$\begin{aligned} OBJECTIVE(S \cup \{v_{i_1}\}, H) &= \sum_{e_{l_{jk}} \in E} \prod_{v_i \in H} n\_p_{ijk}(S \cup \{v_{i_1}\}) \\ &= \sum_{e_{l_{jk}} \in E} \prod_{v_i \in H} n\_p_{ijk}(S) \\ &= OBJECTIVE(S, H) \end{aligned}$$

□

The above proposition will provide a basis for excluding some feasible solution space to be searched.

5) LEVEL 2 SOLUTION

After obtaining the solution of the first layer $MCM^1 = (S^1, H^1)$, use the tolerance set $To = \{(S, H) | MCM^1 - MCM(S, H) \leqslant \varepsilon\}$ to provide a feasible solution space for the second layer. In order to search for the minimum expected monitoring delay in the tolerance set *To*, an interchange algorithm can be used to find the approximate solution.

Let equation 13 is $d(S, H)$. Same as above, $g'(y_{t+1}) = max_H f_{y_{t+1}}(H)$. Then,

(1) Initialization: Let $t = 0$, $y_0 = S$, and $g'(y_0) = H$.

(2) Iteration: Let $y_{t+1} = argmax_{q \in Lat_{t+2}} g'(q)$, $Lat_{t+1} = \{y_{t+1} || y_{t+1} - y_t| - |y_t - y_{t+1}| = 1, V_s \subset y_t, V_s \subset y_{t+1}, MCM^1 - w(y_{t+1}) \leqslant \varepsilon, d(y_{t+1}, g'(y_{t+1})) < d(y_t, g'(y_t))\}, w(y_{t+1}) = MCM(y_{t+1}, g'(y_{t+1}))$.

(3) Judgment: if $Lat_t = \varnothing$, complete the iteration, go to (4); otherwise go to (2).

(4) End: we get the final solution of the interchange algorithm $(y_{t+1}, g'(y_{t+1}))$. It is also the target problem, high coverage The low-delay mobile sensor monitors the decision set of the approximate solution.

B. ALGORITHM IMPLEMENTATION

To simplify the description, we only highlight the key functions and their arguments.

1) FEASIBLE SOLUTION SPACE PRUNING

`FEASIBLESOLUTIONPRU`

This function takes probability transition matrix T_P and static sensor deployment V_s as input, and outputs column vectors *AH*, *ASink*. 1 in *AH* means mobile sensor can be deployed at the corresponding junction, 0 means no deployment. Similarly, 1 in *ASink* means a receiver node can be deployed at the corresponding junction.

This function implements proposition 2 and 3 to prune the feasible solution space. It worth noting that according to those propositions, the solution space pruning won't affect the monitoring coverage rate.

2) MINIMUM COMPLETE UPLOAD FUNCTION MINCOMPLETEUPLOAD

This function takes probability transition matrix T_P and static sensor deployment V_s as input, and returns minimum complete upload column vector *ISink* and the corresponding receiver number *nISink* as output.

This function implements proposition 1 to obtain the complete upload set, under which all the deployed mobile sensors can upload their sensory data. *nISink* is the maximum amount of receivers that the forward greedy algorithm can reach. *ISink* is the starting point of the backward greedy algorithm.

3) UNIDIRECTIONAL GREEDY FUNCTION UNIDIGREEDY

This function takes the following input parameters: probability transition matrix T_P, adjacent probability matrix A_P, static nodes deployment vector V_s, vector AH that mobile sensor can be deployed, vector *ASink* that mobile sensor can be deployed, starting receiver deployment vector *StartSink*, maximum amount of receiver node M, maximum amount of mobile node N, and flag b. It will return the following values: mobile sensors deployment matrix MH, receivers deployment matrix *MSink*, monitoring coverage rate column vector MU, where $MH(i, :)$ denotes the deployment of mobile sensors under i receivers, $MSink(i, :)$ denotes those i receivers, and $MU(i)$ is the corresponding monitoring coverage rate.

When the flag $b = 1$, the algorithm will greedily add receivers to *StartSink* to maximize the monitoring coverage rate. on the contrary, when $b = 0$, the algorithm will greedily take receivers away from *StartSink* while keeping the maximized monitoring coverage rate. During the greedy process, for every receiver deployment vector, it keeps adding mobile sensors from nothing to maximize the monitoring coverage rate, which is also used as the maximum monitoring coverage rate of the corresponding receiver deployment vector. The function implements Equation 16a to 16f in the iterative model II, and the time complexity is $O(MN|V|)$.

4) AN ITERATION OF BIDIRECTIONAL GREEDY FUNCTION BIDIRECTIONALGREEDY

A better result can be obtained by utilizing the differences between forward and backward greedy algorithms.

This function takes the following parameters as inputs: probability transition matrix T_P, adjacent probability matrix A_P, static nodes position vector V_s, AH and *ASink* for preprocessing acceleration, maximum allowed amount of receivers Z, multi-dimensional array MH, *MSink* and MU for recording the change of receivers. The outputs would be: the iterative result RH, *RSink*, RU. Line 4 is used to record forward greedy algorithm results, and Line 6 is used to record backward greedy algorithm results. Lines 4 through 4 correspond to the second step of Section IV-A.3. Lines 8 to 10 select all maximum monitoring coverages with the same number of receivers and record the results in the return parameters The computational complexity of this function is $O(NM|V|^3)$.

