A Submodular Receding Horizon Control Strategy to Distributed Persistent Monitoring

Xiaohu Zhao[1,2,3], Yuanyuan Zou[1,2,3]* and Shaoyuan Li[1,2,3]

[1]Department of Automation, Shanghai Jiao Tong University, Shanghai, China, [2]Key Laboratory of System Control and Information Processing, Ministry of Education of China, Shanghai, China, [3]Shanghai Engineering Research Center of Intelligent Control and Management, Shanghai, China

This paper investigates the multi-agent persistent monitoring problem *via* a novel distributed submodular receding horizon control approach. In order to approximate global monitoring performance, with the definition of sub-modularity, the original persistent monitoring objective is divided into several local objectives in a receding horizon framework, and the optimal trajectories of each agent are obtained by taking into account the neighborhood information. Specifically, the optimization horizon of each local objective is derived from the local target states and the information received from their neighboring agents. Based on the sub-modularity of each local objective, the distributed greedy algorithm is proposed. As a result, each agent coordinates with neighboring agents asynchronously and optimizes its trajectory independently, which reduces the computational complexity while achieving the global performance as much as possible. The conditions are established to ensure the estimation error converges to a bounded global performance. Finally, simulation results show the effectiveness of the proposed method.

Keywords: persistent monitoring, distributed control, receding horizon control, optimization, submodular function

1 INTRODUCTION

In recent years, the multi-agent persistent monitoring problem has received much attention because of the wide range of applications such as smart cities, intelligent transportation, and industrial automation (Nigam, 2014; Yu et al., 2015; Ha and Choi, 2019; Maini et al., 2020). It involves a finite set of targets with dynamical behaviors that need to be monitored, and the main objective is to design the motion strategy for a team of agents equipped with sensors to move between these targets to collect information or minimize the uncertainty metric of targets over a long period of time. In persistent monitoring tasks, due to the dynamic characteristics of the target states, agents need to visit these targets and stay for some time to collect enough information or avoid unbounded estimation errors. In other words, the dynamic target states make the agents perform time-varying monitoring tasks continuously, which makes it difficult for agents to interact with each other effectively and brings challenges to the design of effective monitoring strategies (Rajkumar et al., 2010).

At present, most approaches for multi-agent persistent monitoring problems are exploited in a centralized fashion, such as reinforcement learning (Chen et al., 2020; Liu et al., 2020), approximate dynamic programming (Deng et al., 2017), data-driven (Alam et al., 2018) and others (Smith et al., 2011; Zhao et al., 2018; Asghar et al., 2019; Ostertag et al., 2019; Hari et al., 2020). An incremental sampling-based algorithm is proposed (Lan and Schwager, 2013) to plan a periodic monitoring

trajectory for sensing robots in Gaussian Random Fields. Aiming to solve the road-network persistent monitoring problem, (Wang et al., 2020) designed a heuristic path planning algorithm from a decision-making perspective, which enables a UGV to persistently monitor the viewpoints of the road network without traversing them. (Pinto et al., 2020) described the multi-agent persistent monitoring process as a hybrid system, and the Infinitesimal perturbation analysis (IPA) was adopted to solve this problem.

However, as the number of targets and agents increases, the computational complexity of designing a centralized controller can be overwhelmingly high and unreliable under uncertainty monitoring tasks such as the existence of stochastic disturbances. Therefore, a decentralized approach where agents act upon only local information or communicate with only local targets and agents is more desirable. To overcome this issue, (Cassandras et al., 2012; Zhou et al., 2018; Zhou et al., 2020) proposed some decentralized IPA algorithms which significantly reduce communication costs. However, these approaches often converge to poor locally optimal solutions and still need complex calculations because they involve the gradient parameter mentioned in (Pinto et al., 2020). Moreover, the dynamics of targets could be subject to stochastic uncertainties, and the form of the optimization problem could be subjected to multiple constraints such as agent dynamics, which makes it difficult to solve by the methods mentioned above.

Actually, considering the definition of submodular optimization (Gharesifard and Smith, 2017; Mackin and Patterson, 2018; Lee et al., 2021) and the event-driven nature of the multi-agent persistent problem, distributed model predictive control (DMPC) is an attractive strategy because of its ability to deal with complex constraints effectively and the flexibility of solving optimization problems online (Zeng and Liu, 2015; Mi et al., 2018; Zou et al., 2019; Li and Li, 2020). Within the receding horizon framework, the agents only need to rely on local information to optimize decisions on a short horizon each time, which greatly reduces the computational requirements. (Rezazadeh and Kia, 2021) assigned to each target a concave and increasing reward function, and a distributed sub-optimal greedy algorithm with bounded performance was designed based on the submodularity of the objective function. However, this approach does not take into account the need of dwell time for agents to monitor targets, which is related to the agents' real-time strategies, thus limiting their application. Although (Welikala and Cassandras, 2021) proposed an event-driven receding horizon control strategy for distributed persistent monitoring problems, aiming to optimally control each of the agents in an online distributed manner using only a minimal amount of computational power, yet it only provided a basic distributed control method in the receding horizon control framework, and the cost of transforming the global persistent monitoring problem into multiple distributed optimization subproblems can not be effectively evaluated, which means that the agent strategies derived from local information can be locally optimal.

Based on the discussions mentioned above, the centralized methods require plenty of computing resources, while the existing distributed strategies are difficult to achieve a good monitoring performance. It is still challenging to balance the demand of computing resources and monitoring performance. In this paper, a distributed submodular receding horizon control method was proposed for multi-agent persistent monitoring tasks. In the submodular optimization framework, we first decompose the global optimization objective into multiple local optimization problems driven by monitoring events, then a distributed receding horizon control strategy is proposed, where each agent optimizes its trajectory of the next horizon, including the target to be visited and the dwell time. The utility function that measures the monitoring targets of the agent on each finite horizon is defined and proved to be a submodular function. Additionally, a distributed greedy algorithm using only local information of neighborhood targets and agents is proposed to obtain the optimal solution. Finally, some basic properties of the strategy are analyzed, and the performance of the algorithm is demonstrated by a simulation example. It is worth emphasizing that the algorithm doesn't require any parameters and only involves a finite number of calculations to obtain a bounded monitoring performance.

The structure of the paper is as follows: **Section 2** formulates the persistent monitoring problem. **Section 3** presents the distributed optimization problem under the receding horizon control strategy and the distributed optimal decision of each agent; **Section 4** discusses the submodular properties of the distributed optimization problem, and **Section 5** gives a distributed greedy algorithm. A simulation is given in **section 6** to verify the effectiveness of the algorithm. **Section 7** concludes the paper.

