RESEARCH   **Open Access**

General optimization framework for surface gateway deployment problem in underwater sensor networks

Saleh Ibrahim, Manal Al-Bzoor[*], Jun Liu, Reda Ammar, Sanguthevar Rajasekaran and Jun-Hong Cui

Abstract

The performance of underwater sensor networks (UWSNs) is greatly limited by the low bandwidth and high propagation delay of acoustic communications. Deploying multiple surface-level radio-capable gateways can enhance UWSN performance metrics, reducing end-to-end delays and distributing traffic loads for energy reduction. In this paper, we study the problem of gateway placement for maximizing the cost-benefit of this UWSN architecture. We develop a mixed integer programming (MIP) gateway deployment optimization framework. We analyze the tradeoff between the number of surface gateways and the expected delay and energy consumption of the surface gateway architecture in the optimal case. We used an MIP solver to solve the developed optimization problem and integrated the optimal results to serve as an input for our simulations to evaluate the benefits of surface gateway optimization framework. We investigated the effect of acoustic channel capacity and the underwater sensor node deployment pattern on our solution. Our results show the significant advantages of surface gateway optimization and provide useful guidelines for real network deployment.

Keywords: Surface gateway, Multiple gateways, UWSN deployment, Deployment optimization

1   Introduction

An important component of oceanographic studies is the collection of data from the aquatic environment. Remote sensing has long been employed as a tool to collect aquatic data in underwater monitoring and exploration activities. Recently, in the last decade to be more specific, underwater acoustic sensor networks (UWSNs) have emerged as a new alternative technology enabling underwater monitoring and exploration applications, including scientific, commercial, and military applications [1-5]. Compared to their remote-sensing counterparts, UWSNs have many advantages. UWSNs can provide localized and more precise data acquisition. They can also employ a wider variety of sensors including, but not limited to, chemical, temperature, photo, and motion sensors.

UWSN technology is also replacing traditional underwater instrumentation technology. Traditionally, bulky sensor nodes equipped with data-storage capability are manually deployed in the underwater target space. Each node operates independently for the duration of the mission to collect readings according to a preset program. At the end of the mission, sensor nodes are picked up, and the collected data are retrieved and processed. UWSN technology adds networking capabilities to underwater sensor nodes so that sensor nodes can relay real-time data to an off-shore or even an on-shore control station for immediate analysis. The communication channel can also be used to transmit control signals from the control station to the underwater sensor nodes, which enables interactive control of the underwater sensor network deployment. UWSNs offer many advantages over traditional instrumentation techniques. First off, UWSNs add a real-time reporting functionality that enables a host of new real-time monitoring and warning systems. Another advantage of UWSN is that the sensing mission can be dynamically reconfigured without the need to physically access all the underwater nodes in order to reprogram them. While this particular feature makes the reuse of a UWSN deployment much less costly, it also provides for fixing configuration errors that compensates for unforeseen circumstances

*Correspondence: mbzoor@engr.uconn.edu
Department of Computer Science and Engineering, University of Connecticut, Storrs, CT 06269-2155, USA

and unexpected node failures. This improves the UWSN resilience compared to traditional instrumentation techniques. Therefore, sensor node failures can be detected soon after they occur, allowing early replacement or early abortion of the mission instead of having to wait until the end of the mission only to find that it has failed.

In addition to the usual design challenges faced by terrestrial wireless sensor networks, UWSN technology has to deal with some unique challenges. It cannot use electromagnetic waves for long-range communication due to their quick absorption in water. Acoustic waves are usually considered the practical solution for UWSN communication. The dependency of UWSNs on underwater acoustic communications is particularly challenging. Factors such as the high levels of noise and the channel variability due to temperature, pressure, salinity gradients, and current-induced turbulence add more constraints to the already small bandwidth available for acoustic communication. Moreover, when the Doppler effect (due to mobility) is added to those factors, channel encoding becomes a crucial component to the success of underwater acoustic sensor networks. However, the most limiting factor of underwater acoustic communications is the extremely low propagation speed of sound, which is roughly 1.5 km/s, subject to slight changes due to pressure, temperature, and salinity variations [6]. This is five orders of magnitude slower than the propagation speed of electromagnetic waves. Such high propagation delay can cause high end-to-end delay, which could be greatly limiting for interactive applications and other monitoring applications where response time is critical.

One way to mitigate the high propagation delay in acoustic communications is to deploy multiple surface-level gateways. In sensor networks, each sensor node can monitor and detect environmental events locally and then transfer these measurements through the network to a surface gateway node (also referred to as a sink in UWSNs), which then relays data to the control station. Unlike single-sink networks that use long underwater paths to reach the unique surface sink, in a multiple-sink underwater sensor network, as illustrated in Figure 1, underwater sensor nodes can send data packets toward their nearest surface gateway. A surface gateway then uses electromagnetic waves to forward the packets to the control station. Considering that electromagnetic wave propagation is in orders of magnitude faster than acoustic wave propagation, it is safe to assume that surface gateways can send packets to the control station in negligible time and with relatively small energy consumption since acoustic communications consume much more energy than radio communications [5]. In this way, all the surface gateways (or sinks) form a virtual sink. Although architectures employing multiple surface gateway nodes were mentioned in [5,7], there is no formal study on surface gateway deployment. Neither an analysis on the effect of using multiple surface gateways on the network energy consumption or delay characteristics has been conducted nor a guideline on the deployment of such a multi-sink architecture has been provided.

In this paper, we study the problem of surface gateway deployment and present guidelines for deciding the number and locations of surface gateway nodes given an underwater sensor network deployment scenario. We focus on optimizing the cost of surface gateway deployment, by finding the minimum number and the locations of surface gateway nodes required to achieve a given design objective, which can be communication delay, energy consumption, fault tolerance, or a combination of them. The surface gateway deployment problem is formulated as an optimization problem modeling the routing of data packets from underwater sensor nodes to the virtual sink under link capacity and flow conservation constraints. A variety of objective functions are presented. Our framework provides an optimal gateway selection from given gateway candidate locations which are assumed to be a given. The rest of this paper is organized as follows. In Section 2, we provide a review of related work. In Section 3, the network model and assumptions regarding the surface gateway deployment problem are presented and justified, and the surface gateway deployment problem is formulated as an optimization problem. In Section 4, we evaluate our work, choosing sample problems to analyze the effect of various constraints on the deployment solution quality, problem complexity, and feasibility. Finally, in Section 5, conclusions are drawn, and a future work is presented.