Algorithm 1 An Iteration of Bidirectional Greedy Function BidirectionalGreedy

Input: $T\_P, A\_P, V_s, AH, ASink, MH, MSink, MU, Z$
Output: $RU, RH, RSink$

1: $RH \leftarrow 0_{|V|*Z}, RSink \leftarrow 0_{|V|*Z}, RU \leftarrow 0_{|V|}$, $TH \leftarrow 0_{|V|*Z*Z}$, $TSink \leftarrow 0_{|V|*Z*Z}$, $TU \leftarrow 0_{|V|*Z}$
2: **for** $a = 1$ **to** Z **do**
3: $b \leftarrow a + 1, c \leftarrow a - 1$
4: $[TH(:, a, b : Z), TSink(a, b : Z), TU(a, b : Z)] \leftarrow UnidiGreedy\{T\_P, A\_P, V_s, AH, ASink, MSink(:, a), M, N, 1\}$
5: $TH(:, a, a) \leftarrow MH(:, a)$, $TU(a, a) \leftarrow MU(a)$, $TSink(:, a, a) \leftarrow MSink(:, a)$
6: $[TH(:, a, 1 : c), TSink(a, 1 : c), TU(a, 1 : c)] \leftarrow UnidiGreedy\{T\_P, A\_P, V_s, AH, ASink, MSink(:, a), M, N, 0\}$
7: **end for**
8: **for** $a = 1$ **to** Z **do**
9: $b \leftarrow max(TU(:, a))$, $RH(:, a) \leftarrow MH(:, b, a)$, $RSink(a) \leftarrow TSink(:, b, a)$, $RU(a) \leftarrow TU(b, a)$
10: **end for**
11: **return** $RU, RH, RSink$

5) DELAY REDUCTION FUNCTION DELAYREDUCTION

After obtaining the approximated solution of maximum coverage rate, this function tries to alter the deployment positions of mobile sensors to find the solution with minimum delay.

This function takes the following parameters as inputs: probability transition matrix T_P, adjacent probability matrix A_P, time cost matrix T_T, static nodes position vector V_s, mobile sensor deployment vector H, receiver deployment vector *Sink*, the corresponding coverage rate sU and the size of forbearing set ε. The outputs would be: the mobile sensor deployment vector after delay reduction HL, the corresponding coverage rate sUL and expected monitoring delay lyL. Lines 6 to 11 are the second level solution according to section IV-A.5. Lines 12 to 14 are used to find the position of the mobile sensor input position that can reduce the expected monitoring delay and maximize the monitoring coverage rate. The computational complexity of this function is $O(g|V|^2)$, where g is the iteration of while loop.

6) THE COMPLETE ALGORITHM BGI

Till now, we can combine both forward and backward greedy algorithms together according to Figure 3. Meanwhile, we can also combine the maximum coverage rate L1 and minimum expected delay L2 together based on Figure 2. By doing so, we can obtain Bidirectional Greedy Iterative algorithm (BGI).

This function takes the following parameters as inputs: probability transition matrix T_P, adjacent probability matrix

Algorithm 2 Delay Reduction Function DelayReduction

```
Input: T_P,A_P,T_T,V_s,H,Sink,sU,ε
Output: HL,sUL,lyL
RH ← H,S ← Sink|V_s,sUL ← 0, SU ← 0_{|V|*1},b ← 1
ly ← Computing EMD(T_P, T_T, H, S)
while sign do
    b ← 0
    for i = 1 to |V| do
        if RH(i) > 0 then
            RH(i) ← RH(i) − 1
            for j = 1 to |V| do
                RH(j) ← RH(j) + 1
                tU ← Computing MC(T_P, A_P, RH, S)
                tly ← Computing EMD(T_P, T_T, RH, S)
                if sU − tU < ε&&tly < ly&&tU > sUL then
                    sUL ← tU,HL ← RH,ly ← tly,b ← 1
                end if
                RH(j) ← RH(j) − 1
            end for
        end if
        if b == 1 then
            BREAK
        end if
    end for
end while
return HL,sUL,lyL
```

Algorithm 3 The Complete Algorithm BGI

```
Input: T_P,A_P,T_T,V_s,N,M,ε
Output: sU,ly,H,Sink
1: [AH, ASink] ← FeasibleSolutionPru(T_P, V_s)
2: [ISink, nISink] ← MinCompleteUpload(T_P, V_s)
3: [BH, BSink, BU] ← UnidiGreedy{T_P, A_P, V_s, AH, ASink
   , ISink, nISink, N, 0 }
4: [FH, FSink, FU] ← UnidirectionalGreedy{T_P, A_P, V_s, AH, ASink
   , 0_{|V|*1}, nISink, N, 1}
5: while FU ≠ BU do
6:     for i = 1 to |V| do
7:         if FU(i) < BU(i) then
8:             FH(:, i) ← BH(:, i),FU(i) ← BU(i),
               FSink(:, i) ← BSink(:, i)
9:         end if
10:    end for
11:    [BH, BSink, BU] ← BidirectionalGreedy{T_P, A_P, V_s, AH, ASink
       , FH, FSink, FU, nISink}
12:    [FH, FSink, FU] ← BidirectionalGreedy{T_P, A_P, V_s, AH, ASink
       , BH, BSink, BU, nISink}
13: end while
14: [H, sU, ly] ← DelayReduction{T_P, A_P, T_T, V_s
    , FH(M), FS(M), FU(M), ε}
15: return sU,ly,H,Sink
```

A_P, time cost matrix T_T, static nodes position vector V_s, maximum amount of mobile node N, maximum amount of receiver node M and the size of forbearing set ε. The outputs would be: the monitoring coverage rate sU and expected monitoring delay ly, the mobile sensor deployment vector H and the receiver deployment vector $Sink$. Line 1 to 13 correspond to maximum coverage rate L1. Line 14 minimum expected delay L2.

Line 1 performs algorithm acceleration. Line 2 gets the minimum complete upload receiver deployment location. Line 3 and line 4 correspond to the first step of Section IV-A.3. Line 5 to 13 correspond to the second step and third step of Section IV-A.3. Line 14 is used to reduce the expected monitoring delay. The computational complexity of this function is $O(rNM|V|^3)$, where r is the number of iteration loop.