2 PROBLEM STATEMENT

Consider M targets and N mobile agents equipped with sensing capabilities in an undirected graph $\mathcal{G} = (\mathcal{V}, \zeta)$, where every vertex in $\mathcal{V} = \{1, 2, \ldots, M\}$ represents a target, and the edge set $\zeta \subseteq (i, j): i, j \in \mathcal{V}$ represents the trajectory segments available for agents to travel between targets i and j, the travel time over nodes i and j is represented by an associated value ρ_{ij}. The neighbor set of target i is $\mathcal{N}_i$, including all points directly connected to it. The main focus of this problem is to persistently reduce the uncertainty of targets by exploiting mobile agents as target state sensors.

Target Dynamics. The target states are dynamically changing with time, and the uncertainty estimation of target i is denoted by $R_i(t)$ according to the following model:

$$\dot{R}_i(t) = \begin{cases} 0 & if\ R_i(t) = 0\ \&\ A_i \le B_i r_a \\ A_i - \sum_{a \in \Gamma} B_i r_a & otherwise, \end{cases} \tag{1}$$

Where $R_i(t)$ increases with a prespecified rate A_i if no agent is visiting it, and decreases with a rate of $\sum_{a \in \Gamma} B_i r_a - A_i$ otherwise. $\sum_{a \in \Gamma} B_i r_a$ indicates the joint monitoring capability of agents to target i and Γ is the set of current agents on target i. Each agent can have different monitoring capabilities, denoted by r. Note that $R_i(t) \ge 0$ for all t.

Agent sensing model. In this graph topology, agents $a \in N$ can only be on the vertex $i \in \mathcal{V}$ or edges in ζ. If an agent is on a vertex i, it means that this agent is monitoring the current target i;

otherwise, it is on the way to the next target j. Assume that the velocity of each agent is $v = \{0, 1\}$, the velocity is 0 if an agent stays at a target, otherwise the velocity becomes 1.

Assumption 1. For any given agent $a \in \Psi$, it's monitor capability r_a will always guarantee $B_i r_a - A_i \geq 0$ and $B_j r_a - A_j \geq 0$, where $j \in \mathcal{N}_i$.

The goal of the persistent monitoring tasks is to minimize the average estimation error of the targets, aiming to solve the following optimization problem:

$$min \quad J = \frac{1}{T} \int_0^T \sum_{i=1}^M R_i(t)dt, \tag{2}$$

subject to target dynamics in (1).

Note that some studies have considered the correlation between the monitoring capability of an agent and the distance to the target. This paper simplifies the monitoring model of agents, where the targets on a given topology represent the detection range of the agent. In fact, only when the agent reaches the target, will it perform the monitoring task. This simplification does not affect the effectiveness of the proposed method.

3 DISTRIBUTED RECEDING HORIZON STRATEGY

A complete solution to the problem (2) requires the determination of the access trajectory and dwell time of the agents, which is usually computationally intractable. Moreover, the process requires global information about all agents and targets. It has been mentioned that the agent will leave the current target to perform the next monitoring task after completing the current monitoring target task, i.e., the global monitoring process of agents is composed of multiple continuous switching events. Each event represents the descent of a target's estimation error, and the length of the event is determined by the target state described in (1). Therefore, the problem (2) can be formulated as:

$$J = \frac{1}{T} \sum_{i=1}^M \sum_{k=0}^K \int_{t_0}^{t_1} R_i(t)dt, \tag{3}$$

Where $k = 0, 1, \ldots K$ is the number of times the target i has been visited in period T, and $J_i(t_0, t_1) = \int_{t_0}^{t_1} R_i(t)dt$ denotes the local contribution to the global objective J by a target $i \in \mathcal{V}$ during a time period $t \in [t_0, t_1]$.

The form of problem (3) contains multiple block optimization problems. This enables us to use an event-driven receding horizon control strategy to solve this problem, which has been widely used in multi-agent cooperative control and other fields.

Assumption 2. Each agent has access to the following information and communication.

- Once an agent a reaches a target i, it obtains the current uncertain states and models of all targets in the local set $\mathcal{N}_i$, but only when agent reaches the target can make $R_i(t)$ decrease.

- Agent communicates with the target in local set $\mathcal{N}_i$ to know whether it is monitored by other agents, and if so, the current agent communicates with them.

Remark 1. Both the communication and monitoring actions of agents depend on the local graph topology. Agents with the same accessible targets can communicate between them, and the current target acts as a relay for the communication. This communication constraint is much more relaxed than the centralized method, especially for applications such as marine environmental monitoring where communication is expensive. This assumption can be relaxed to the task scenario considering the monitoring distance under certain conditions.

Definition 1. If an agent a performs a monitoring task at target i, then only targets in the local set $\mathcal{N}_i(a, t)$ are allowed to access the next time. The local agent set of agent a is $\mathcal{A}_a(t) = \{\mathcal{N}_i(a, t) \cap \mathcal{N}_j(b, t) \neq \varnothing\}$, including agents that can access the same target in $\mathcal{N}_i$, that is, if there are more than one agent have access to the same target in $\mathcal{N}(a, t)$, communication can be established between them.

As mentioned in **Section 2**, in order to minimize the optimization objective 3, agents need to determine the dwell time at each target first. Under Assumption 1, when an agent a completes the task of target i at time t_i and goes to j, according to (1), it holds that $R_j(t) \geq 0$ for all $t \in [0, T]$. Besides, agent needs to leave the current target j immediately once the monitoring task is completed. Therefore, R_j is maximum when agent a reaches agent j, and then begins to decrease until it reaches 0 during $[t_0, t_1]$. It follows that its optimal dwell time u_j^k in j is

$$u_j^k(t_i, t_1) = \frac{R_j(t_i) + \rho_{ij}A_j}{B_j r_a - A_j}. \tag{4}$$

Lemma 1. Suppose agent a ready to go to the next target i, at time t_0, and arrives target i at time t_b, finish the monitor task of target i at time t_1. The local objectives during $t \in [t_0, t_1]$ can be formulated as $J_i(t_0, t_1) = \frac{u_i^0}{2}(2R_i(t_0) + A_i u_i^0) + \frac{u_i^1}{2}(2(R_i(t_0) + A_i u_i^0) - (B_i r_a - A_i)u_i^1)$, $J_j(t_0, t_1) = \frac{u_i^0 + u_i^1}{2}(2R_j(t_0) + A_j(u_i^0 + u_i^1))$, where $j \in \mathcal{N}_i$, $u_i^0 = t_b - t_0$ indicates the travel time to i, and $u_i^1 = t_1 - t_b$ denotes the visiting time at target i.

