Dynamic Sensor Nodes Distribution with Coordinated Autonomous
Vehicles for Environment Pollution Monitoring and Modeling

Alfa Budiman1, Wenbo Wu1, Edisson A. Naula-Duchi2, Patricia Portillo Jiménez2, Hanifeh Imanian1, Pierre Payeur1, Luis E. Garza-Castañón2, Abdolmajid Mohammadian1, Eric Lanteigne1
1Faculty of Engineering, University of Ottawa
800 King Edward, Ottawa, ON, Canada
abudi097@uottawa.ca; wwu039@uottawa.ca; himania3@uottawa.ca;
ppayeur@uottawa.ca; majid.mohammadian@uottawa.ca; eric.lanteigne@uottawa.ca
2School of Engineering and Sciences, Tecnologico de Monterrey
Eugenio Garza Sada 2501, Monterrey 64849, México;
A00825462@tec.mx; A01400595@tec.mx; legarza@tec.mx

Abstract – This research aims toward collecting air and water samples over opportunistically selected locations to monitor pollutants distribution. To support precise sensor nodes deployment over a variety of terrains and changing conditions, not only appropriate sensor devices must be designed, but means for deployment must also be carefully studied, developed, and implemented. This paper investigates methodologies to efficiently distribute environment sensor nodes while maximizing space coverage, minimizing acquisition time, and leveraging the benefits of autonomous robotic agents to carry environmental sensors to strategic locations. The research contributes to fill existing gaps in local and global sensor networks for environment pollution monitoring by developing innovative technologies to dynamically deploy sensor nodes using mobile unmanned ground, air and water vehicles. The dispatch of dynamic sensor nodes on autonomous robotic agents to collect measurements on pollution can efficiently cover territories of different size, automatically detect areas where pollution varies significantly or reaches concerning levels, and strategically concentrate data acquisition over those regions to support the formation of more accurate data-centric pollutants dispersion models.

Keywords: Environmental data collection; pollution monitoring; sensor nodes distribution; coordinated autonomous vehicles; dynamic replanning.

1. Introduction
The climate crisis impacting the planet motivates countries around the world to work toward environment sustainability and reduction of pollution. Atmospheric pollution has detrimental health consequences on the population, while water pollution is impacting ecosystems and threatens food production. Current methods to collect data and model pollutants concentration and dispersion often involve a combination of large-scale surveys with airborne remote sensing, and small-scale local surveys where sensors are installed as static nodes. Environmental drones are used for even abatement at altitudes above ground level in a specific geographic region but obtain relatively low-density measurements. Fixed air quality monitoring stations impose limited geographical distribution, and involve high cost of installation and operation, which restrains broader deployment of sensor networks. Surface and underwater autonomous vehicles have been deployed sporadically to search for specific water body pollutant sources, while marine pollution monitoring is a fundamental component of current environmental legislation and aims to sustainably protect marine ecosystems. Remote sensing technologies along with machine learning algorithms play an increasingly important role in the accurate detection and monitoring of pollution in coastal waters like oil spill slicks or desalination outflows, assisting scientists in forecasting their trajectories, developing clean-up plans, taking timely and urgent actions, and applying effective treatments to contain and alleviate adverse effects [1-6]. However, machine learning-based approaches are intrinsically dependent on the availability of massive and accurate measurements which in turn involve efficient data collection processes.

This research leverages the potential of autonomous robotic platforms, namely unmanned ground (UGV), aerial (UAV), and water (UWV) vehicles, equipped with specialized sensors to collect air or water samples, to optimally distribute sensor nodes in the environment [7], and to efficiently monitor pollution. While simply densifying samples collection can provide
large volume of data for mapping environmental pollution, data collection is time-consuming, and information can become redundant. Intuitively, areas where pollution distribution parameters vary deserve closer examination compared to areas where air or water quality is relatively uniform in order to understand and map pollutants distribution phenomena. Therefore, proactive sensing strategies are designed to accurately capture the variability of pollution distribution, support the on-going evolution toward data-centric and machine learning-based modeling approaches, and eventually lead to accurate models of the environment which will inform policy makers and guide environmental interventions.