C. ALGORITHM ANALYSIS

For layer one monitoring coverage rate, let the approximate solution of the algorithm to be $F^1 = MCM(S^1, H^1)$, and the optimal solution to be $F^{op} = MCM(S^{op}, H^{op})$. To simplify the analysis, similarly to literature [23], we assume that we can obtain the optimal sensor deployment set H, then we have the following theorem:

Theorem 8: The lower bound of the approximation ratio of the first layer of the BGI algorithm to maximize the monitoring coverage is:

$$F^1/F^{op} \geqslant max\{1-(1-1/M)^M, 1-(1-1/(|V|-M))^{|V|-M}\} \geqslant 1-(1-2/|V|)^{|V|/2}$$

Proof: Since $H^1 = H^{op}$, according to Theorem 6, this is a submodular optimal problem. Because of the greedy algorithm, then,

(1) For the forward greedy algorithm part, assume that the approximate solution of the forward greedy algorithm is $MCM(S1^1, H^{op})$. According to the literature [8],

$$\frac{g_{H^{op}}(S1^1) - g_{H^{op}}(\varnothing)}{g_{H^{op}}(S^{op}) - g_{H^{op}}(\varnothing)} \geqslant 1-(1-1/M)^M$$

At the same time, $MCM(S, H) = g_H(S)$ and $MCM(\varnothing, H) = g_{H^{op}}(\varnothing) \geqslant 0$, so,

$$\frac{MCM(S1^1, H^{op})}{MCM(S^{op}, H^{op})} = \frac{g_{H^{op}}(S1^1)}{g_{H^{op}}(S^{op})} \geqslant \frac{g_{H^{op}}(S1^1) - g_{H^{op}}(\varnothing)}{g_{H^{op}}(S^{op}) - g_{H^{op}}(\varnothing)} \geqslant 1-(1-1/M)^M$$

(2) For the backward greedy algorithm part, according to Proposition 4, it can be seen that the backward greedy algorithm is equivalent to starting from the $S = 1_{|V|*1}$ and starting from $ISink$, for greedy will surely go through $ISink$.

Assume that the approximate solution of the backward greedy algorithm is $MCM(S2^1, H^{op})$. According to the literature [8], similar to (1):

$$\begin{aligned}\frac{MCM(S2^1, H^{op})}{MCM(S^{op}, H^{op})} &= \frac{g_{H^{op}}(S2^1)}{g_{H^{op}}(S^{op})} \\ &\geqslant \frac{g_{H^{op}}(S2^1) - g_{H^{op}}(\varnothing)}{g_{H^{op}}(S^{op}) - g_{H^{op}}(\varnothing)} \\ &\geqslant 1 - (1 - 1/(|V| - M))^{|V|-M}\end{aligned}$$

Then, $F^1 \geqslant max\{MCM(S1^1, H^{op}), MCM(S2^1, H^{op})\}$ So,

$$\begin{aligned}F^1/F^{op} &\geqslant \frac{max\{MCM(S1^1, H^{op}), MCM(S2^1, H^{op})\}}{MCM(S^{op}, H^{op})} \\ &= max\{1-(1-1/M)^M, 1-(1-1/(|V| - M))^{|V|-M}\} \\ &\geqslant 1 - (1 - 2/|V|)^{|V|/2}\end{aligned}$$

□

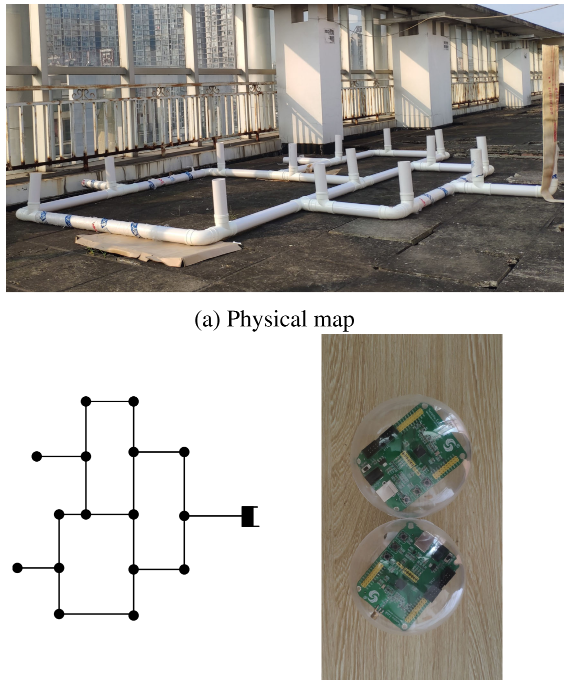
V. EVALUATION

In this section, we evaluate the performance of BGI algorithm. Firstly, we build a small-scale WDN as a testbed to measure the test error rate between theory and experiment. Then, we use the simulation program to test the performance of BGI on a large-scale WDN. We compare BGI with Greedy y [28] to analyze the performance of BGI in formula error rate, approximation ratio, monitoring coverage rate and monitoring delay. Finally, we compare the BGI with X-Y [28] to analyze the performance of BGI in computing time, for the reason that X-Y is faster than Greedy y and Greedy y is better than X-Y in other aspects.

A. TESTBED EVALUATION

As shown in Figure4(a),we assembled a small-scale WDN testbed using PVC pipes. In this figure, each vertical pipe was connected in front of each horizontal pipe junction to facilitate the input of mobile sensors. Limited by the experimental conditions, we adopt a specific WDN structure, so that the flow velocity ratio of each pipe can be simply determined. The structure of WDN is as shown in Figure 4(b), with a total of 20 pipes and 17 junctions. Figure 4(c) shows the mobile sensors we use. Due to the small-scale of testbed, the receiver communication range can cover multiple junctions, which is different from the urban WDNs. Therefore, we only test the test error ratio between the theoretical and actual moving probabilities of the mobile sensor on the testbed. Take the theoretical movement probability as P, and the measured arrival probability is P^*, then the test error rate is $|(P - P^*)/P^*|$.

By repeatedly putting mobile sensor into the testbed and counting the number of times it reaches a specific junction, we can get the probability of mobile sensor reaching the junction. Figure 5 shows two groups of test error rate curves. Among them, abscissa represents the number of experiments, and ordinate represents the test error rate. In the first group, the input position of the mobile sensor is the closest junction to the water source, while in the second group, the input position of the mobile sensor is the adjacent lower junction (in Fig. 4(b)). It can be seen that with the increase of the number of experiments, the test error rate gradually decreases. The two curves in Figure5 are always higher than 0.04, which is caused by many factors, such as unstable water source, unstable center of gravity of mobile sensor and so on. In addition, the first group is longer than the second group in distance between mobile sensor insert junction and target junction. Therefore, the curve of the first group is higher than the second group.