2 Related work

The deployment problem for sensor nodes was studied extensively for terrestrial sensor networks [8]. The closest to our work from the placement strategies in terrestrial networks are those for relay node and multi-tier sensor network architecture [9-11]. However, most terrestrial deployment problems assume a static two-dimensional (2D) architecture, and not much attention was received for multi-gateway deployment in UWSN. A triangular-grid deployment pattern for 2D UWSN was proposed in [12]. The objective is to minimize the number of sensors needed to achieve the sensing and communication coverage of a target area. An interesting attempt to formulate the 3D UWSN point-coverage deployment problem as an integer linear program (ILP) is presented in [13]. The solution of the ILP decides relay node deployment, routing, and link scheduling. Throughout this work, it is assumed that there is a single sink for the entire UWSN deployment and only the overall power consumption is used as an optimization objective. Neither of these two studies considers the multiple-sink network architecture.
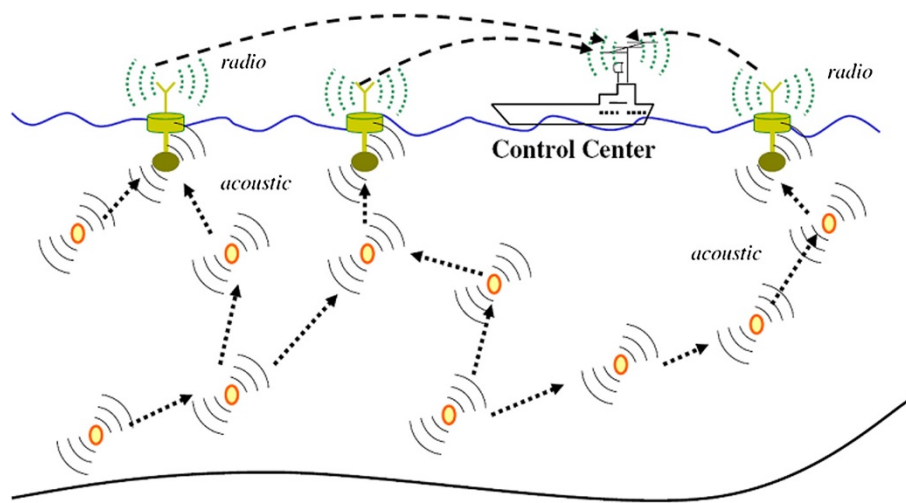
Figure 1 Using multiple surface gateways.

The only research study in the frame of multiple sinks we found is [7], in which Seah and Tan investigated the use of multi-sink architecture to enhance the underwater sensor network reliability. In this study, the same message is directed to more than one of the multiple sinks, with the assumption that if any of the sinks gets the message, then it is considered delivered successfully. The simulation results showed that high-reliability benefits can be achieved at the cost of reasonable increase in energy consumption. The surface gateway (i.e., sink) deployment problem was not considered in this work. In a parallel research effort, [14] studied the problem of placing multiple mobile data collectors in both delay-tolerant and delay-constrained underwater acoustic sensor networks. The authors defined candidate data collection stations as the maximal overlapping regions (MORs) of surface circles corresponding to underwater node communication ranges. They developed an $O\left(n^2 \log n\right)$ algorithm for finding MORs. An earlier work in [15] was the first to address the underwater surface gateway deployment problem and formulate it as an optimization problem. The problem of surface-level gateway placement has been addressed by later research effort in[16]. The authors used surface gateways deployment as a mean to guarantee connectivity and survivability (tolerance to single node failure). They proposed an approximation algorithm for choosing a minimal subset of candidate locations where SGs may be deployed. An effort by the authors in [17] addressed the deployment for a mobile multiple-sink architecture in UWSN. They used a prediction-based deployment strategy to cater for the mobility of underwater nodes. However, our work differs from all above by formulating the problem to find the best candidate locations that satisfies a set of flow conservation constraints, interference constraints, number of gateway constraints, and performance constraints

in addition to a set of delay and energy-consumption objectives.

3 Gateway deployment optimization generalized formulation

There are two approaches to handle the surface gateway deployment problem, (1) solving the underwater deployment and the surface-level deployment problems jointly or (2) solving each of them separately. It is understood that solving both underwater and surface deployment problems jointly will lead to optimal solutions that are better, or at least as good as, the outcome of the two-phase approach. However, since the objective of the research presented here is to analyze the effect of surface gateway deployment on the overall underwater sensor network performance, we fix the underwater deployment and therefore opt for the latter option. Thus, we assume that there is a pre-existing underwater deployment that has been reached by a way or another.

The surface gateway deployment problem is formulated as a combinatorial graph optimization problem. The nodes of the graph represent underwater sensors and candidate surface gateway positions, and the problem is to find the subset of the candidate surface locations that maximizes a certain performance metric, satisfying a set of flow conservation constraints, interference constraints, and a deployment cost constraint (on the number of surface gateways) or a performance constraint (such as maximum end-to-end delay or minimum reliability level).

The selection of the candidate positions is sophisticated enough to be considered as a separate problem to on its own and is discussed by our work in [18]. For the purpose of producing a generalized formulation, the set of candidate surface points is considered a given and assumed

to satisfy connectivity constraints as a precondition. This means that each underwater node has to have at least one connected path to one or more candidate surface locations, taking into account the communication ranges of the involved intermediate nodes.

Associated with each underwater sensor node is a packet generation rate. Surface gateway nodes have to collect all generated data packets. The surface gateway deployment problem is formulated as a combinatorial optimization problem. The formulation consists of a basic set of constraints that can be augmented with a variety of objective functions.

3.1 Assumptions

The functionality of the UWSN considered is assumed to be mainly collecting data from underwater sensor nodes and transmitting the collected data samples at regular time intervals to a central station through one of the surface-level gateways. We assume that most of the traffic, therefore, will be flowing from the underwater nodes to the surface gateways. Inter-node communication (for purposes such as synchronization, collaborative sensing, data fusion, etc.) is assumed to be small enough to ignore.

Together, the set of surface-level gateways forms a *virtual sink* for the underwater sensors because the propagation delay, the energy consumption, and the reliability of transmitting the received packet by a gateway to the central station over a direct or multi-hop radio path are far superior to underwater communication links. It is reasonable, therefore, to assume that a packet delivered to one of the surface gateways is delivered to the central station with high reliability and negligible delay and energy cost.

For simplicity, we assume that the data link protocol uses only fixed-length packets and that all links have the same bit rate. Consequently, the packet transmission time is consistent throughout the network, and if the transmission scheme is slotted, the packet transmission time is conventionally called the timeslot. Our deployment formulation can be adjusted to include a variable packet length and a variable bit rate at different links without compromising the quality of the solution. This will require a pre-existing knowledge of these parameters which will increase number of inputs and will result in a more complex formulation of the flow constraints. However, we are interested in providing a more generalized formulation of the gateway optimization problem to be adjusted and tuned for a more specified application that may require varying some of the inputs we assume is a fixed.

3.2 Definitions

The network is modeled as a graph, in which nodes represent the underwater sensors and surface gateways, and edges represent pair-wise communication links.

3.2.1 Nodes

Let $\mathcal{U}$ be the set of all underwater sensor nodes and $\mathcal{T}$ be the set of candidate surface node positions.

Let $\mathcal{V}$ be the set of all nodes, i.e.,

$$\mathcal{V} = \mathcal{U} \cup \mathcal{T}. \tag{1}$$

Let $\mathcal{C}_v$ be the set of nodes within the communication range of an underwater node u, i.e.,

$$\mathcal{C}_u = \{v : v \in \mathcal{V}, v \neq u, d(u, v) \leq R_u^C\}, \ \forall u \in \mathcal{U}, \tag{2}$$

where $d(v, w)$ denotes the Euclidean distance between the two nodes u and v, and R_u denotes the maximum acoustic communication range of the underwater node u.

3.2.2 Edges

Let the set of edges $\mathcal{E}$ be the set of all possible communication links, i.e.,

$$\mathcal{E} = \{e(u, v) : u \in \mathcal{U}, v \in \mathcal{C}_u\}. \tag{3}$$

Let $\mathcal{E}_u^O$ denote the set of outgoing links of an underwater node u, i.e.,

$$\mathcal{E}_u^O = \{e(u, v) : v \in \mathcal{C}_u\}, \ \forall u \in \mathcal{U}. \tag{4}$$

Since surface gateways do not transmit data packets on their underwater acoustic interface,

$$\mathcal{E}_t^O = \phi, \ \forall t \in \mathcal{T} \tag{5}$$

and $\mathcal{E}_v^I$ denote the set of ingoing links to a node v, i.e.,

$$\mathcal{E}_v^I = \{e(u, v) : e(u, v) \in \mathcal{E}\}, \ \forall v \in V. \tag{6}$$

3.2.3 Data generation and link flow rates

Let τ be the packet transmission time, called the timeslot in slot. Let g_u be the average packet generation rate at node $u \in \mathcal{U}$, i.e., the expected number of generated packets during the packet transmission time τ, and let G be the total data generation rate of the entire network, which should be equal to the average packet arrival rate at the virtual sink, i.e., the number of packets expected to arrive at the virtual sink during the packet transmission time τ,

$$G = \sum_{u \in \mathcal{U}} g_u. \tag{7}$$

Let f_e be the average flow per packet time in edge e measured in packets per packet time.