This paper proposes global sampling methodologies to conduct dynamic deployment of pollution sensor nodes in the environment and explores mobile platforms to demonstrate the potential of the approaches to fill the gap between fixed sensor nodes and large-scale airborne environment monitoring methods.

2. State-of-the-Art Autonomous Agents Coordination for Environment Sensors Distribution

Strategies for detecting and locating pollution sources can leverage modern data-centric representations and efficient coordination schemes to deploy sensor nodes while satisfying the operational constraints imposed by the environment, the scale of a given area for data collection and the characteristics of the pollution to be detected. Recent technological advances enable the development of multi-agent robotic systems to form distributed sensor networks capable to measure air and water quality and detect tiny amounts of particles therein. The coordination of sensor nodes builds upon multi-agent task allocation (MATA) mechanisms to dynamically determine which sensor agent should be deployed in what location, with the objective to achieve optimal overall data sampling performance. The literature offers various solutions to the MATA problem, from deterministic analytical methods that provide optimal navigation of agents at the price of computational complexity; bio-inspired heuristic methods such as ant colony [8] and particle swarm that provide close to optimal solutions with improved computational efficiency; numerical and iterative market-based strategies [9] that rely on centralized auctioneer mechanisms with the goal to minimize overall cost; or even computational methods inspired by fuzzy logic, probabilistic theory or reinforcement learning. Intentional cooperation [10] explicitly assigns a specific task to each sensor node agent. The Hungarian algorithm [11] is useful when the requirement is to perform all sensing tasks by assigning exactly one agent to every task and exactly one task to every agent in such a way that the total cost is minimized. While these coordination schemes assume that all agents are equal, the latest allocation methods [12] rather consider specialized individual agents based on their respective embedded functionalities, such as the nature of sensors they carry.

While task allocation entails identifying and matching sensing tasks to robotic agents in an optimal manner, such as sensor’s compatibility or cost minimization, task coordination entails facilitating the conduct of those tasks by agents in an effective and conflict-free manner. In the context of environmental sensing and monitoring, tasks are related to the dynamic deployment of sensor nodes over UGVs, UAVs or UWVs to collect air or water samples and measure contamination level at diverse locations. The collected information is then used to dynamically replan the navigation of autonomous vehicles carrying sensors and progressively direct them towards a source of pollution or opportunistically increase the sampling density over selected areas that are most representative of the natural pollutant dispersion process.

Classical task allocation and coordination can be formulated as an extension of the traveling salesman problem [13]. Overall, these solutions seek to optimize some sort of objective function. Given a set of $M$ robotic agents, $R = \{R_i | i = 1, 2, ..., M\}$, carrying sensor nodes, and $N$ tasks, $T = \{T_j | j = 1, 2, ..., N\}$, to conduct air or water sampling operations, the cost for sensor node $i$ to perform sample collection task $j$ is equal to $c_{ij}$. This cost is a function, $f(R_i, T_j)$, of the sensor node and the sample collection procedure. Assuming that $A = \{A_i | i = 1, 2, ..., M\}$ defines the set of sample collection allocations in which $A_i$ is the set of samples assigned to sensor node $i$, $A_i = \{T_{ij} | j = 1, 2, ..., N\}$, an optimal allocation, $A_{opt}$, is the one in which as many sample collection tasks as possible are conducted by all agents while the cost is minimal.

$$A_{opt} = \arg\min_A \sum_{i=1}^M \sum_{j=1}^N (c_{ij}) = \arg\min_A \sum_{i=1}^M \sum_{j=1}^N f(R_i, T_j)$$ (1)
To coordinate the robotic sensor nodes to their assigned sample collections effectively and without conflict, agents’ heterogeneity that involves differences between robotic agents and sensor nodes is advantageously leveraged for environmental monitoring. For example, a UAV possesses different characteristics to that of a UUV on the areas it can reach in the environment (e.g., collecting air samples at altitudes versus water samples at sea). Conversely, a UAV can use its superior mobility and field of view to provide global information (e.g., detected oil spill over the water, or plastic waste accumulation over the shore) to a UUV or a UGV, which then can use their superior endurance and payload to carry heavier sensing equipment on-site, or traverse areas unreachable by other agents such as a river or lake. An example of such coordination is discussed in [14] in which a UAV provides aerial observation of a ground target to be investigated by a UGV. Such a dynamic optimization of data collection using heterogeneous mobile robotic agents as sensor nodes in environmental surveys can substantially increase knowledge about pollutant transport pathways and feed modern machine learning based environmental models that support the development of policies and methodologies to protect the environment.