(a) Physical map

(b) Structure chart

(c) Mobile sensors

FIGURE 4. Testbed.

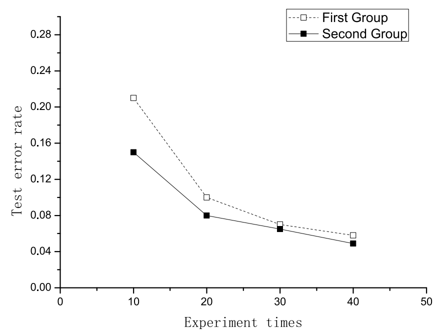
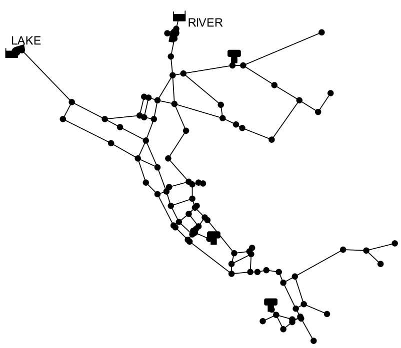


FIGURE 5. The comparison of error rate.

B. SIMULATION PROGRAM EVALUATION

We adapt the WDNs model from EPANET [31], which is a widely used WDN simulator from US Environmental Protection Agency. As dipicted in Figure 6, two networks are used in our simulation. Figure 6(a) is the largest network that EPANET offers, which contains 96 junctions and 117 pipes.



(a) $|V| = 96, |E| = 117$ (b) $|V| = 36, |E| = 40$

FIGURE 6. Network Topology.

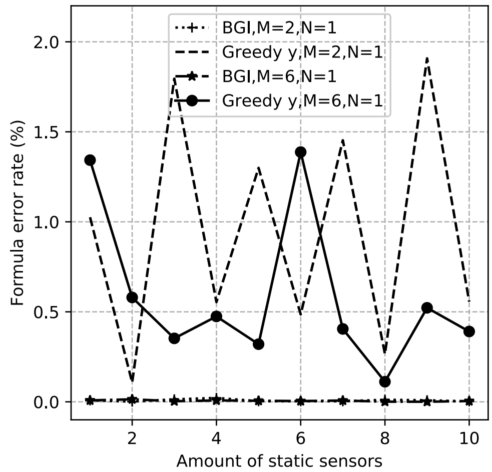


FIGURE 7. The comparison of formula error rate.

Figure 6(b) is relatively smaller that has 36 junctions and 40 pipes. We import the pipe data from EPANET into MATLAB 2016 and calculate the node travelling probability according to Equation 1. The deployment of static nodes is randomly generated for each (M, N). By this mean, after obtaining pipe length, water speed and static sensor deployment, we can calculate coverage rate and expected delay for each mobile sensor and receiver deployment.

In the following sections, we compare the performance of our BGI algorithm with Greedy y in formula error rate, approximation ratio of coverage rate, evaluate monitoring coverage rate, expected monitoring delay and X-Y in computing time. We use the WDN in Figure 6(a) to evaluate the formula error rate and approximation ratio of coverage rate, as the optimal solution can be found with exhaustive enumeration thanks to the small-scale. Then, we use the larger network in Figure 6(a) to evaluate monitoring coverage rate, expected monitoring delay and computing time.

1) FORMULA ERROR RATE EVALUATION

Since the probability models used in our BGI algorithm and Greedy y algorithm are different. For the sake of fairness, we define a Probability Dependence Model (PDM) to calculate expected coverage rate. In this model, the paths of all mobile nodes are simulated to calculate the overall expected coverage rate. The process is repeated many times to find out the mean expected coverage rate. It is obvious that if the repeated times are sufficient, the final result will be close to the actual expected coverage rate. We use R to denote the closed-form expression, and R^* to denote the result of PDM. And we define the formula error rate as $|(R - R^*)/R^*|$.

According to Equation 5 and 6, the upload probability is:

$$R = \sum_{v_j \in S} (t\_p_{i,j} - \sum_{v_g \in S, g<j} m\_p_{i,g}(S)\, t\_p_{g,j})$$

In Greedy y algorithm, the probabilities of arriving at each receiver are independent, then, the upload probability is:

$$R = 1 - \prod_{v_j \in S} (1 - t\_p_{i,j})$$

The formula error rates of BGI and Greedy y under different M and $|V_s|$ are illustrated in Figure 7, in which x-axis shows the number of static nodes and y-axis is the formula error rate. PDM is the mean of 20,000 simulations. Each point is obtained by 50 iterations, and in every iteration, we use the same number of static and mobile sensors with different deployment. It worth noting that the two lines representing BGI are almost coincident.

The results show that the error rates of Greedy y are between 0.1% and 2%, which is larger than that of BGI (within 0.07%). We can conclude that our BGI algorithm outperforms Greedy y in precisely describing the error rate.

2) APPROXIMATION RATIO EVALUATION

We compared the approximation ratio of both algorithm with small-scaled network in Figure 6(b). The optimal solution is obtain by exhaustive enumeration. Approximation ratio is defined as F/F^*, where F is the coverage rate under the compared solutions, and F^* is the coverage rate under optimal solution. The comparison results of both algorithm under different N, M, and V_s are shown in Figure 8, in which x-axis is the number of static nodes and y-axis is the approximation ratio. Every data point is the mean of 50 iterations of experiments, and in each iteration we use the same number of static nodes, but different deployment.

The results show that the approximation ratios of BGI is quite close to 1 and the two lines are almost coincident. In comparison, the approximation ratios of Greedy y are between 0.89 and 0.98. Therefore, our BGI algorithm can more precisely approximate to the optimal solution. In addition, the real approximation ratio is much higher than the theoretical lower bound, which is $1 - (1 - 2/|V|)^{|V|/2} \approx 0.642$.