Let f_v^I be the average total flow into node v per packet time, i.e.,

$$f_v^I = \sum_{e \in \mathcal{E}_v^I} f_e, \ \forall v \in \mathcal{V}. \tag{8}$$

Let f_u^O be the average total flow out of node u per packet time

$$f_u^O = \sum_{e \in \mathcal{E}_u^O} f_e, \ \forall u \in \mathcal{U}. \tag{9}$$

For gateways, there is no underwater outgoing flow, and therefore,

$$f_t^O = 0, \ \forall t \in \mathcal{T}. \tag{10}$$

3.2.4 Gateway presence indicator
Let x_t be a binary variable that indicates whether a surface gateway is to be deployed at candidate location t, i.e.,

$$x_t = \begin{cases} 1 \text{ if a node deployed at } t, \\ 0 \text{ otherwise} \end{cases}, \quad \forall t \in \mathcal{T}. \tag{11}$$

3.2.5 Link scheduling
Let T be the schedule length, i.e., the number of time slots in a single period of the schedule.

Let $h_{e,t}$ be a binary variable that indicates whether link e is scheduled for transmission during slot t of the schedule.

$$h_{e(u,v),t} = \begin{cases} 1 & u \text{ transmits to } v \text{ at timeslot } t \\ 0, & \text{otherwise} \end{cases},$$

$$\forall u \in \mathcal{U}, e \in \mathcal{E}, 0 \leq t < T. \tag{12}$$

3.2.6 Performance parameters
The most important performance aspects of any network are the delay and energy-consumption characteristics.

Delay Let υ be the average propagation velocity of sound waves in water.

Let $\tau_{u,v}^P$ be the propagation delay from node u to node v. Thus,

$$\tau_{u,v}^P = \frac{d(u,v)}{\upsilon}, \ \forall u \in \mathcal{U}, v \in \mathcal{V}. \tag{13}$$

Let τ_u^Q be the average transmission queuing time including the expected channel access delay at node u.

Therefore, the total packet transmission delay, $\tau_{u,v}$, over link e (u,v) is

$$\tau_{u,v} = \tau_u^Q + \tau + \tau_{u,v}^P, \ \forall u \in \mathcal{U}, e \in \mathcal{E}_u^O. \tag{14}$$

Energy consumption Let π_e^S be the transmission energy required for transmitting one data packet over the underwater acoustic link e (u,v), π^L be the listening/sleeping average energy consumption per packet time, and π_v^R be the reception energy per packet. For surface gateways, the reception power is taken to include the energy required to forward a packet to the central station over radio.

Therefore, the total power consumption π_v of node v per packet time is

$$\pi_v = \pi^L + \pi_v^R f_v^I + \sum_{e \in \mathcal{E}_v^O} \pi_e^S f_e, \ \forall v \in \mathcal{V}. \tag{15}$$

Note that when the underwater sensor nodes use only one transmission power level π^S, the total energy-consumption formula reduces to

$$\pi_v = \pi^L + \pi_v^R f_v^I + \pi^S f_v^O, \ \forall v \in \mathcal{V}. \tag{16}$$

Due to the nature of the shared communication medium, contention for channel access can occur, and some transmission attempts may end up in failure. Transmission failures can also be caused by uncorrectable transmission errors. Let us assume that the transmission success probability on link e is a constant ρ_e that depends on the channel utilization. This means that on the average, a successful transmission will occur every $\dfrac{1}{\rho_e}$ attempts. The average number of retransmissions for link e is hence n_e such that

$$n_e = \left(\frac{1 - \rho_e}{\rho_e} \right), \ \forall e \in \mathcal{E}. \tag{17}$$

We assume that, on the average, every failed transmission attempt will cost the sender $\pi_e^{S'}$ and will cost the receiver $\pi_v^{R'}$. Including the energy cost of retransmissions, we get the following generalized formula for total power consumption:

$$\pi_v = \pi^L + \sum_{e \in \mathcal{E}_v^I} \left[\left(\pi_v^R + n_e \pi_v^{R'} \right) f_e \right]$$
$$+ \sum_{e \in \mathcal{E}_v^O} \left[\left(\pi_e^S + n_e \pi_e^{S'} \right) f_e \right], \ \forall v \in \mathcal{V}. \tag{18}$$

Using any of above Equations 15, 16, and 18, the average energy consumed collectively by the network for a single transmission over link e (u,v) can be written as follows:

$$\pi_{e(u,v)} = \frac{\pi^L}{f_e} + \left(\pi_e^S + \pi_v^R + n_e \left(\pi_e^{S'} + \pi_v^{R'} \right) \right). \tag{19}$$

3.3 Constraints
The constraints can be classified into deployment constraints, flow conservation constraints, and interference constraints.

3.3.1 Deployment constraints
- Number of surface gateways

If the objective of the optimization is to minimize delay or energy consumption using a limited number of surface nodes, N, the following constraint can be used to limit the number of gateways deployed to at most N:

$$\sum_{t \in \mathcal{T}} x_t \leq N. \tag{20}$$

- Gateway presence indicator constraints

No more than one gateway can be deployed at any candidate location:

$$x_t \in \{0, 1\}, \quad \forall t \in \mathcal{T}. \tag{21}$$

3.3.2 Flow constraints
The scenario considered in this work will be a monitoring network, where instrumentation data flow from

underwater sensor nodes and through the network to a common sink station. Therefore, the analysis is limited to the (possibly multi-path) route from each underwater sensor to the virtual sink. Control traffic flowing from the central station down to the underwater sensor nodes or inter-node traffic for the purpose of synchronization or localization or other functions than sensing data transfer will be ignored.

- Gateway presence constraints

Data can only be received at locations where surface gateways are deployed. This constraint can be written as follows:

$$f_t^I \le x_t G, \quad \forall t \in \mathcal{T}. \tag{22}$$

This means that the total flow f_t^I into candidate gateway location t has to be zero if $x_t = 0$; otherwise, f_t^I can grow as large as the maximum potential flow in any link in the network which is equal to the total data generation rate G in the network, thus rendering the constraint void.

- Per-node flow conservation constraints

Flow conservation implies that for underwater sensor nodes, the sum of the average flows leaving a node equals the sum of the flows entering that node plus the local data generation rate.

$$f_u^O - f_u^I = g_u, \quad \forall u \in \mathcal{U}. \tag{23}$$

- End-to-end flow conservation constraints

Flow conservation also implies that the total data generation rate of all underwater nodes must equal the total data absorption rate by the virtual sink composed of all surface gateways since each packet, generated by any source, must eventually be received by a surface node (before being relayed to the sink).

$$\sum_{t \in \mathcal{T}} f_t^I = G. \tag{24}$$

- Flow allocation constraints

The average flow in any edge e during a schedule period cannot exceed the total number of time slots in which the link is active

$$T \cdot f_e \le \sum_{t=1}^{T} \left(h_{e,t} \right), \ \forall e \in \mathcal{E}. \tag{25}$$

Note that since we assume the schedule is periodic, there is no need to add a constraint to enforce the reception of a packet before sending it out. If a packet is sent before being received, this means that it was stored from the previous period. It follows that the maximum per-hop queuing delay will not exceed $T - 1$ timeslots.