3. Methodology and Experimental Validation

To survey and analyze contamination levels in the environment, dense sampling provides ample measurements to generate accurate spatial pollution maps. However, deploying mobile sensor nodes at numerous locations sees a major increase in acquisition time, operational cost, and faces limitations on autonomous mobile agents’ autonomy, often imposed by battery life. Different optimization approaches are considered in this section, along with a description of the experimental testbed and early experimental results.

3.1. Centralized Programmatic Control for Environment Sensor Nodes Coordination

The optimization of sensor nodes distribution is first studied in a classical perspective by minimizing acquisition time and sampling effort while maximizing the coverage of the samples collection area for data-centric pollution distribution mapping and analysis. To support early development, UAVs are deployed in an indoor environment to automatically identify points of interest which are assigned to different UGV agents for data collection. A UAV equipped with a color camera uses its superior mobility and elevated point of observation to localize points of interest, identified via visual markers distributed on the ground as depicted in Figure 1. The UAV’s pose, estimated via embedded visual inertial odometry and external pose tracker sensors, is considered to determine the location of different points of interest, \((x_j, y_j)\), on the ground [15]. The latter are then allocated to individual robotic agents, \(R_i\) with coordinates \((x_i, y_i)\), either sequentially or simultaneously to distribute the environment sensing workload. Each UGV’s trajectory is planned to define a series of waypoints that serve to navigate the mobile sensor agents toward the locations where they perform air or water samples collection.

![Waypoint Allocation](image)

**Fig. 1:** Locations of interest and waypoints determination from a UAV embedded vision sensor looking downward.

In a first validation phase, waypoints assigned to each sensor node are selected such that the total travel distance for the entire group of mobile sensor nodes to conduct the environmental survey is minimized. A greedy algorithm [16] is used to calculate an optimal allocation, \(A_{opt}\), of waypoints for each agent. Through dynamic replanning, this approach calculates the locally optimal solution at each timestep, iterating through all waypoints until they are all assigned to robotic agents, or if constraints can no longer be satisfied. The cost function of Eq. (1) is defined based on Euclidean distance as:
As a refinement step, we also propose to minimize a global imbalance index, $V$, to promote even distribution of the air or water samples collection workload, where $c_i$ is the cost for agent $R_i$ to execute a given sample collection, and $C$ is the cost of all $M$ agents to execute all their assigned sample collection processes.

$$V = \sqrt{\sum_{i=1}^{M} \left( c_i - \frac{C}{M} \right)^2}$$

(3)

The greedy algorithm approach was initially tested in simulation. Figure 2 presents experimental results on one test case considering 3 robotic sensor agents (1-A, 2-B, 3-C) and 10 waypoints (0 to 9), respectively without imbalance index enforcement (Figure 2a), and with imbalance index enforcement (Figure 2b). The results of the simulation are detailed in Table 1, showing the distance travelled per sensor node in each case, the total travel distance, and the global imbalance index achieved. It can be observed that while the overall travel distance is less when the imbalance index is not enforced, the workload distribution among the different sensor nodes does not use every agent effectively, even leading to a single sensor node performing all sample collections sequentially along a complex path rather than sharing the samples collection process. Conversely, when the imbalance index, $V$, is enforced and minimized in the workload distribution process, all sensor nodes are put to contribution and share the data collection process by each performing a subset of the samples collection in parallel over the surveyed area, at the expense of a slightly increased total travel distance. Also, a significantly better balance in the workload distribution among the agents is achieved, which is reflected in a smaller global imbalance index, $V$, value. Moreover, since each sensor node is less solicited, the overall autonomy of the autonomous robotic agents that carry the sensors and the capability to pursue additional samples collection through an iterative process are significantly improved.