There are two reasons why the actual approach rate is much higher than the theoretical lower bound. First, the scale of WDN used in the test is small, and it is easy for BGI to obtain an optimal solution. However, considering the huge

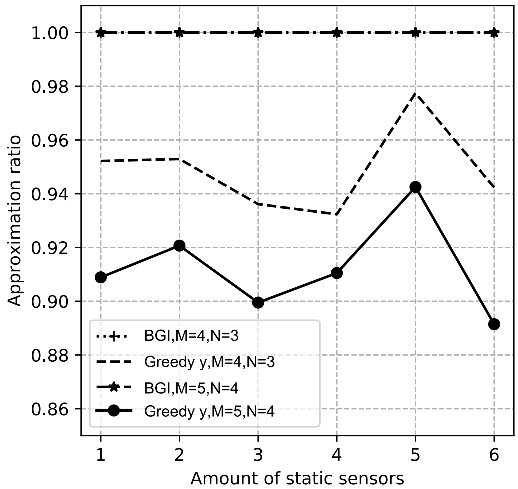


FIGURE 8. **The comparison of approximation ratio.**

amount of computation to enumerate the optimal solution, it is difficult to test on a large-scale WDN. Second, 0.642 is the lower bound of the theoretical lower bound. According to property 8, the theoretical lower bound can be regarded as a valley-shaped curve, and 0.642 is located at the bottom of the curve.

3) MONITORING COVERAGE RATE EVALUATION

We use the network with 96 junctions in Figure 8(a) to evaluation the monitoring coverage rate of BGI and Greedy y. The performance of BGI and Greedy y under different N, M and $|V_s|$ are depicted in Figure 9. Each point is the mean results of 50 iterations, with the same amount of sensor number but different deployment.

Figure 9(a) shows the results of both algorithms with the same number of mobile sensors ($N = 25$) and increasing number of receiver nodes. X-axis in this figure is the number of receiver nodes and y-axis shows the monitoring coverage rate. The dashed line in the figure shows the increase of coverage rate with different number of receivers, ranging from 0.51 to 0.84. While the solid line is the coverage rate of Greedy y, increasing from 0.51 to 0.81, with the increase of number of receivers.

Both curves are concave downward, which matches the monotone description in Theorem 4. In addition, the BGI curve is mostly on top of the Greedy y one, which means our BGI algorithm outperforms Greedy y in the simulation.

Figure 9(b) shows the results of both algorithms with the same number of receiver nodes ($M = 8$) and increasing number of mobile sensors. The dashed line in the figure shows the increase of coverage rate with different number of mobile nodes, ranging from 0.49 to 0.89. In contrast, the solid line is the coverage rate of Greedy y, increasing from 0.46 to 0.87, with the increase of number of mobile nodes.

Both curves are concave downward, which matches the monotone description in Theorem 6. In addition, the BGI curve is mostly on top of the Greedy y one, which means our BGI algorithm outperforms Greedy y in the simulation.

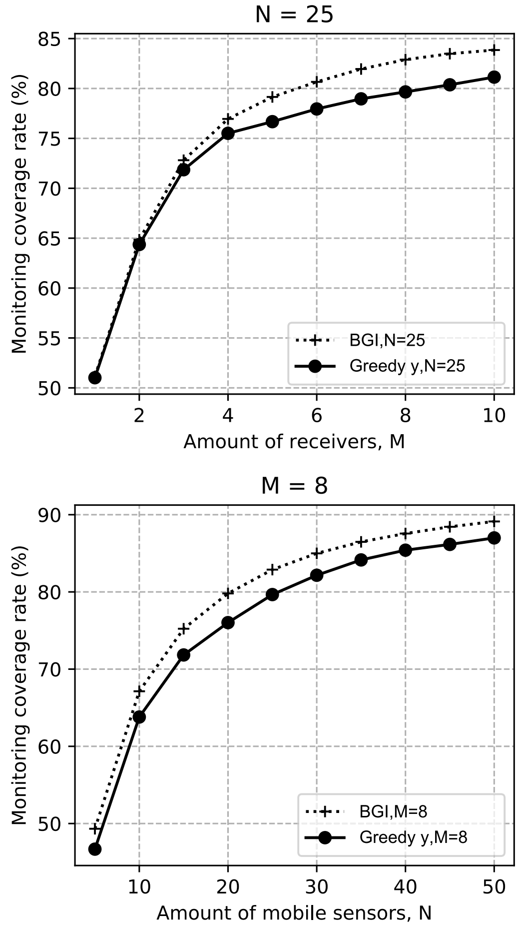


FIGURE 9. **The comparison of monitoring coverage rate.**

4) EXPECTED MONITORING DELAY EVALUATION

To evaluate the expected monitoring delay, we use the same experiments as the monitoring coverage rate ones in the previous section. The performance of BGI and Greedy y under different N, M and $|V_s|$ are depicted in Figure 10.

Figure 10(a) shows the results under same mobile sensors ($N = 25$) but different receivers. The expected monitoring delays of both solutions increase with the number of receivers. For our BGI algorithm, the expected monitoring delay increases from 6,300 to 17,600. While the Greedy y's delay increases from 7,400 to 18,500. The increase of delay is because that the system tends to deploy receivers to the far ends of the WDNs when the amount of receivers grows, and it takes longer for mobile sensors to reach them. Overall, we can see that the BGI curve is under that of the Greedy y, which means BGI has shorter expected monitoring delay than Greedy y.