Proof. According to the definition of optimization objective 2, J can be discribed as $J = \frac{1}{T}[\sum_{j \in M \backslash i} \int_0^T R_j(t)dt] + \frac{1}{T}[\int_0^{t_0} R_i(t)dt + \int_{t_0}^{t_1} R_i(t)dt + \int_{t_1}^T R_i(t)dt]$. Based on (1), $J_i(t_0, t_1) = \frac{R_i(t_0) + R_i(t_1)}{2}(t_1 - t_0) = \frac{t_1 - t_0}{2}[2R_i(t_0) + \dot{R}_i(t)(t_1 - t_0)]$. In the period $[t_0, t_1]$, the time of agent a at target i is divided into two sections, including monitoring period and movement period. Meanwhile, the target j in the neighborhood changes at the original rate, as shown in **Figure 1**. Then, $J_i(t_0, t_1) = \frac{u_i^0}{2}(2R_i(t_0) + A_i u_i^0) + \frac{u_i^1}{2}(2(R_i(t_0) + A_i u_i^0) - (B_i r_a - A_i)u_i^1)$ for $j \in \mathcal{N}_i$, where $u_i^0 = t_b - t_0$ indicates the travel time to i, and $u_i^1 = t_1 - t_b$ denotes the visiting time at target i, and $J_j(t_0, t_1) = \frac{u_i^0 + u_i^1}{2}(2R_j(t_0) + A_j(u_i^0 + u_i^1))$. This completes the proof.
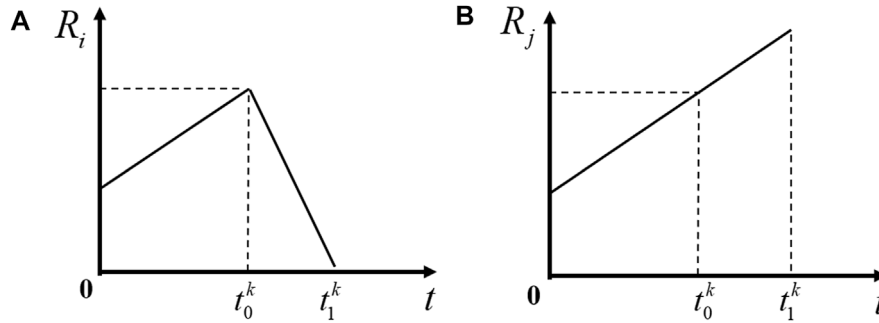
FIGURE 1 | The real-time estimation error R of agents (**(A)** presents the real-time estimation error of target i when it is monitored by agents, and **(B)** presents the real-time estimation error of target i if it is not monitored by agents).

When agent a decides the next visiting target, the local objective J_i cannot guarantee global performance, agent a needs to consider the neighborhood targets and agents. The local optimization objective can be described as

$$\overline{J}_i(t_0, t_1) = J_i(t_0, t_1) + \sum_{j \in \mathcal{N}_j} J_j(t_0, t_1). \tag{5}$$

In the design of each controller, the interaction of agents is distributed since only the states of neighboring nodes and agents rather than global information are considered.

4 SUBMODULAR UTILITY FUNCTION

To achieve the global monitoring performance, the amount of reduced target estimation error during each event is evaluated by adopting the utility function. When target i is monitored by agents for the kth time, since the global optimization objective of agent $a \in N$ is related to the total time of monitoring task, the utility function can be defined as the difference between the uncertainty of the target in the remaining time $[t_0^k, T]$ before and after the start of the kth monitoring, i.e

$$U_i^k = \int_{t_0^k}^{T} R_i(t)dt - \int_{t_0^k}^{T} R_i^k(t)dt, \tag{6}$$

where $R_i^k(t)$ indicates the estimation error of target i under the strategy before the kth monitoring, $R_i(t)$ denotes the estimation error of target i under the strategy after the start of kth monitoring, t_0^k is the starting time when one agent decides to access target i.

In the process of multi-agent persistent monitoring tasks, the utility of targets will decrease marginally with the number of agents dwelling on them.

Lemma 2. The utility function of target i is submodular and satisfies $U_i^k(\varnothing) = 0$.

Nondecreasing: $U_i^k(\tilde{P}) \leq U_i^k(P), \forall \tilde{P} \subset P \subset \Psi.$

Submodular: $U_i^k(P \cup \{a\}) - U_i^k(P) \leq U_i^k(\tilde{P} \cup \{a\}) - U_i^k(\tilde{P}),$ $\forall \tilde{P} \subset P \subset \Psi$ and $a \subset \Psi.$

Proof. When target i is not being monitored by any agent $a \in N$, it holds that $U_i^k(\varnothing) = 0$ due to the definition of U_i^k in function **(6)**. Define $L_{\tilde{P}} = \sum_{b \in \tilde{P}} B_i r_b$ and $L_P = \sum_{b \in P} B_i r_b$, where $\tilde{P} \subset P \subset N/\{a\}, \{a\} \subset N, \Delta_i^k = P - \tilde{P}$. According to **(4)**, one agent spend at most $t_i = \frac{R_i(t_0)}{\sum B_i r_a - A_i}$ to make the estimation error of target i decrease to 0, then the cumulative monitoring error J_i^k of target i in the remaining time can be expressed as:

$$
\begin{aligned}
J_i &= \int_{t_0^k}^{T} R_i(t)dt \\
&= \int_{t_0^k}^{t_i} \frac{R_i^2(t)}{\sum B_i r_a - A_i}dt \\
&= \frac{R_i^2(t_0)}{2(\sum B_i r_a - A_i)},
\end{aligned}
$$

While,

$$
\begin{aligned}
U_i^k(\tilde{P}) - U_i^k(P) &= \int_{t_0^k}^{T} R_i(t)dt - \int_{t_0^k}^{T} R_{i,\tilde{P}}^k(t)dt - \left(\int_{t_0^k}^{T} R_i(t)dt - \int_{t_0^k}^{T} R_{i,P}^k(t)dt \right) \\
&= \int_{t_0^k}^{T} R_{i,P}^k(t)dt - \int_{t_0^k}^{T} R_{i,\tilde{P}}^k(t)dt \\
&= \frac{R_i^2(t_0)}{2} \left(\frac{1}{\sum_{b \in \tilde{P}} B_i r_b + \sum_{a \in \Delta_i^k} B_i r_a - A_i} - \frac{1}{\sum_{b \in \tilde{P}} B_i r_b - A_i} \right) \\
&= -\frac{R_i^2(t_0)}{2} \frac{\sum_{a \in \Delta_i^k} B_i r_a}{\left(\sum_{b \in \tilde{P}} B_i r_b + \sum_{a \in \Delta_i^k} B_i r_a - A_i \right)\left(\sum_{b \in \tilde{P}} B_i r_b - A_i \right)} < 0,
\end{aligned}
\tag{7}
$$