![Fig. 2: Simulation of atmospheric or water samples collection via dynamic waypoints allocation](image)

**Table 1: Experimental evaluation of performance under two modes of waypoints allocation for samples collection.**

<table>
<thead>
<tr>
<th>Agent</th>
<th>Travel Distance</th>
<th>Imbalance Index</th>
<th>Travel Distance</th>
<th>Imbalance Index</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-A</td>
<td>112.1</td>
<td>-   -</td>
<td>44.2</td>
<td>-   -</td>
</tr>
<tr>
<td>2-B</td>
<td>0</td>
<td>-   -</td>
<td>53.6</td>
<td>-   -</td>
</tr>
<tr>
<td>3-C</td>
<td>0</td>
<td>-   -</td>
<td>34.6</td>
<td>-   -</td>
</tr>
<tr>
<td>Global</td>
<td>112.1</td>
<td>91.5</td>
<td>132.4</td>
<td>13.4</td>
</tr>
</tbody>
</table>
3.2. Experimental Autonomous Robotic Platforms for Sensor Nodes Deployment

To experimentally validate the proposed dynamic environment sensors deployment methods beyond simulation, several robotic platforms equipped with specialized sensor nodes are developed. These include a quadcopter and a single-rotor UAV, a modified UGV and a UWV platform.

An off-the-shelf RB5 quadcopter drone, shown in Figure 3a, equipped with PX4 autopilot firmware for flight control and a single board computer running Ubuntu 18.08 with robot operating system (ROS) Melodic version is experimentally evaluated. The UAV-embedded PX4 and ROS communicate with a central control computer through MAVROS using the MAVLink protocol, which enables programmatic flight control of the drone with limited intervention from the operator, as well as access to and recording of embedded sensor data through ROS topics. Equipped with an embedded color camera, it is used to detect points of interest on the ground, and for collecting air samples at various elevations. Alternatively, a custom single rotor unmanned aerial vehicle (SR-UAV), shown in Figure 3b, is also in the design phase to help with surveying the environment and detect points of interest. The design consists of a mechanical structure where a single rotor and four control fins are placed symmetrically in the lower part of the aircraft and directly below a propeller. The control fins are directly attached to servos, so that an angle of attack is obtained to produce a torque in each of the main axes [17]. The system involves up of 5 control outputs: one that comes from the thrust of the engine and four from the regulation of the movement of the control fins. The required torque is generated by using the four fin controls, to maintain both positive and negative angle of attack behaviour. The flight control system includes all the degrees of freedom of the SR-UAV, thus separating it into two controllers: one being the translational (x, y, z) and the other the attitude ($\phi_x$, $\phi_y$, $\phi_z$). As such, the design provides the ability to move laterally and to fly forward. The prototype shown in Figure 3b was built with the aim of determining the best material and aerodynamics of the components. Benefits of a single-rotor UAV over a quadcopter include increased versatility and higher payload to accommodate embedded sensors due to lower structural weight involved by a single actuator, and the capability to custom design the platform to perform well in the considered environment. Moreover, various control schemes can be considered to ensure accurate trajectory control performance, which in turn leads to accurate sampling of pollution measurements.

As shown in Figure 3c, the Turtlebot3 waffle is a differential drive robot. It embeds a microcontroller connected to a Raspberry Pi minicomputer which runs Ubuntu 20.04 with ROS Noetic, also enabling programmatic control and access to sensor data via ROS topics.

To support pollution samples collection in the water, a Thor Robotics TrenchRover 110 ROV designed for underwater research is used, as shown in Figure 3d. A custom microfluidic sensor device [18] is embedded on the UWV to measure nitrate concentration in the water. The compact robotic platform (360x200x200 mm) has four vector propellers, arranged in such a way that there are two horizontal and two perpendicular control axes for manoeuvrability. The robot carries a frontal camera and two frontal LED spotlights. It operates through a AT9S Pro remote controller. Since electromagnetic waves fade very quickly in water, this vehicle uses a tether cable to connect the vehicle and the remote control. The compact size of the vehicle permits a small impact on the aquatic environment but has the disadvantage of a reduced payload. Hence, for the integration of the microfluidic device, the required hardware must be attached as appendages to the vehicle. Other important

Fig. 3: UAVs for points of interest detection and sensor nodes deployment: a) RB5 quadcopter drone, b) single-rotor-unmanned aerial vehicle (SR-UAV); UGV and UWV for sensor nodes deployment: c) Turtlebot3 waffle, d) TrenchRover 110 ROV.