3.3.3 Interference constraints

Because they use a shared medium to communicate, each node will be able to utilize only a fraction of the underwater channel bandwidth. The exact formulation of interference constraints will depend heavily on the medium access control protocol being used by the UWSN. Without knowing the specifics of the medium access protocols, we can assume a randomized protocol with either one of two medium access schemes, namely slotted vs. unslotted schemes. It has been shown in literature that due to the relatively high propagation delay of acoustic signals, slotted protocols behave as inefficiently as unslotted protocols do [19-22]. Therefore, we can assume that the channel utilization cannot exceed the theoretical maximum of approximately 0.18. Let b be the channel bit rate. We define the effective bandwidth B as the throughput at the maximum utilization. Hence,

$$B \le 0.18b. \tag{26}$$

To avoid collision at the receiver, a node is able to receive successfully only if (1) it is not transmitting and (2) all nodes that are within its interference range are silent for the duration of the reception. To formulate this constraint, we first define the interfering node set $\mathcal{I}_v$ for receiver v as the set of nodes whose transmissions can potentially collide with data packets sent to receiver v. Thus,

$$\mathcal{I}_v = \left\{ u : u \in \mathcal{U}, d\left(u, v\right) \le R_u^I \right\}, \ \forall v \in \mathcal{V}. \tag{27}$$

A pessimistic formulation of the effect of contention on constraining the available bandwidth for each receiver v can be written as follows:

$$f_v^O + \sum_{u \in \mathcal{I}_v} f_u^O \le B, \ \forall v \in \mathcal{V}. \tag{28}$$

This formulation is called pessimistic because that assumes that transmissions that could interfere with the reception at node v are going to happen in distinct time intervals without any overlapping. This is the worst case because it leaves the minimum possible bandwidth for v to receive its own intended transmissions. In reality, however, some of these transmissions from interfering neighbors of node v can occur simultaneously and successfully. Consider, for instance, two interfering neighbors of v, called u_1 and u_2, v cannot receive any signals as long as it can hear either u_1 or u_2 is transmitting. Now, if u_1 and u_2 are sending to w_1 and w_2, respectively, it is possible that both transmissions can succeed simultaneously if u_1 is not within w_2's interference range and u_2 is not within w_1's interference range.

Because surface nodes do not transmit data packets underwater, the interference constraint formula for surface gateways reduces to

$$\sum_{u \in \mathcal{I}_t} f_u^O \le B, \ \forall t \in \mathcal{T}. \tag{29}$$

For each underwater node in the interfering node set of a certain receiver, we define the *vulnerable slots* of a link $e = (u, v)$ with respect to a transmitter w as follows:

$$\mathcal{P}_{u,v,w} = \left\{ k : k \in \mathbb{Z}, \lfloor \frac{\delta_{u,v,w}}{\tau} \rfloor \le k \le \lceil \frac{\delta_{u,v,w}}{\tau} \rceil \right\}.$$

$$\delta_{u,v,w} = \tau_{w,v}^P - \tau_{u,v}^P, \quad \forall (u, v) \in \mathcal{E}, w \in \mathcal{I}_v \quad (30)$$

This means that node u cannot send to node v during slot t if node w is transmitting during time slots $t' = (t - i)$ modulo T, $\forall i \in \mathcal{P}_{u,v,w}$. Note that $t - i$ could be negative, which means that there is potential interference with packets w sends during earlier schedule periods. Therefore, we write the interference avoidance constraint as follows:

$$h_{(u,v),t} \le 1 - h_{(w,z),t'}, \, t' = (t - i) \text{ modulo } T, \, \forall (u, v) \in E, \, w \in \mathcal{I}_v.$$
$$(w, z) \in \mathcal{E}_w^O, i \in \mathcal{P}_{u,v,w}, \quad t \in \{1, \ldots, T\} \quad (31)$$

In addition, we have to prevent self-interference by enforcing the following constraint:

$$\sum_{e \in \mathcal{E}_v^O} h_{e,t} \le 1, \quad \forall v \in \mathcal{U}, t \in \{1, \ldots, T\}. \quad (32)$$

This means that a node cannot transmit on more than one link during the same time slot.

T is the number of time slots a packet spends waiting in the transmission queuing at any node. Since the queuing delay can vary from 0 up to $T - 1$ time slots, we estimate $L = \frac{T-1}{2}$. In order to for this estimate to be accurate and in order to insure the best possible performance, a search for the minimum feasible schedule length will be necessary. We do so following the method pointed out in [13]. Namely, we formulate the problem starting from $T = 2$ and attempt to solve the MIP problem. As long as the MIP is infeasible, we continue to increase the schedule length T until we find the minimum T for which the ILP formulation is feasible.

3.4 Objective functions

The objective function determines the goal of the deployment optimization. In general, the objective can be (1) a collective measure, such as minimizing the average end-to-end delay and minimizing the total energy-consumption rate or (2) an extreme measure, such as minimizing the worst case end-to-end delay or maximizing the worst case node lifetime.

3.4.1 Minimizing expected end-to-end delay

The objective here is to minimize the expected end-to-end delay for all packets. The end-to-end delay for a packet is the sum of the per-hop delay over the entire path from the source that generates the packet to virtual sink (i.e., the gateway) that receives it. As expressed earlier in

Equation 14, the per-hop delay consists of three components: queuing and channel access delays, transmission time, and propagation delay. Let $K_{u,t}$ be the set of all active paths from source underwater node u to surface gateway candidate location t. Let $k_i = (e_1, e_2, \cdots, e_m)$ be one of the paths in $K_{u,t}$, where $e_1, e_2, \cdots, e_m$ are links that constitute the path k_i. The end-to-end delay $\tau_{k_i}^E$ of path k_i can be written as

$$\tau_{k_i}^E = \sum_{e \in k_i} \tau_e, \quad \forall u \in \mathcal{U}, t \in \mathcal{T}, k_i \in K_{u,t}. \quad (33)$$

Let f_{k_i} be the rate of data flow from u to t through path k_i. It follows that the average end-to-end delay between u and t is

$$\mathrm{E}\left[\tau_{u,t}^E\right] = \frac{\sum_{k_i \in K_{u,t}} f_{k_i} \tau_{k_i}^E}{\sum_{k_i \in K_{u,t}} f_{k_{u,t}^i}}. \quad (34)$$

The flow in each link in terms of the path flows can be written as follows:

$$f_e = \sum_{u \in \mathcal{U}} \sum_{t \in \mathcal{T}} \sum_{k_i \in K_{u,t}:e \in k_i} f_{k_i}. \quad (35)$$

Let δ_{e,k_i} be a set of decision variables that indicate whether link e is part of path k, i.e.,

$$\delta_{e,k_i} = \begin{cases} 1, & e \in k_i \\ 0, & \text{otherwise} \end{cases}. \quad (36)$$

Using Equation 36, Equation 35 can be rewritten as follows:

$$f_e = \sum_{u \in \mathcal{U}} \sum_{t \in \mathcal{T}} \sum_{k_i \in K_{u,t}} \delta_{e,k_i} f_{k_i}. \quad (37)$$

The average end-to-end delay between u and the virtual sink (all surface gateway candidate location) is

$$\mathrm{E}\left[\tau_u^E\right] = \frac{\sum_{t \in \mathcal{T}} \sum_{k_i \in K_{u,t}} f_{k_i} \tau_{k_i}^E}{g_u}, \, \forall u \in \mathcal{U}, \quad (38)$$

where $g_u = \sum_{t \in \mathcal{T}} \sum_{k_i \in K_{u,t}} f_{k^i}$.