Which means that the utility function of target i is nondecreasing. Similarly,

(a) $U_i^k(\tilde{P} \cup \{a\}) - U_i^k(\tilde{P}) = \frac{R_i^2(t_0)}{2} \frac{B_i r_a - \sum_{b \in \tilde{P}} B_i r_b}{\left(\sum_{b \in \tilde{P}} B_i r_b - A_i \right)(B_i r_a - A_i)},$
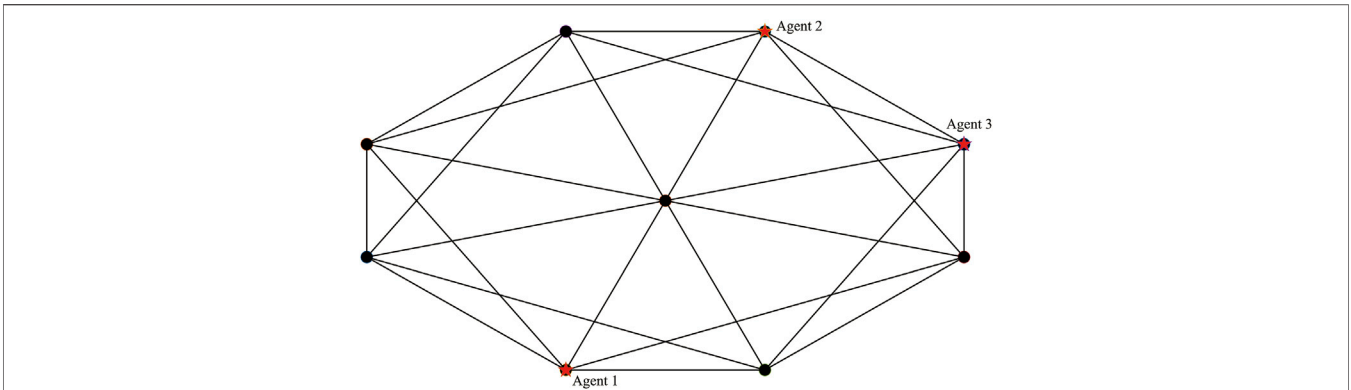
FIGURE 2 | Initial position of targets and agents (The black points represent the targets and the five-pointed stars represent the position of agents).
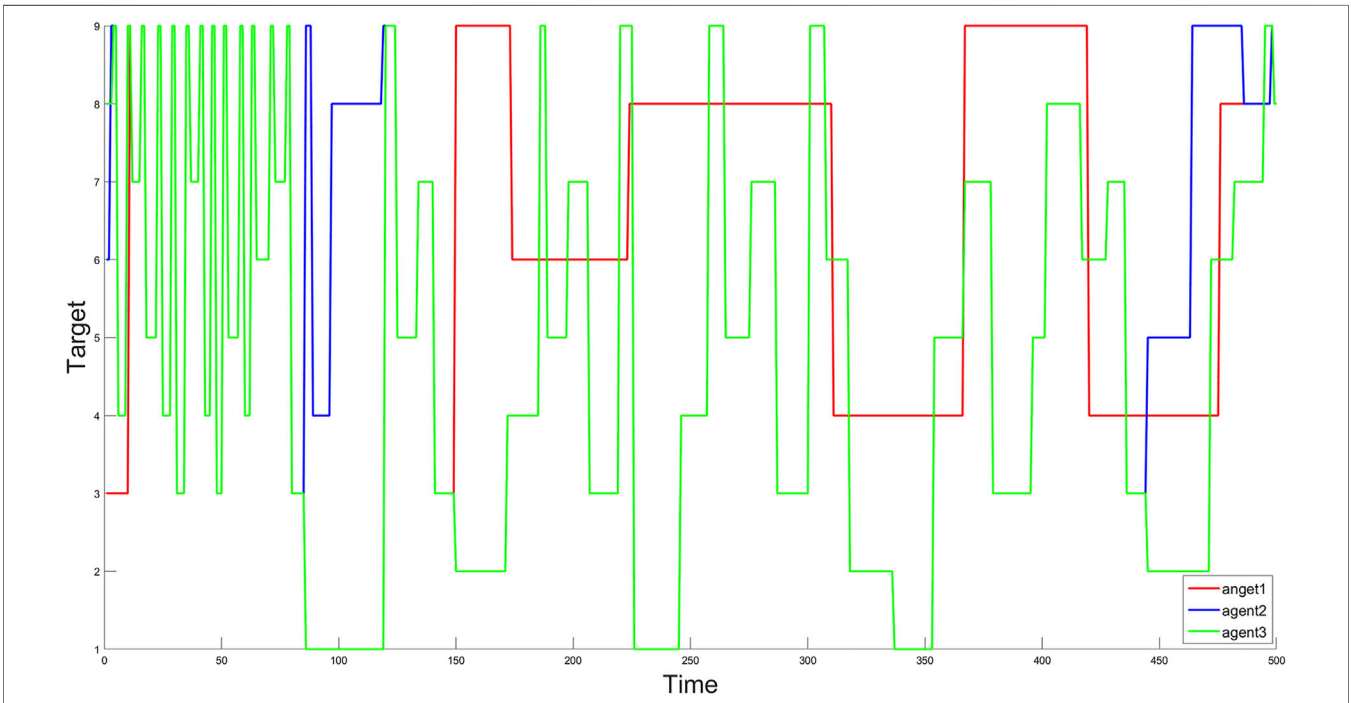
FIGURE 3 | The real-time monitoring targets of agents (1–9 on the *Y*-axes repersent the labels of targets, and the *X*-axes denotes the monitoring time *t*).

(b) $U_i^k(P \cup \{a\}) - U_i^k(P) = \frac{R_i^2(t_0)}{2} \frac{B_i r_a - \sum_{b \in P} B_i r_b}{(\sum_{b \in P} B_i r_b - A_i)(B_i r_a - A_i)}$,

thus the submodulariry of the utility function can be proved by simple comparison.

Lemma 3. For local optimization objective 5, the targets in the local set $\mathcal{N}_i^k$ are monitored by agents, and the local utility function $U_{\mathcal{N}_i}^k = \bar{J}_i^0(t_0^k, T) - \bar{J}_i^k(t_0^k, T)$ is nondecreasing and submodular, and $U_{\mathcal{N}_i}^k(\varnothing) = 0$, where $\bar{J}_i^0$ denotes the initial strategy of local monitoring tasks.