Figure 10(b) shows the results under same receivers ($M = 8$) but different mobile sensors. The expected monitoring delays of both solutions decrease with the number of mobile sensors. For our BGI algorithm, the expected monitoring delay decreases from 19,600 to 12,600. While the Greedy y's delay decreases from 20,200 to 14,400. The

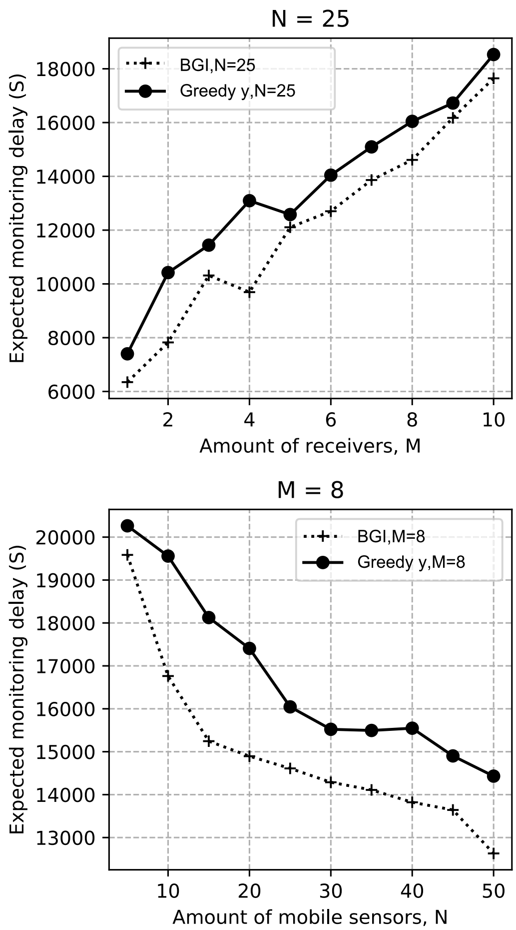
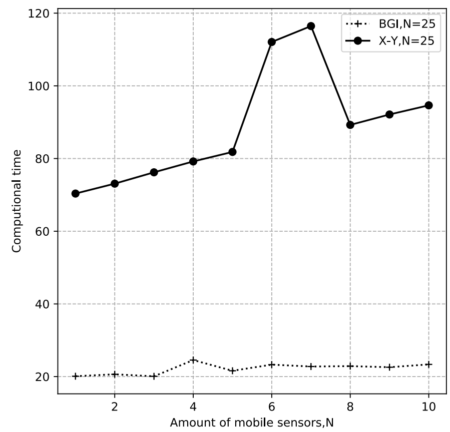
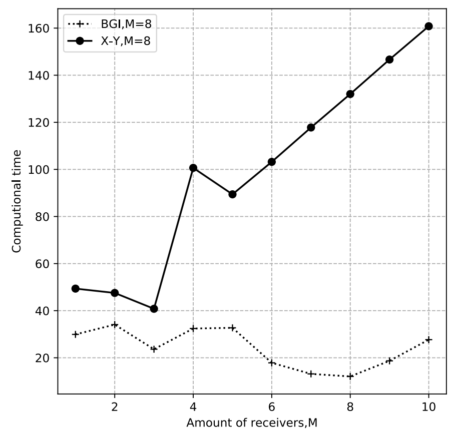


FIGURE 10. **The comparison of expected monitoring delay.**



(a) $M = 8$



(b) $N = 25$

FIGURE 11. **The comparison of computational complexity.**

decrease of delay is because the more mobile sensors we deploy, the closer they are to the receivers. It also can be observed that BGI's curve is under that of Greedy y, which means BGI performs better in term of expected monitoring delay.

5) COMPUTATIONAL COMPLEXITY EVALUATION

To evaluate the computational complexity, we use the same experiments as those for evaluating monitoring coverage rate. The performance of BGI and X-Y under different N and M are dipicted in Figure 11. Among them, the basic unit of computational time is the specified CPU time of MATLAB timing function.

Figure 11(a) shows how the time costs (y-coordinate) grow with the number of receiver nodes (x-coordinate) for both algorithms under the same amount of mobile sensors ($N = 25$). The BGI curve increases from 20 to 23. Differently, the computational cost for X-Y rises from 70 to 94.

The BGI curve is independent from the amount of receivers, because the algorithm just calculates all the coverage rate from 1 to $|ISink|$ receivers.

Figure 11(b) illustrates the time costs of both algorithms under the same amount of receivers ($M = 8$). The time cost of BGI vary between varies between 20 and 24. But the cost for X-Y rises from 49 to 160. The insights for this is similar as Figure 11(a).

It should be noted that the literature [28] points out that the computational complexity of X-Y algorithm is $O((N + M) * (|V|)^2)$. However, this complexity considers O (1) to complete a calculation of monitoring coverage. According to the formulas given in [28], it actually takes $O(M * V^2)$ time. So in fact, the computational complexity of X-Y is $O(M(N + M) * |V|^4)$. It is slightly larger than $O(rNM * |V|^3)$, for $r < |V|$, so the BGI curve is under the X-Y curve.

VI. CONCLUSION

We have addressed the problem of using mobile sensors to monitor urban water pipelines with low delay and high coverage. The problem has been formulated as monitoring coverage rate maximization and expected monitoring delay minimization problem to find the solution for receiver and mobile nodes deployment. The problem has been proven to be NP-hard. Then, we have unified the problem into a forbearing stratified multi-objective optimization model. By utilizing submodular set function property, we have designed greedy algorithms to find the solutions. Finally, extensive simulation and analysis have been conducted to validate the feasibility and effectiveness of our solution.

In real-world applications, the water pressure, speed and even direction could change dynamically. In addition, the pollution could possibly spread with water. Therefore, the dynamic nature of water flow should be considered. In future works, we will go on studying the pipeline monitoring with dynamically changing fluid, and design solutions to be able to track the pollution spreading in time.

REFERENCES

[1] N. B. of Statistics. (2017). *China Statistical Yearbook*. [Online]. Available: http://www.stats.gov.cn/tjsj/ndsj/2017/indexch.htm

[2] X. Bing, "Discussion on improvement of water quality in municipal water distribution network," *Water Wastewater Eng.*, vol. 12, p. 3, Dec. 2002.

[3] A. J. Whittle, L. Girod, A. Preis, M. Allen, H. B. Lim, M. Iqbal, S. Srirangarajan, C. Fu, K. J. Wong, and D. Goldsmith, "WaterWiSe@SG: A testbed for continuous monitoring of the water distribution system in singapore," in *Proc. 12th Annu. Conf. Water Distrib. Syst. Anal.*, 2010, pp. 1362–1378.