Now, the overall expected end-to-end delay can then be written as:

$$\mathrm{E}\left[\tau^E\right] = \frac{\sum_{u \in \mathcal{U}} \sum_{t \in \mathcal{T}} \sum_{k_i \in K_{u,t}} f_{k_i} \tau_{k_i}^E}{\sum_{u \in \mathcal{U}} g_u} \quad (39)$$

Using Equations 35, 36, 33, and 7, the overall expected end-to-end delay can be rewritten as follows:

$$\mathrm{E}\left[\tau^E\right] = \frac{\sum_{u \in \mathcal{U}} \sum_{t \in \mathcal{T}} \sum_{k_i \in K_{u,t}} f_{k_i} \sum_{e \in k_i} \tau_e}{G} = \frac{1}{G} \sum_{e \in \mathcal{E}} (f_e.\tau_e). \quad (40)$$

Also, the corresponding objective function will be

$$\text{Minimize} \left\{ \frac{1}{G} \sum_{e \in \mathcal{E}} (f_e.\tau_e) \right\}. \quad (41)$$

3.4.2 Minimizing expected energy consumption

The objective is to minimize the expected end-to-end energy consumption, i.e., the energy consumed in the network for a packet to travel from its source to a sink. The average energy, π_e, consumed collectively by the network in order to transmit one packet on link e was formulated in Eqaution 19. The expected end-to-end per packet energy consumption for a path $k_i \in K_{u,t}$ between source node u and sink t can be written as

$$\pi_{k_i}^E = \sum_{e \in k_i} \pi_e, \quad \forall u \in \mathcal{U}, t \in \mathcal{T}, k_i \in K_{u,t}. \tag{42}$$

Following a derivation similar to that given in subsection 3.4.1, the overall expected end-to-end per packet energy consumption can be shown to be

$$\mathrm{E}[\pi^E] = \frac{1}{G}\left(\sum_{e \in E} f_e . \pi_e\right). \tag{43}$$

Also, the corresponding objective function will be

$$\text{Minimize} \left\{ \frac{1}{G}\left(\sum_{e \in E} f_e . \pi_e\right) \right\}. \tag{44}$$

3.4.3 Minimizing worst case per-source average delay

When the optimization objective is to minimize the overall average end-to-end delay, some sources maybe excessively penalized with a much longer delay. In monitoring applications where uniform response time is favored, the optimization objective has to be minimizing the worst case, per-source average delay. In other words, the objective is to minimize the average delay observed by packets from each source taken separately, to guarantee the best possible worst case scenario. In order to define this objective, we define a set of new flow variables, $f_{e,u}$ to denote the portion of the flow in link e that is generated originally by node u. If no path from u to the virtual sink uses link e, this implies that $f_{e,u} = 0$.

The flow conservation constraints can be rewritten to be source specific as follows:

$$\sum_{e \in \mathcal{E}_v^O} f_{e,u} - \sum_{e \in \mathcal{E}_v^I} f_{e,u} = 0, \ \forall u \in \mathcal{U}, v \in \mathcal{V}, u \neq v$$

$$\sum_{e \in \mathcal{E}_u^O} f_{e,u} = g_u, \ \forall u \in \mathcal{U}$$

$$f_{e,u} = 0, \ \forall u \in \mathcal{U}, e \in \mathcal{E}_u^I \tag{45}$$

$$\sum_{t \in \mathcal{T}} \sum_{e \in \mathcal{E}_t^I} f_{e,u} = g_u, \ \forall u \in \mathcal{U}$$

The end-to-end per-source average delay τ_u^E can be written as follows:

$$\tau_u^E = \frac{1}{g_u} \sum_{e \in \mathcal{E}} f_{e,u} \tau_e, \ \forall u \in \mathcal{U}. \tag{46}$$

We define an upper bound U_τ for the per-source average delay,

$$\tau_u^E \leq U_\tau. \tag{47}$$

Also, the optimization objective is to minimize this upper bound,

$$\text{Minimize} \{U_\tau\}. \tag{48}$$

4 Performance evaluation

We conducted extensive simulation to evaluate our gateway deployment optimization framework. We assumed that acoustic transceivers of all nodes, both underwater nodes and surface gateways, to be homogeneous and, therefore, the communication range is assumed to be constant for all nodes. We assume that sensor nodes are either stationary or that their motion is correlated strongly enough to assume that their relative locations are fixed. We adopt the pessimistic interference model. To reduce the problem complexity, we assume a lightly loaded, i.e., such that channel access delays and queuing delays can be safely ignored. Since both channel access delay and queuing delay are functions of network load (i.e., flow), keeping the network very lightly loaded justifies the assumption that the queuing delay is constant, hence allowing the linearization of the formulation and the use of LP solvers. When the load is increased, the non-linear formulation can be solved similarly by piece-wise linearization algorithms, such as the Frank-Wolfe algorithm. Although the problem can be solved for a choice of optimization goals, we limit our focus on the simplest optimization goals, namely minimizing the average delay and minimizing the average power consumption. The LP solver uses a brute-force search algorithm to find the optimal gateway deployment locations by numerating all possible candidate solutions and checking whether each satisfies the problem's statement.

4.1 Case study

In order to evaluate the benefits and the performance of the gateway deployment optimization techniques, we used the following simulation setting.

Throughout the experiments, we fixed the packet length $L = 400$ bits, the underwater acoustic propagation velocity $\upsilon = 1.5$ km/s, and the transmission power is set to a constant of 1 W/s per packet time. The communication range for the underwater modems for all nodes is fixed at $R^C = 150$. We also fixed the area of deployment to a square area of 600 m $\times$ 600 m horizontal extent and fixed the candidate gateway deployment positions to a 5×5 square mesh of points spaced 150 m apart .

The depth of all underwater sensors is arbitrarily set to 100 m, such that each of the underwater sensors,

Figure 2 Uniform deployment.