Proof. For submodularity, we consider $\tilde{P} \subset P \subset \Psi/\{a\}$, and $a \subset \Psi$. Define $\tilde{P}_1 = \tilde{P} \cup \{a\}$, $\tilde{P}_2 = P \cup \{a\}$ and $\pi = (a, i, u_i^k) \in \Psi \times H \times M$, where $u_i^k \in H$ for $i \in \mathcal{N}_i$ denotes the

possible dwell time set of each target visited by possible agents. Then

$$
\begin{aligned}
U_{\mathcal{N}_i}^k(P \cup \{a\}) - U_{\mathcal{N}_i}^k(P) &= J^k(t_0^k, T) - J_{\tilde{P}_2}(t_0^k, T) \\
&\quad - (J^k(t_0^k, T) - J_{\tilde{P}}(t_0^k, T)) \\
&= J_P(t_0^k, T) - J_{\tilde{P}_2}(t_0^k, T) \\
&= U_{\mathcal{N}_i}^k(P) - U_{\mathcal{N}_i}^k(\tilde{P}_2) \\
&\leq U_{\mathcal{N}_i}^k(\tilde{P}) - U_{\mathcal{N}_i}^k(\tilde{P}_1).
\end{aligned}
\tag{8}
$$

For monotonicity, for any $P \subset \Psi$ and $a \in \Psi$, $\Pi \subset \Psi \times H \times M$ and $\pi \in \Psi \times H \times M$, $U_{\mathcal{N}_i}^k(\Pi \cup \{\pi\}) - U_{\mathcal{N}_i}^k(\Pi) = U_{\mathcal{N}_i}^k(P \cup \{a\}) -$
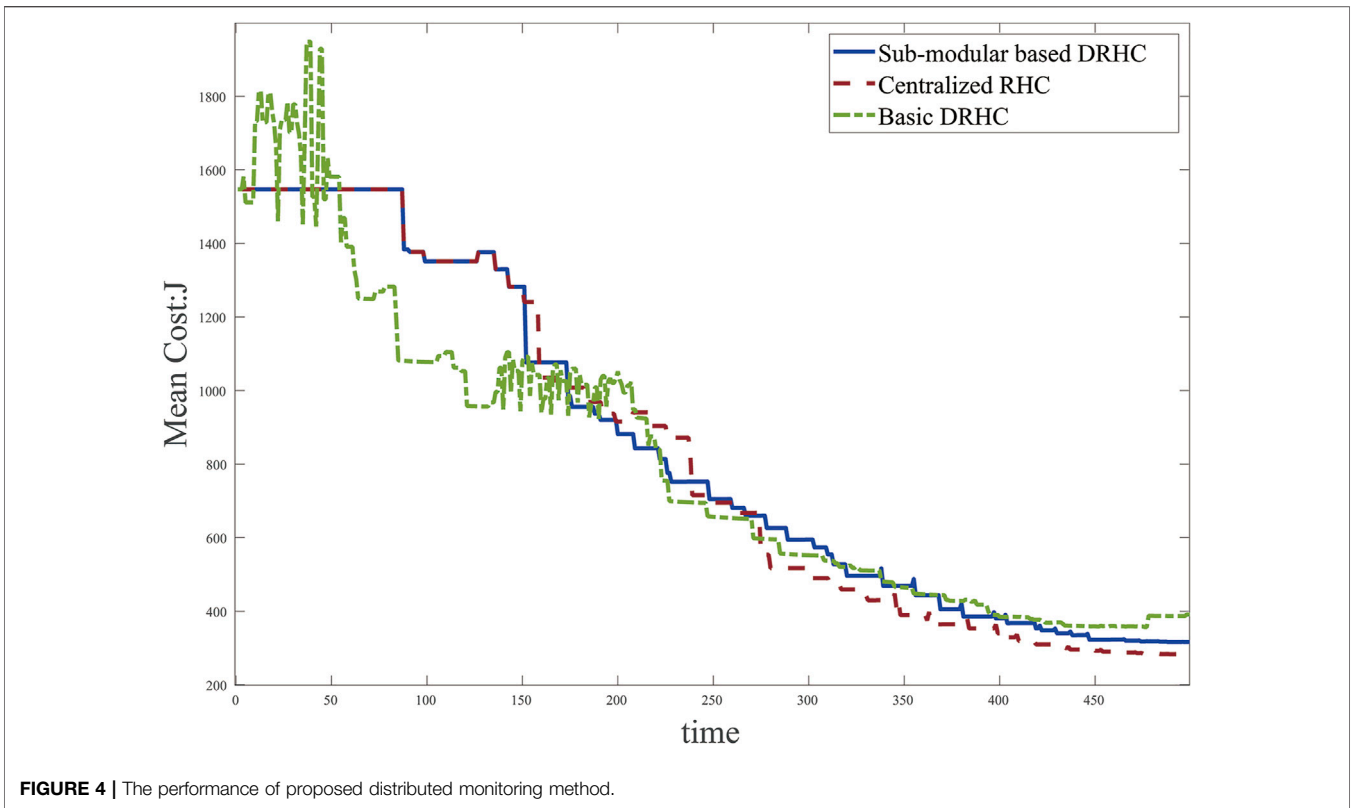
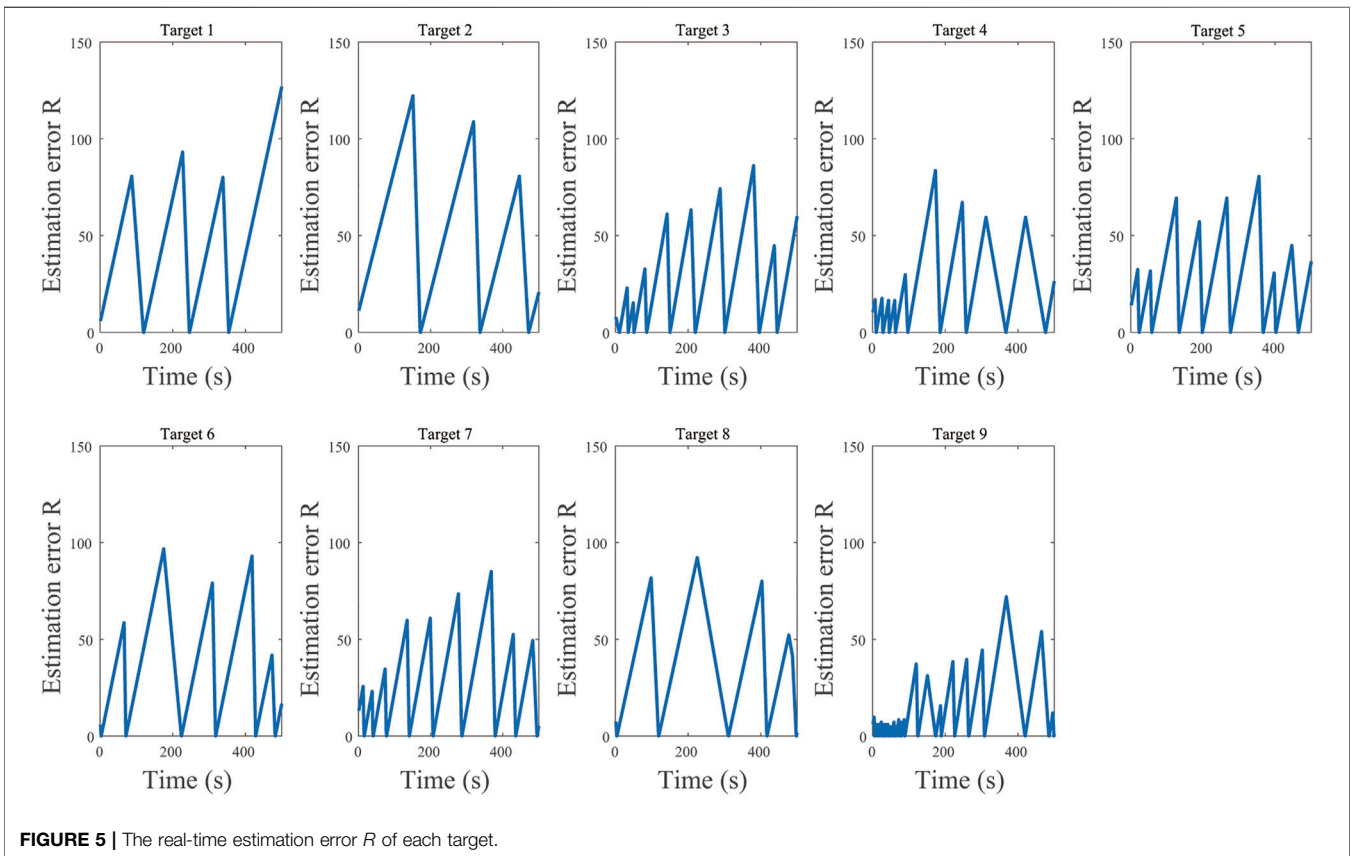**FIGURE 4 |** The performance of proposed distributed monitoring method.