[4] A. J. Whittle, M. Allen, A. Preis, and M. Iqbal, "Sensor networks for monitoring and control of water distribution systems," in *Proc. 6th Int. Conf. Struct. Health Monit. Intell. Infrastruct. (SHMII)*, Hong Kong, Dec. 2013.

[5] C. Hu, M. Li, D. Zeng, and S. Guo, "A survey on sensor placement for contamination detection in water distribution systems," *Wireless Netw.*, vol. 24, no. 2, pp. 647–661, 2018.

[6] R. Puust, Z. Kapelan, D. A. Savic, and T. Koppel, "A review of methods for leakage management in pipe networks," *Urban Water J.*, vol. 7, no. 1, pp. 25–45, 2010.

[7] E. Magistretti, J. Kong, U. Lee, M. Gerla, P. Bellavista, and A. Corradi, "A mobile delay-tolerant approach to long-term energy-efficient underwater sensor networking," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, Mar. 2007, pp. 2866–2871.

[8] S. Basagni, L. Bölöni, P. Gjanci, C. A. Petrioli, C. A. Phillips, and D. Turgut, "Maximizing the value of sensed information in underwater wireless sensor networks via an autonomous underwater vehicle," in *Proc. IEEE INFOCOM*, Apr./May 2014, pp. 988–996.

[9] T. T.-T. Lai, W.-J. Chen, Y.-H. T. Chen, P. Huang, and H.-H. Chu, "Mapping hidden water pipelines using a mobile sensor droplet," *ACM Trans. Sensor Netw.*, vol. 9, no. 2, p. 20, 2013.

[10] C. Hu, X. Shu, X. Yan, D. Zeng, W. Gong, and L. Wang, "Inline wireless mobile sensors and fog nodes placement for leakage detection in water distribution systems," in *Software: Practice and Experience*, to be published. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/spe.2631

[11] N. Sankary and A. Ostfeld, "Multiobjective optimization of inline mobile and fixed wireless sensor networks under conditions of demand uncertainty," *J. Water Resour. Planning Manage.*, vol. 144, no. 8, 2018, Art. no. 04018043.

[12] A. Rasekh, R. Wu, W. W. A. W. Salim, and M. K. Banks, "Operation of mobile sensors for monitoring municipal drinking water distribution systems," in *Proc. World Environ. Water Resour. Congr.*, 2014, pp. 362–367.

[13] L. Li-Dong and G. Wen-You, "A heuristic algorithm for packing problem based on forbearing stratified strategy,"*J. Comput. Appl.*, vol. 38, no. 4, pp. 1195–1200, 2018.

[14] G. L. Nemhauser, L. A. Wolsey, and M. L. Fisher, "An analysis of approximations for maximizing submodular set functions—I," *Math. Program.*, vol. 14, no. 1, pp. 265–294, 1978.

[15] T.-T. Lai, Y.-H. Chen, P. Huang, and H.-H. Chu, "Pipeprobe: A mobile sensor droplet for mapping hidden pipeline," in *Proc. 8th ACM Conf. Embedded Networked Sensor Syst.*, 2010, pp. 113–126.

[16] J.-H. Kim, G. Sharma, N. Boudriga, and S. S. Iyengar, "SPAMMS: A sensor-based pipeline autonomous monitoring and maintenance system," in *Proc. 2nd Int. Conf. Commun. Syst. Netw. (COMSNETS)*, Jan. 2010, pp. 1–10.

[17] M. A. Suresh, R. Stoleru, E. M. Zechman, and B. Shihada, "On event detection and localization in acyclic flow networks," *IEEE Trans. Syst., Man, Cybernetics, Syst.*, vol. 43, no. 3, pp. 708–723, May 2013.

[18] M. A. Suresh, R. Stoleru, R. Denton, E. Zechman, and B. Shihada, "Towards optimal event detection and localization in acyclic flow networks," in *Distributed Computing and Networking*, L. Bononi, A. K. Datta, S. Devismes, and A. Misra, Eds. Berlin, Germany: Springer, 2012, pp. 179–196.

[19] M. A. Suresh, L. Smith, A. Rasekh, R. Stoleru, M. K. Banks, and B. Shihada, "Mobile sensor networks for leak and backflow detection in water distribution systems," in *Proc. IEEE 28th Int. Conf. Adv. Inf. Netw. Appl.*, May 2014, pp. 673–680.

[20] N. Oliker and A. Ostfeld, "Inclusion of mobile sensors in water distribution system monitoring operations," *J. Water Resour. Planning Manage.*, vol. 142, no. 1, 2015, Art. no. 04015044.

[21] D. M. Chatzigeorgiou, K. Youcef-Toumi, A. E. Khalifa, and R. Ben-Mansour, "Analysis and design of an in-pipe system for water leak detection," in *Proc. ASME Int. Design Eng. Tech. Conf. Comput. Inf. Eng. Conf.* New York, NY, USA: Amer. Soc. Mech. Eng., 2011, pp. 1007–1016.

[22] X. Li, W. Yu, X. Lin, and S. S. Iyengar, "On optimizing autonomous pipeline inspection," *IEEE Trans. Robot.*, vol. 28, no. 1, pp. 223–233, Feb. 2012.

[23] T. T.-T. Lai, W.-J. Chen, K.-H. Li, P. Huang, and H.-H. Chu, "TriopusNet: Automating wireless sensor network deployment and replacement in pipeline monitoring," in *Proc. ACM/IEEE 11th Int. Conf. Inf. Process. Sensor Netw. (IPSN)*, Apr. 2012, pp. 61–71.

[24] D. Trinchero and R. Stefanelli, "Microwave architectures for wireless mobile monitoring networks inside water distribution conduits," *IEEE Trans. Microw. Theory Techn.*, vol. 57, no. 12, pp. 3298–3306, Dec. 2009.

[25] D. Trinchero, L. Cisoni, R. Stefanelli, A. Abu-Dayya, M. O. Hasna, A. Kadri, and T. M. Khattab, "Mobile wireless sensor networks applied to the survey of water infrastructures," in *Proc. IEEE SENSORS*, Nov. 2010, pp. 690–693.