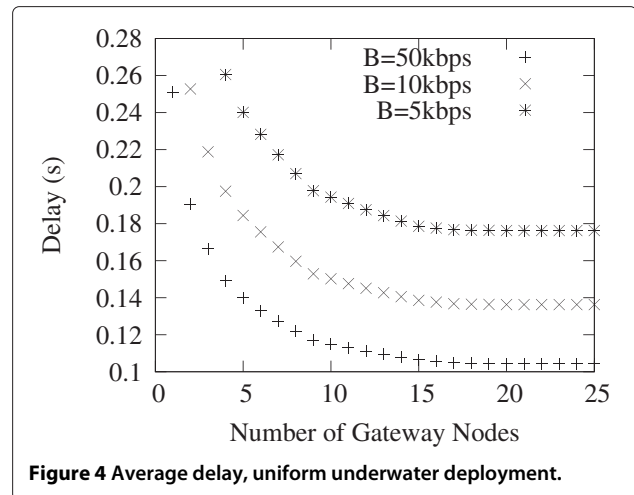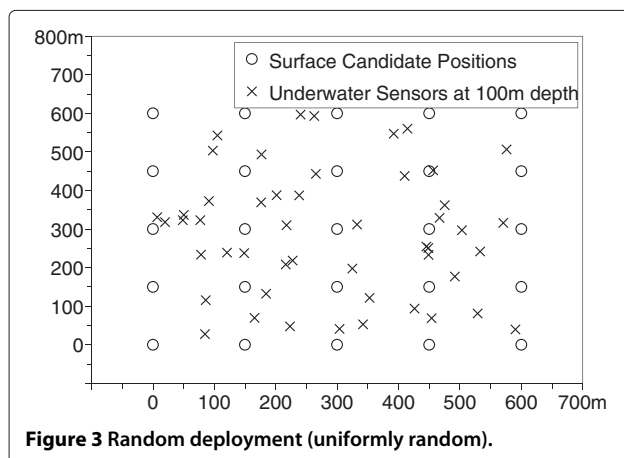**Figure 4 Average delay, uniform underwater deployment.**
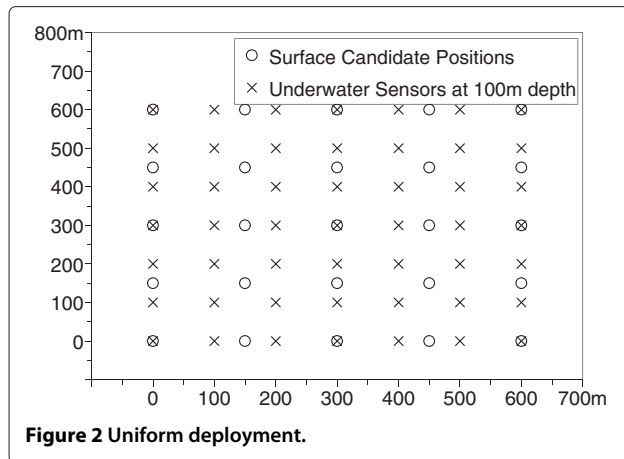
regardless of its horizontal location, is within the communication range of at least one surface gateway candidate position, thus satisfying the connectivity requirement. This guarantees that an optimal solution can be found by setting a large-enough limit on the number of surface gateway nodes, N. Finally, the data generation rate at each underwater sensor is set to 1 packet per second. The acoustic channel effective bit rate B is varied among 5, 10, and 50 kbps. Accordingly, the packet transmission time $\tau =$ 80, 40, and 8 ms, respectively, and the data generation rate $g = 0.08$, 0.04, and 0.008 packets per packet time, and the energy consumption per packet transfer $\pi^S = 80$, 40, and 8 W, respectively.

4.2 Results

To characterize the benefits and limitations of the deployment of surface gateways, we analyze the effect of the following factors:

- Number of gateway nodes

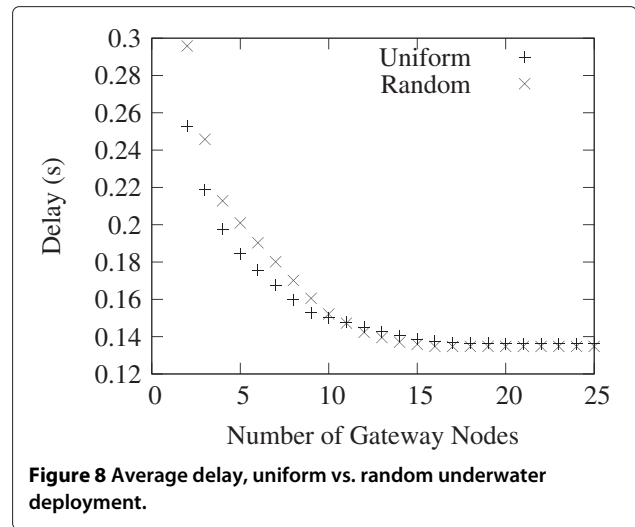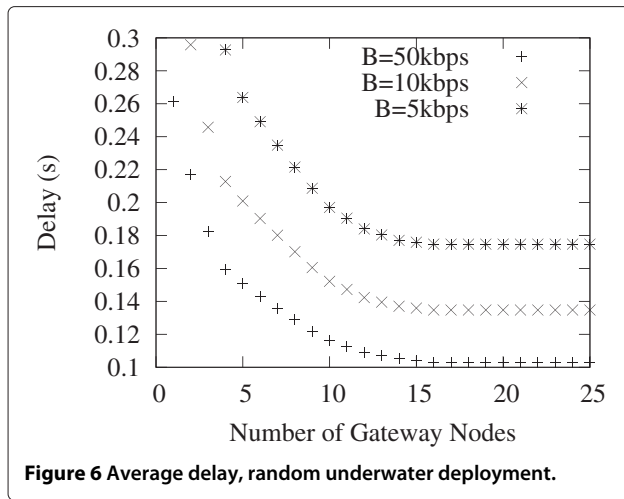When the number of allowed surface-level gateway nodes increase, the performance characteristics, average delay or average power consumption, is expected to improve. To verify that, we vary the number of allowed surface nodes from 1 to 25 nodes and solve the optimization problem to get the optimal average delay or energy. Results show that an increase in the number of surface gateways can dramatically enhance performance, especially when the network is lightly loaded.

- Network load

Intuitively, when the ratio of total data generation rate to the per-node channel bandwidth increases, the minimum number of surface-level nodes required to make the problem feasible increases. This is due to the fact that surface gateways will saturate with incoming traffic and, therefore, more nodes will be needed to handle the additional traffic corresponding to the increased data generation rate. On the other hand, increasing channel capacity reduces the network load, and consequently, our assumptions about ignoring queuing delays become more

Figure 3 Random deployment (uniformly random).

Figure 5 Average energy, uniform underwater deployment.

Figure 6 Average delay, random underwater deployment.

Figure 8 Average delay, uniform vs. random underwater deployment.
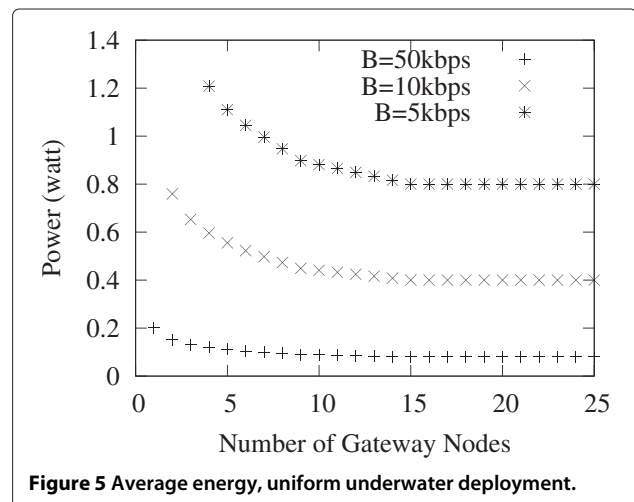
realistic. To demonstrate the effect of channel capacity on the quality of the solution, we solve the deployment optimization problem for different link capacities, namely 5, 10, and 50 kbps.