FIGURE 5 | The real-time estimation error *R* of each target.

TABLE 1 | The mean cost and computing time of 200 simulations.

Method	Mean cost J	Computing time t of running 200 simulations
Sub-modular based DRHC	299.32	204.2676s
Basic DRHC	390.97	229.5898s
Centralized RHC	275.13	387.2561s

$U_{\mathcal{N}_i}^k(P) \geq 0$ due to Lemma 2. Obviously, $U_{\mathcal{N}_i}^k(\varnothing) = 0$ follows from $U_i^k(\varnothing) = 0$ for each $i \in M$.

Lemma 4. For global optimization objective 3, the global utility function $U^k = J_i^k(t_0^k, T) - J_i(t_0^k, T)$ is nondecreasing and submodular. Moreover, $U_{\mathcal{N}_i}^k(\varnothing) = 0$.

Proof. For submodularity, we consider $\tilde{P} \subset P \subset \Psi/\{a\}$, and $a \subset \Psi$. Define $\tilde{P}_1 = \tilde{P} \cup \{a\}, \tilde{P}_2 = P \cup \{a\}$, then

$$
\begin{aligned}
U^k(P \cup \{a\}) - U^k(P) &= J^k(t_0^k, T) - J_{\tilde{P}_2}(t_0^k, T) \\
&\quad - (J^k(t_0^k, T) - J_P(t_0^k, T)) \\
&= J_P(t_0^k, T) - J_{\tilde{P}_2}(t_0^k, T).
\end{aligned}
\tag{9}
$$

Since the cost of monitoring local targets is submodular by Lemma 3, it can be obtained that

$$
\begin{aligned}
J_P(t_0^k, T) - J_{\tilde{P}_2}(t_0^k, T) &= U^k(P) - U^k(\tilde{P}_2) \\
&\leq U^k(\tilde{P}) - U^k(\tilde{P}_1).
\end{aligned}
\tag{10}
$$

For monotonicity, for any $P \subset \Psi$ and $a \in \Psi$, $\Pi \subset \Psi \times H \times M$ and $\pi \in \Psi \times H \times M$, it can be obtained that $U^k(\Pi \cup \{\pi\}) - U(\Pi) =$

$U^k(P \cup \{a\}) - U^k(P) \geq 0$ due to Lemma 2 and Lemma 3. Obviously, $U^k(\varnothing) = 0$ follows from $U_i^k(\varnothing) = 0$ for each $i \in M$.

Remark 2. The analysis of the utility function of targets is equal to the objective 3. Lemma 2–4 proves that the designed utility function is a submodule function, it reduces the complexity of target performance analysis under a distributed receding horizon control strategy.

5 DISTRIBUTED GREEDY ALGORITHM

A distributed greedy algorithm is proposed in this section, which does not require synchronous actions of agents, meaning that at each visiting time t, an agent takes its action independently, and the motion of an agent is determined by the completion of the event of monitoring the current target. If the agent is going to move, then the agent will perform the decision-making procedures as shown in algorithm 1. Otherwise, the agent will continue to execute the previous strategy.

In algorithm 1, if there is more than one agent in the local set $\mathcal{A}$ to make decisions at the same time, only the one with high monitoring ability will be calculated, and the rest will make decisions at the next computing time in turn. This ensures that agents fully consider the behavior of other agents in the neighborhood to make more rational decisions.

Remark 3. It is possible that agents can not complet the next monitoring tasks in the remaining time t_r, considering that time T is limited. The situation that the remaining time t_r is less than the

---

Algorithm 1 | The distributed greedy strategy for Multi-agent persistent monitoring

1 **Input:** Intial target states and the position of each agent, the monitor time T.
2 **Output:** The strategy for the agent's next access target and the dwell time.
3 $\mathcal{N}_i \leftarrow$ Neighborhood target set of target i.
4 **On-line stage:** At time $t \in [0, T]$:
5 **for** *agent* $a \in \Psi$ **do**
6      $Z_{a,t} \leftarrow$ the current agent set that is located within the target set $\mathcal{N}_i$.
7   **if** *agent* a *is going to complete the current task at time* t_i **then**
8   **if** $t_i + \rho_{ij} \geq T, \; j \in \mathcal{N}_i$ **then**
9   stay at the current target.
10   **end**
11   **else**
12          $P_i \leftarrow$ the set of agents that have already selected i in previous time.
13   agent a calculates the dwell time and local cost of completing each target task in set $\mathcal{N}_{i,t}$, using function (4) and (5).
14          $\gamma_{j,a}^k \leftarrow U_i^k(\tilde{A}) \leq U_i^k(A)$ according to (6).
15   next visiting target $j \leftarrow argmax_{j \in \mathcal{N}_i} \gamma_{j,a}^k$.
16   send the strategy to the agents in $Z_{a,t}$.
17   **end**
18   **end**
19   agent a executes the current policy;
20 **end**

---

travel time ρ_{ij} to any next target is considered first. According to 5 and Lemma 1, the global estimation error of an agent going anywhere will be greater than stay at the current target. In this case, the optimal strategy of the agent is to stay at the current target. In another situation, the remaining time t_r can be less than the optimal dwell time. The local estimation error of agent staying at the current target is $\overline{J}_i = \sum\limits_{i \in \mathcal{N}_i} R_i t_r$, and the local estimation error of agent going to other target $j \in \mathcal{N}_i$ is $\overline{J}_j = \sum\limits_{j \in \mathcal{N}_i} R_j(t_1)\rho_{ij} + \sum\limits_{j \in \mathcal{N}_i} R_j(t_2)(t_r - \rho_{ij})$, where t_1, t_2 are the moments to visit and arrive the next target, respectively. In this case, the optimal strategy of an agent is to access the target corresponding to the smallest $\overline{J}$.