For the purpose of testing deployment algorithms, an off-the-shelf Turtlebot3 waffle UGV, shown in Figure 3c, was considered given its rapid integration capability under ROS. The Turtlebot3 waffle is a differential drive robot. It embeds a microcontroller connected to a Raspberry Pi minicomputer which runs Ubuntu 20.04 with ROS Noetic, also enabling programmatic control and access to sensor data via ROS topics.

To support pollution samples collection in the water, a Thor Robotics TrenchRover 110 ROV designed for underwater research is used, as shown in Figure 3d. A custom microfluidic sensor device [18] is embedded on the UWV to measure nitrate concentration in the water. The compact robotic platform (360x200x200 mm) has four vector propellers, arranged in such a way that there are two horizontal and two perpendicular control axes for manoeuvrability. The robot carries a frontal camera and two frontal LED spotlights. It operates through a AT9S Pro remote controller. Since electromagnetic waves fade very quickly in water, this vehicle uses a tether cable to connect the vehicle and the remote control. The compact size of the vehicle permits a small impact on the aquatic environment but has the disadvantage of a reduced payload. Hence, for the integration of the microfluidic device, the required hardware must be attached as appendages to the vehicle. Other important
parameters are the energy source and the time required for taking a measurement with the microfluidic device as they condition the autonomy of the vehicle. Having an external power source connected through the tether cable handles the power issue and allows this research to focus on the sampling rate of the microfluidic sensor device.

3.3. Optimization with Opportunistic Distribution of Sensor Nodes

Beyond minimizing travel distance as in Section 3.1, increased efficiency on samples collection can be achieved by leveraging opportunistic sensing strategies. An algorithm which was introduced for 3D range data acquisition [19] formulates an estimated improvement measure to automatically define the areas where additional data sample collection leads to optimal knowledge acquisition about the environment. It is here considered as a method to efficiently distribute the environmental sensor nodes carried by air, water or ground robotic agents working in collaboration. The methodology aims to collect only sparse samples where environment pollution distribution is relatively uniform, while dynamically densifying the distribution of sensor nodes over regions where the contamination parameters vary more extensively.

The proposed pipeline for opportunistic distribution of sensor node agents is shown in Figure 4. Initially, sensor nodes are uniformly but sparsely distributed, which enables a rapid collection of samples over a wide area and coarse pollution mapping to support future data sampling runs. Any area where measured pollution parameters vary becomes a region of interest (ROI), over which the sensor nodes will be dynamically redeployed to collect additional samples that will increase the knowledge cumulated in an environment pollution map. The method to determine the optimal locations for subsequent selective data acquisition builds upon ordinary kriging [20], which computes spatial autocorrelation among available samples to estimate the measurements and variance expected at unsampled locations. Among those locations, those with highest kriging variance, which corresponds to the highest uncertainty on the measurements, are selected as the next locations to visit for sensor nodes to collect samples. The process is repeated iteratively, forming a series of unsampled locations whose coordinates are defined as waypoints for navigating the sensor nodes with autonomous robotic agents and collecting additional measurements. The opportunistically collected samples supplement the initial sparsely distributed measurements, enabling the creation of spatial pollution maps with rich information while minimizing the number of data samples collected.