- Underwater deployment pattern

If the set of candidate surface gateway positions is pre-set, the locations of underwater sensors and the distribution of data generation load among them are expected to affect the benefit of adding more surface gateways. If underwater sensors are clustered in groups, less surface gateways are expected to feasibly route all traffic to the surface, compared to the case when underwater sensors are spread evenly over the deployment area. On the other hand, clustering increases the odds of collision and, in the case of high traffic loads, can negatively affect delay and energy consumption. To study the effect of the deployment scheme of the given set of underwater sensors on the
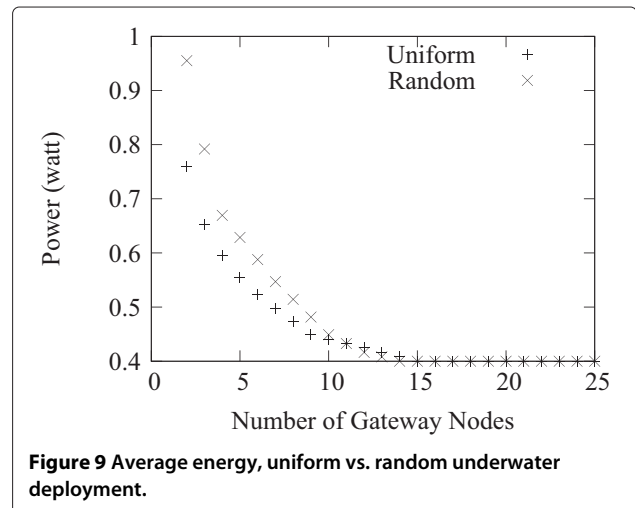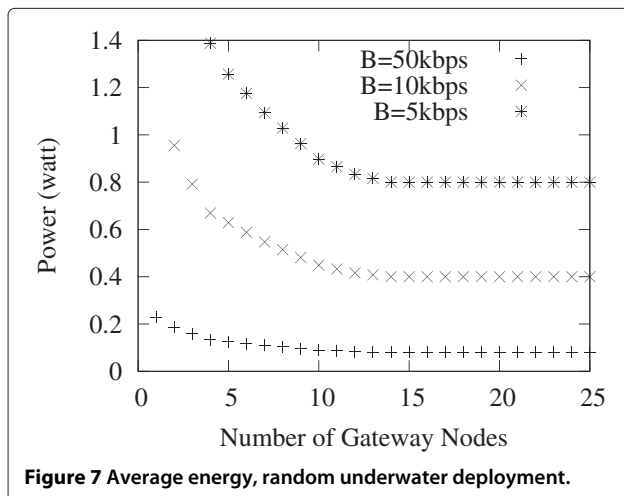
result of the surface gateway deployment optimization, we use two underwater deployment patterns.
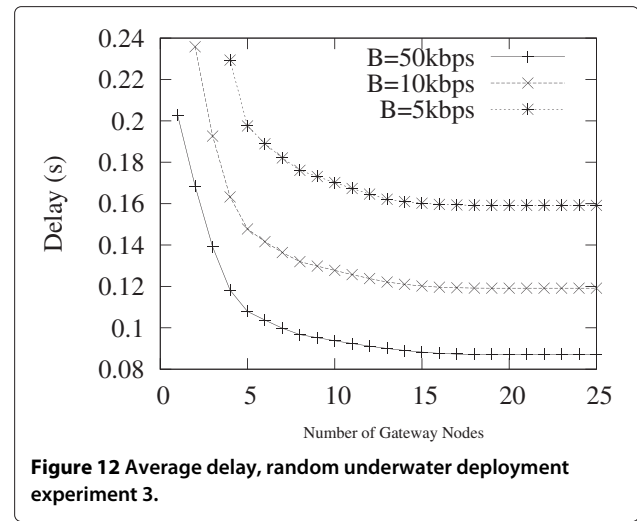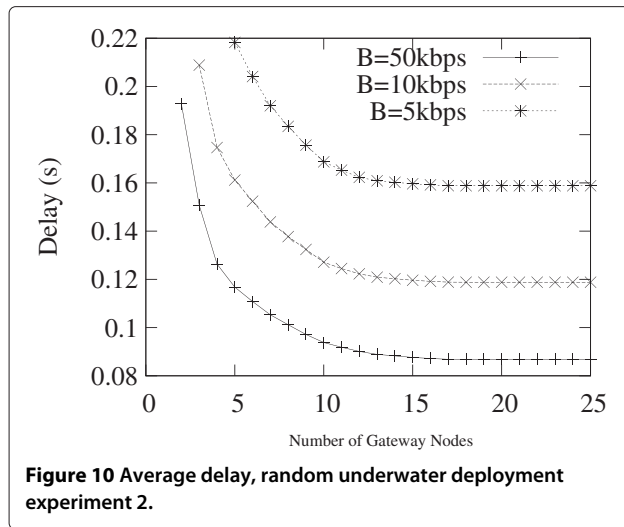
The uniform underwater deployment was chosen because the uniformity of the solution simplifies the process of verifying the results. The chosen underwater deployment consists of a 7×7 planar mesh of sensor nodes. The distance between two adjacent nodes is 100 m, and therefore, the nodes cover the entire 600×600-m area. This problem setting is illustrated in Figure 2.

It was similar to the uniform underwater deployment, except that the 49 underwater sensor nodes are distributed at random within the 600×600-m underwater area. This problem setting is illustrated in Figure 3.

4.2.1 Effect of number of gateway nodes
Results show that an increase in the number of surface gateways can dramatically enhance performance,

Figure 7 Average energy, random underwater deployment.

Figure 9 Average energy, uniform vs. random underwater deployment.

Figure 10 Average delay, random underwater deployment experiment 2.

Figure 12 Average delay, random underwater deployment experiment 3.

especially when the network is lightly loaded. For example, Figure 4 shows that the expected delay corresponding to $B = 50$ kbps can be reduced from 0.26 to 0.16 s using four surface gateways instead of one. It is also worth noting that the improvement gained by adding a surface gateway diminishes as the number of surface gateways increase. After a certain number of surface gateways, depending on underwater deployment and other factors, any additional surface nodes have negligible effect on the performance of the network. This is due to the fact that at that point, all underwater nodes communicate with the a surface gateway at the candidate position nearest to each of them, and therefore, further addition of surface nodes becomes redundant.

4.2.2 Effect of network load
Simulation results show that the performance improvement resulting from the addition of more surface gateway

nodes diminishes when the network load increases, confirming our expectations as explained before. For example, Figure 5 shows that the heavily loaded network, corresponding to the channel capacity of 5 kbps has a smaller dynamic range of $0.2/0.08 = 2.5$ than the lightly loaded network corresponding to $B = 50$ kbps, whose dynamic range is equal to $1.2/0.8 = 1.5$. Figures 4 and 5 show the effect of varying the channel capacity, in the case of uniform underwater deployment, on the expected delay and expected energy consumption, respectively. Figures 6 and 7 show the effect of varying the channel capacity in the case of random underwater deployment, on the expected delay and expected energy consumption, respectively.

4.2.3 Effect of underwater deployment pattern
Figures 8 and 9 compare the results of the uniform underwater deployment case and the random underwater

Figure 11 Average energy, random underwater deployment experiment 2.

Figure 13 Average energy, random underwater deployment experiment 3.
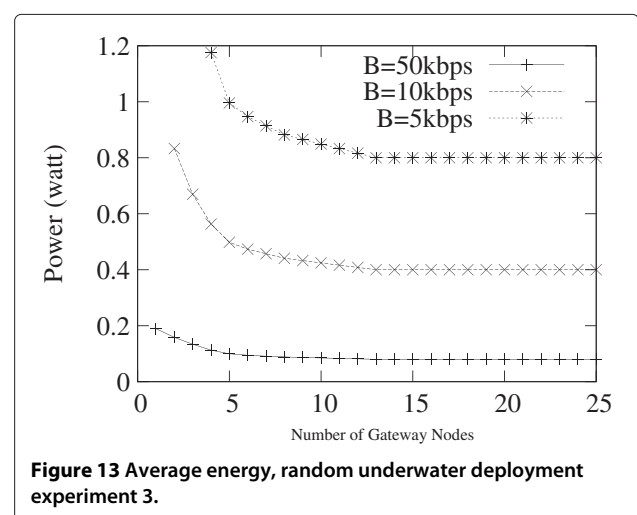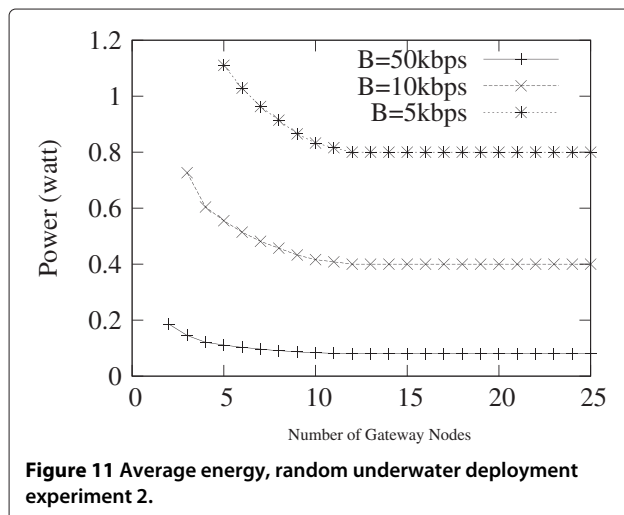
deployment case at a fixed channel capacity of 10 kbps. When the number of surface gateways is small, the randomly distributed underwater deployment suffers more congestion and therefore performs slightly poorer than the uniformly distributed counterpart in both delay and energy-consumption metrics. When the number of surface gateways increases, the effect of congestion diminishes, and the effect of clustering grows stronger. After a certain number of surface gateways, the energy consumption of the randomly distributed underwater deployments converges to that of the uniformly distributed case because, eventually, each underwater node becomes one hop away from a surface gateway. Although Figure 8 shows that the delay of the randomly distributed case eventually becomes lower than the uniformly distributed case is not necessarily always true. The problem instance shown in Figure 3 happened to have an average distance between underwater nodes and candidate surface positions, lower than that in the uniformly distributed case in Figure 2, and therefore, the randomly distributed underwater deployment case exhibits less average propagation delay. The experiments were repeated for several random underwater deployments to confirm that the results exhibit the same trends. Figures 10, 11, 12 and 13 illustrate two more samples for the random deployment results.