Assumption 3. Let $\kappa_\pi = U(\Pi \cup \pi) - U(\Pi)$, each agent a has to access target i at time t with a strategy $\pi = \{a, i, u_i^k\}$, suppose that for any $\pi \in \Pi - \overline{\Pi}$, there exist a constant $\nu \in (0, 1)$ that makes $\kappa_{\pi_2} > \nu\kappa_{\pi_1}$, where $\kappa_{\pi_1} = U(\Pi \cup \pi_1) - U(\Pi)$ and $\kappa_{\pi_2} = U(\Pi \cup \pi_1 \cup \pi_2) - U(\Pi \cup \pi_1)$.

Obviously, this assumption can be satisfied easily due to the submodularity of utility functions.

Theorem 1. The cost of constructing feasible strategies of agents by **Algorithm 1** is of order $O(\sum\limits_{a \in N} K_a D)$, where K_a denotes the total number of observed targets in time T, $N = max_{i \in M}\mathcal{N}_i$ and $D = max_{a \in N_i}\mathcal{A}_i$. Let $J^\star$ be an optimal solution of **(2)** and $\overline{J}$ be the output of **Algorithm 1**, then $J_i - J^\star < 2(J_0 - \overline{J})$, where J_0 represents the global cost when there is no agent monitoring targets.

Proof. The proposed algorithm can obtain a solution within limited computing resources is proved first. According to the definition of persistent monitoring tasks, agents moving to the next target will take ρ_{ij} seconds. For a finite time T, the number of events K_a would satisfy $K_a < T/(\rho_{ij})_{min}$, where $(\rho_{ij})_{min}$ denotes the minimum travel time in the graph. For each agent, at the kth decision, it needs to calculate the utility of a limited number of potential access targets in the neighborhood after being monitored. The number of calculations in this process is only determined by the number of targets in the neighborhood. That is, the agent only needs to make decisions $K_a D$ times in $t \in [0, T]$. Meanwhile, agent decisions are independent of each other. Communication between agents relies on neighborhood targets, which only occurs when the agent makes its next access decision. This kind of communication can be asynchronous each time a new access target is determined. This shows that for the global objective in 2, only $O(\sum\limits_{a \in \Psi} K_a D)$ complexity of communication and computing is needed to obtain a global solution at most.

The performance boundary of the algorithm is discussed next. Suppose $U^\star$ as the optimal utility to objective 2, and $U(\Pi)$ as the final solution based on the algorithm 1, then, it can be formed that $U(\Pi \cup \hat{\Pi}) \overset{(e)}{=} U(\hat{\Pi}) + \sum\limits_{\pi \in \{Pi - \hat{\Pi}\}} \kappa_\pi(\Pi)$. Considering the fact that $\Pi^\star = \cup_{a \in M, i \in \Psi}(a, i, u_i^k)$, $\sum\limits_{\pi \in \{Pi - \hat{\Pi}\}} \kappa_\pi(\Pi)$ can be redescribed as $\sum\limits_{i \in M} \sum\limits_{k=0}^{K} \sum\limits_{\pi \in \{Pi - \hat{\Pi}\}} \kappa_\pi(\Pi)$. From the submodularity of U

mentioned in Lemma 4, it follows that $U(\hat{\Pi}) + \sum\limits_{i \in M} \sum\limits_{k=0}^{K} \sum\limits_{\pi \in \{Pi - \hat{\Pi}\}} \kappa_\pi(\Pi) \leq U(\hat{\Pi}) + \sum\limits_{i \in M} \sum\limits_{k=0}^{K} \sum\limits_{\pi \in \{Pi - \hat{\Pi}\}} \kappa_\pi(\Pi_k)$, that is:

$$U(\Pi \cup \hat{\Pi}) \leq U(\hat{\Pi}) + \sum_{i \in M} \sum_{k=0}^{K} \sum_{\pi \in \{Pi - \hat{\Pi}\}} \kappa_\pi(\Pi_k). \tag{11}$$

Similarly, considering the submodularity bound of U defined in Assumption 3,

$$\begin{aligned} U(\Pi \cup \hat{\Pi}) &= U(\Pi^\star) + \sum_{\pi \in \{Pi - \hat{\Pi}\}} \kappa_\pi(\Pi) \\ &= U(\Pi^\star) + \sum_{i \in M} \sum_{k=0}^{K} \sum_{\pi \in \{Pi - \hat{\Pi}\}} \kappa_\pi(\Pi) \\ &\geq U(\Pi^\star) + \nu \sum_{i \in M} \sum_{k=0}^{K} \sum_{\pi \in \{Pi - \hat{\Pi}\}} \kappa_\pi(\Pi_k). \end{aligned} \tag{12}$$

Combine **Eqs 11, 12**, then

$$\begin{aligned} U(\Pi^\star) &\leq U(\hat{\Pi}) + (1 - \nu) \sum_{i \in M} \sum_{k=0}^{K} \sum_{\pi \in \{Pi - \hat{\Pi}\}} \kappa_\pi(\Pi_k) \\ &\leq U(\hat{\Pi}) + (1 - \nu)U(\hat{\Pi}) \\ &\leq 2U(\hat{\Pi}). \end{aligned} \tag{13}$$

According to the utility function defined in 4, $J_i - J^\star < 2(J_0 - \overline{J})$, which means that $\frac{J_0 - \overline{J}}{J_0 - J^\star} \geq \frac{1}{2}$ and the proof is concluded.

Remark 4. At each time t, if the agent completes monitoring of the current target, which means $R(t) = 0$, the agent will immediately move to the next target. It's the key to ensuring the efficiency guarantee of the algorithm. **Algorithm 1** returns feasible strategies of agents when they visit the next targets each time, making objective J decrease continuously.