[26] M. A. Suresh, W. Zhang, W. Gong, R. Stoleru, A. Rasekh, and M. K. Banks, "Toward optimal monitoring of flow-based systems using mobile wireless sensor networks," *ACM Trans. Sensor Netw.*, vol. 11, no. 3, p. 48, 2015.

[27] N. Sankary and A. Ostfeld, "Inline mobile sensors for contaminant early warning enhancement in water distribution systems," *J. Water Resour. Planning Manage.*, vol. 143, no. 2, 2016, Art. no. 04016073.

[28] R. Du, C. Fischione, and M. Xiao, "Flowing with the water: On optimal monitoring of water distribution networks by mobile sensors," in *Proc. 35th Annu. IEEE Int. Conf. Comput. Commun. (IEEE INFOCOM)*, Apr. 2016, pp. 1–9.

[29] T.-W. Kuo, K. C.-J. Lin, and M.-J. Tsai, "Maximizing submodular set function with connectivity constraint: Theory and application to networks," *IEEE/ACM Trans. Netw.*, vol. 23, no. 2, pp. 533–546, Apr. 2015.

[30] K. Oikonomou, M. Parvania, and R. Khatami, "Optimal demand response scheduling for water distribution systems," *IEEE Trans Ind. Informat.*, vol. 14, no. 11, pp. 5112–5122, Nov. 2018.

[31] L. A. Rossman, "Epanet 2 users manual," Water Supply Water Resour. Division, Nat. Risk Manage. Res. Lab., Cincinnati, OH, USA, Tech. Rep. , 2000, vol. 45268, pp. 1–200.

JUNBIN LIANG received the B.E. and M.S. degrees from Guangxi University, Nanning, China, and the Ph.D. degree from Central South University, Changsha, China, in 2000, 2005, and 2010, respectively. He is currently a Professor with Guangxi University, Nanning, China. His research interests include sensor-cloud systems, fog computing, and distributed computing.

XUEYUE XUE was born in Yanchen, Jiangsu, China, in 1992. He received the B.S. degree in computer science and technology from Soochow University, Suzhou, China, in 2014. He is currently pursuing the M.S. degree in computer system architecture with Guangxi University, Nanning, China, where he has been a graduate student with the School of Computer and Electronic Information, since 2016.

His research interests include wireless sensor networks and intelligent services.

ZONGJIAN HE received the B.Eng. and M.Sc. degrees from Tongji University, Shanghai, China, and the Ph.D. degree from the Department of Computing, The Hong Kong Polytechnic University, China, in 2004, 2007, and 2015, respectively.

He is currently a Senior Research IT Specialist with the Centre for eResearch, University of Auckland, Auckland, New Zealand. His research interests include wireless sensor networks, vehicular networks, participatory sensing applications, and mobile computing.

VICTOR C. M. LEUNG (S'75–M'89–SM'97–F'03) received the B.A.Sc. degree (Hons.) in electrical engineering from The University of British Columbia (UBC), in 1977, and the Ph.D. degree in electrical engineering, in 1982.

He attended the graduate school at UBC on a Natural Sciences and Engineering Research Council Postgraduate Scholarship. From 1981 to 1987, he was a Senior Member of Technical Staff with Microtel Pacific Research Ltd. (later renamed MPR Teltech Ltd.), Burnaby, Canada, where he specialized in the planning, design, and analysis of satellite communication systems. He also held a part-time position as a Visiting Assistant Professor with Simon Fraser University, in 1986 and 1987. In 1988, he was a Lecturer with the Department of Electronics, Chinese University of Hong Kong. He returned to UBC as a Faculty Member, in 1989, where he is currently a Professor and the inaugural holder of the TELUS Mobility Industrial Research Chair in advanced telecommunications engineering with the Department of Electrical and Computer Engineering. He is the Director of the Laboratory for Wireless Networks and Mobile Systems (WiNMoS) and a member of the Institute for Computing, Information and Cognitive Systems, UBC. He also holds or has held Guest/Adjunct Professor appointments with Jilin University, Beijing Jiaotong University, South China University of Technology, The Hong Kong Polytechnic University, Beijing University of Posts and Telecommunications, and the Hefei University of Technology. He has authored or coauthored more than 1100 papers in archival journals and international conferences, as well as 40 book chapters. He has also coauthored/coedited 14 books. His research interests are in the broad areas of wireless networks and mobile systems.

Dr. Leung is a Fellow of the Academy of Science, Royal Society of Canada, a Fellow of the Engineering Institute of Canada, a Fellow of the Canadian Academy of Engineering, and a Voting Member of the ACM. He was a recipient of the 2011 UBC Killam Research Prize, the IEEE Vancouver Section Centennial Award, the 2017 Canadian Award for Telecommunications Research, and the 2018 IEEE ComSoc TGCC Distinguished Technical Achievement Recognition Award. He has coauthored papers that won the 2017 IEEE ComSoc Fred W. Ellersick Prize, the 2017 IEEE Systems Journal Best Paper Award, and the 2018 IEEE ComSoc CSIM Best Journal Paper Award. He was awarded the APEBC Gold Medal as the Head of the graduating class with the Faculty of Applied Science. He is serving on the editorial boards of the IEEE TRANSACTIONS ON GREEN COMMUNICATIONS AND NETWORKING, the IEEE TRANSACTIONS ON CLOUD COMPUTING, IEEE ACCESS, the *Computer Communications*, and several other journals. He has previously served on the editorial boards of the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS (Wireless Communications Series and Series on Green Communications and Networking), the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, the IEEE TRANSACTIONS ON COMPUTERS, the IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY, the IEEE WIRELESS COMMUNICATIONS LETTERS, and the *Journal of Communications and Networks*. He has guest-edited many special journal issues, and has served on the Technical Program Committee of numerous international conferences. He has provided leadership to the technical program committees and organization committees of many international conferences. He is a Registered Member of the Engineers and Geoscientists, BC, Canada. He was a Distinguished Lecturer of the IEEE Communications Society, from 2009 to 2012.