5 Conclusions

In this paper, we provided a generalized optimization framework for the surface gateway deployment problem. We demonstrated how to use the formulation to find the optimal placement of surface gateways with respect to a variety of optimization goals and a set of flow conservation constraints, interference constraints, and deployment constraints. We assisted our work by incorporating the optimal gateway optimization framework results in simulating the operation of the UWSN with multiple surface gateways. Our simulation results confirmed the potential for performance improvement using multiple surface gateways, such as reducing both average delay and energy consumption. It was also shown that the effect of the added surface gateways depends on the channel capacity or the network loading level, as well as the given underwater sensor deployment pattern. Our work presented here helps to pave the way for a wide variety of future research. One possible improvement is to optimize underwater and surface deployments jointly. Another is to consider mobility of surface gateways and a static UWSN with some level of location certainty models and integrate it as a new design parameter in our framework.

Competing interests
The authors declare that they have no competing interests.

References
1. G Xie, J Gibson, *A Networking Protocol for Underwater Acoustic Networks*. (CS Department, Naval Postgraduate School Operations Research Department Monterey, CA 93943, 2000)
2. JG Proakis, EM Sozer, JA Rice, M Stojanovic, Shallow water acoustic networks. IEEE Commun. Mag. **39**(11), 114–119 (2001)
3. IF Akyildiz, D Pompili, T Melodia, Underwater acoustic sensor networks: research challenges. Ad Hoc Netw. **3**(3), 257–279 (2005)
4. J Heidemann, Wei Ye, J Wills, A Syed, Yuan Li, in *Proc. of IEEE Wireless Communications and Networking Conference, WCNC 2006*. Research challenges and applications for underwater sensor networking (Las Vegas, 3–6 April 2006)
5. JH Cui, J Kong, M Gerla, S Zhou, The challenges of building scalable mobile underwater wireless sensor networks for aquatic applications. IEEE Netw. **20**(3), 12–18 (2006)
6. L Berkhovskikh, Y Lysanov, *Fundamentals of Ocean Acoustics*. (Springer-Verlag, Berlin Heidelberg, 1982)
7. WKG Seah, HP Tan, in *Proc. of the First International Conference on Integrated Internet Ad Hoc and Sensor Networks, InterSense 2006*. Multipath virtual sink architecture for wireless sensor networks in harsh environments (Nice France, 30–31 May 2006)
8. MF Younis, K Akkaya, Strategies and techniques for node placement in wireless sensor networks: a survey. Ad. Hoc. Netw. **6**(4), 621–655 (2008)
9. B Hao, H Tang, G Xue, in *Proc. of IEEE workshop on High Performance Switching and Routing, HPSR 2004*. Fault-tolerant relay node placement in wireless sensor networks: formulation and approximation (Phoenix, 2004), pp. 246–250
10. Y Hou, Y Shi, HD Sherali, SF Midkiff, On energy provisioning and relay node placement for wireless sensor networks. Wireless Commun. IEEE Trans. **4**(5), 2579–2590 (2005)
11. Q Wang, K Xu, G Takahara, H Hassanein, in *Proc. of IEEE Global Telecommunications Conference, GLOBECOM 2005*. Locally optimal relay node placement in heterogeneous wireless sensor networks (St. Louis, 2005)
12. D Pompili, T Melodia, IF Akyildiz, in *Proc. of First ACM International Workshop on Underwater Networks, WUWNet 2006*. Deployment analysis in underwater acoustic wireless sensor networks (Los Angeles, 25 September 2006), pp. 48–55
13. L Badia, M Mastrogiovanni, C Petrioli, S Stefanakos, M Zorzi, in *Proc. of First ACM International Workshop on Underwater Networks, WUWNet 2006*. An optimization framework for joint sensor deployment, link scheduling and routing in underwater sensor networks (Los Angeles, 25 September 2006)
14. W Alsalih, H Hassanein, S Akl, in *Proc. of IEEE Local Computer Networks Conference, LCN 2008*. Delay constrained placement of mobile data collectors in underwater acoustic sensor networks (Montreal, 14–17 Oct 2008), pp. 91–97
15. S Ibrahim, JH Cui, RA Ammar, in *Proc. of IEEE Military Communications Conference, MILCOM 2007*. Surface-level gateway deployment for underwater sensor networks (Orlando, 29–31 Oct 2007), pp. 1–7
16. S Misra, SD Hong, G Xue, J Tang, Constrained relay node placement in wireless sensor networks: Formulation and approximations. IEEE/ACM Trans. Netw. **18**(2), 434–447 (2010)
17. J Liu, X Han, M Al-Bzoor, M Zuba, JH Cui, RA Ammar, S Rajasekaran, in *Proc. of IEEE Symposium on Computers and Communications, ISCC 2012*. PADP: Prediction assisted dynamic surface gateway placement for mobile underwater networks (Cappadocia, 1–4 July 2012)
18. S Ibrahim, RA Ammar, JH Cui, in *Proc. of IEEE Symposium on Computers and Communications, ISCC 2009*. Geometry-assisted gateway deployment for underwater sensor networks (Sousse, 5–8 July 2009)
19. Y Xiao, Y Zhang, JH Gibson, GG Xie, in *Proc. of IEEE Second Int. Conf. on Embedded Softw. and Syst., ISCESS 2009*. Performance Analysis of p-Persistent Aloha for Multi-hop Underwater Acoustic Sensor Networks (Zhejiang, 25–27 May 2009)

20. S De, P Mandal, SS Chakraborty, On the characterization of Aloha inunderwater wireless networks. Math. Comput Model. **53**, 2093–2107 (2011)
21. AA Syed, W Ye, J Heidemann, B Krishnamachari, in *Proc. of the second workshop on Underwater networks, WuWNet 2007*. Understanding spatio-temporal uncertainty in medium access with ALOHA protocols (Montereal, 2007), pp. 41–48
22. JH Gibson, GG Xie, Y Xiao, H Chen, in *Proc. of MTS/IEEE Oceans 2007 Conference*. Analyzing the performance of multi-hop underwater acoustic sensor networks (Scotland, June 2007)