![Block diagram of proposed opportunistic sensor nodes distribution and samples collection.](image)

Experiments were conducted to validate the feasibility of the proposed sensor nodes deployment method. A dataset of 900 samples was simulated with spatial location and value to form a high-density ground truth data representation, shown in Figure 5a, where colors represent environmental parameter values. Figure 5b depicts the corresponding spatial map of estimated values using ordinary kriging. Then the procedure described in Figure 4 is implemented. Sparse sampling extracts uniformly distributed points from the ground truth map at a given density, here 100 samples (Figure 5c). A new map of measurements is formed with the sparsely subsampled dataset (Figure 5d). Regions where values significantly vary become ROIs for opportunistic data collection and candidate sampling locations are introduced. Estimates of the measurement value and corresponding variance on each candidate are calculated with ordinary kriging. Candidates with the highest kriging variance then define the waypoints for robots to strategically distribute sensor nodes at the next sampling iteration (Figure 5e). The spatial map is updated with opportunistically collected measurement values (Figure 5f). The root mean square error (RMSE) between the ground truth spatial map and the estimated spatial map is used as a performance metrics. To validate
the proposed method, performance is monitored over cases with different initial sparse sampling density. Table 2 reports for cases where initial subsampling involves respectively 36, 100, or 225 data collection locations.

![Fig. 5: Opportunistic selection of sensor node locations: a-b) ground truth data with 900 samples, and corresponding spatial measurements map; c-d) location of initial 100 uniformly distributed sparse samples, and corresponding measurements map; e-f) opportunistically expanded data sampling in ROI, and corresponding measurements map.](image)

<table>
<thead>
<tr>
<th></th>
<th>36 uniform samples</th>
<th>with opportunistic samples</th>
<th>100 uniform samples</th>
<th>with opportunistic samples</th>
<th>225 uniform samples</th>
<th>with opportunistic samples</th>
</tr>
</thead>
<tbody>
<tr>
<td>RMSE</td>
<td>1.0992</td>
<td>0.9004</td>
<td>1.1204</td>
<td>0.8460</td>
<td>1.1063</td>
<td>0.8438</td>
</tr>
</tbody>
</table>

Based on achieved RMSE against ground truth data, the spatial map of measurements generated with opportunistically selected sample locations improves in accuracy over the one relying on uniformly distributed sparse samples only. As selected samples are collected where uncertainty is higher, they contribute to increase knowledge over those regions.

### 3.4. Data-Centric Environmental Pollution Modeling

As an example of application in pollution monitoring to predict oil pollution concentration in the marine environment [3], the multilayer perceptron (MLP) method, a feedforward backpropagation artificial neural network, is employed. The developed model receives spatial data of oil pollution concentration in the water body and applies MLP on the dataset to train the model [4]. After validation, the trained model can be used to predict oil pollution distribution in the case of an oil spill event, which in turn can guide the deployment of sensor nodes embedded on UWVs to conduct additional data sampling over critical areas as pollutants keep dispersing.

The data used for initial training of the model are obtained from simulation due to the lack of substantial and publicly available datasets on oil pollution in the water. Considering a simulation-to-real knowledge transfer development strategy and building upon the proposed dynamic sensor nodes deployment approaches introduced in previous sections, the next step consists of dynamically deploying sensor nodes, to collect water samples in an opportunistic manner and where most needed to maximize knowledge acquisition, rather than relying only on simulation data or on human manual selection of the preferred sample collection locations. The overall objective of the optimized data sampling process supported by dynamically distributed sensor nodes is to feed pollution dispersion models with the most relevant information, which eventually improves the accuracy of the modeling and provides a better prediction of pollutants dispersion.

### 4. Conclusion

This paper introduces key concepts for an automated selection of optimal locations to conduct data samples collection in the environment with the goal to accurately but efficiently monitor and model air or water pollution. To achieve this objective, innovative approaches to dynamically deploy specialized sensor nodes over air, ground or water autonomous vehicles are presented. The research contributes to the emergence of a next generation of environmental survey mechanisms and data-centric pollution models that prove critical for informing policy makers and guiding interventions in the environment. Future work will involve continuous development of the robotic platforms as sensors carriers, improve scalability and integration with specialized sensor devices, and further refine and experimentally validate the design of optimal task allocation and coordination strategies for sensor nodes distribution in the environment.
Acknowledgements

The authors acknowledge the financial support from the University of Ottawa-Tecnológico de Monterrey Seed Grant Program toward this research, as well as support from the Natural Sciences and Engineering Research Council of Canada under the Discovery grant program. We also acknowledge CONACyT (México) for the support given to students at Tecnológico de Monterrey through the scholarships program.

References