Remark 5. The monitoring performance is affected by the number of targets and agents, the monitoring and movement ability of agents, and the distance between targets, etc. As the number of targets increases in a persistent monitoring task with a given number of agents, the monitoring performance of the proposed method could be worse. Thus, the computation burden could be heavier. Theorem 1 proves that the monitoring performance of the proposed method is close to the result of the centralized method, even if the graph size increases.

6 SIMULATION

Consider three agents equipped with different monitor capabilities, they are used to move between nine targets to minimize the total estimation error in a given graph. The position of targets are [2, 1], [2, 2], [3, 0], [3, 3], [4, 0], [4, 3], [5, 1], [5, 2], [3.5], [1.5], and the initial position of each agent are at target 3, 6, and 8, as shown in **Figure 2**. The dynamic set of targets is defined as A = [0.87; 0.74; 1.09; 1.1; 1.02; 0.93; 1.05; 0.87; 1.2]; B = [1.69; 1.95; 1.95; 2.18; 2.37; 2.93; 2.60; 1.95; 2.60]; And the ability set of agents is r = [1; 1.5; 2]; The initial

uncertainty set of targets is R_0 = [5; 10.5; 8.7; 9.2; 13; 9.4; 12; 10.5; 5]; The overall time period is T = 500 and the time step is 1. The resluts are shown in **Figures 3–5**.

Figure 3 provides the real-time monitoring targets of the agents during $t \in [0, 500]$. Agents perform monitoring tasks in different sequences after starting from their initial targets. The dwell time of the agent at each target is represented by the length of the horizontal line, and the vertical axes represents different targets. As described above, the agent can switch between different targets by only using limited local information. In the process of monitoring tasks, agents may obtain different strategies at the same target, which is mainly affected by the states of other target points in the neighborhood and the monitoring states of agents in the neighborhood.

The result of this simulation is shown in **Figure 4**, in comparison with the centralized receding horizon control method and the basic distributed receding horizon control strategy designed in Welikala and Cassandras (2021). From the observed results, the performance of proposed sub-modular based DRHC method is closer to the centralized method. The optimal result of the centralized RHC approach is an upper bound of the proposed distributed method, which is a trade-off for monitoring performance and calculating costs.

The real-time estimation errors for each target are displayed in **Figure 5**. The estimation error of each target increases continuously until the agents arrive, and decreases to 0 when agents dwell at the current target.

To further illustrate the effect of the proposed method, 200 times simulations are run to obtain an average cost, aiming to reduce the impact of different initial positions of agents on monitoring performance. The results are shown in **Table 1**. It can be concluded that the proposed method can converge to a bounded global performance while costing less computing time.

7 CONCLUSION

In this paper, a submodular receding horizon control strategy is proposed for distributed multi-agent persistent monitoring problem, aiming to obtain a guaranteed bound monitoring performance with the finite number of calculations. Considering the submodular characteristics of the monitoring objective function, the global monitoring objective is decomposed into multiple optimization problems based on the receding horizon control strategy, and the optimal dwell time of each agent in a target is determined first. The utility function of an agent monitoring targets is proved to be submodular, based on which a distributed greedy algorithm is proposed to obtain optimal strategies for agents to persistent monitor targets. In particular, the analysis is presented to prove each agent can only use limited computing resources to obtain a guaranteed performance. Numerical results show that the proposed method has better stability and convergence speed. Future work will pay more attention to the performance assurance of the proposed algorithm in stochastic environments.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding authors.

AUTHOR CONTRIBUTIONS

XZ is responsible for the conceptualization, the simulation and the writing of this paper. YZ and SL are responsible for the funding acquisition, the conceptualization of the main idea and the revisions of the paper.

FUNDING

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fcteg.2021.786877/full#supplementary-material

REFERENCES

Alam, T., Reis, G. M., Bobadilla, L., and Smith, R. N. (2018). "A Data-Driven Deployment Approach for Persistent Monitoring in Aquatic Environments," in 2018 Second IEEE International Conference on Robotic Computing (IRC) (IEEE), 147–154. doi:10.1109/irc.2018.00030

Asghar, A. B., Smith, S. L., and Sundaram, S. (2019). "Multi-Robot Routing for Persistent Monitoring with Latency Constraints," in 2019 American Control Conference (ACC) (IEEE), 2620–2625. doi:10.23919/acc.2019.8814485

Cassandras, C. G., Lin, X., and Ding, X. (2012). An Optimal Control Approach to the Multi-Agent Persistent Monitoring Problem. *IEEE Trans. Automatic Control.* 58, 947–961. doi:10.1109/cdc.2012.6426992

Chen, J., Baskaran, A., Zhang, Z., and Tokekar, P. (2020). Multi-Agent Reinforcement Learning for Persistent Monitoring. arXiv [Preprint]. Available at: https://arxiv.org/abs/2011.01129

Deng, K., Chen, Y., and Belta, C. (2017). An Approximate Dynamic Programming Approach to Multiagent Persistent Monitoring in Stochastic Environments with Temporal Logic Constraints. *IEEE Trans. Automat. Contr.* 62, 4549–4563. doi:10.1109/tac.2017.2678920

Gharesifard, B., and Smith, S. L. (2017). Distributed Submodular Maximization with Limited Information. *IEEE Trans. Control Netw. Syst.* 5, 1635–1645. doi:10.1109/TCNS.2017.2740625

Ha, J.-S., and Choi, H.-L. (2019). On Periodic Optimal Solutions of Persistent Sensor Planning for Continuous-Time Linear Systems. *Automatica.* 99, 138–148. doi:10.1016/j.automatica.2018.10.005

Hari, S. K. K., Rathinam, S., Darbha, S., Kalyanam, K., Manyam, S. G., and Casbeer, D. (2020). Optimal Uav Route Planning for Persistent Monitoring Missions. *IEEE Trans. Robotics.* 37, 550–566. doi:10.1109/TRO.2020.3032171

Lan, X., and Schwager, M. (2013). "Planning Periodic Persistent Monitoring Trajectories for Sensing Robots in Gaussian Random Fields," in 2013 IEEE International Conference on Robotics and Automation (IEEE), 2415–2420. doi:10.1109/icra.2013.6630905

Lee, J., Kim, G., and Moon, I. (2021). A mobile Multi-Agent Sensing Problem With Submodular Functions Under a Partition Matroid. *Comput. Operations Res.* 132, 105265. doi:10.1016/j.cor.2021.105265

Li, H., and Li, X. (2020). Distributed Model Predictive Consensus of Heterogeneous Time-Varying Multi-Agent Systems: With and Without Self-Triggered Mechanism. *IEEE Trans. Circuits Syst.* 67, 5358–5368. doi:10.1109/TCSI.2020.3008528

Liu, Y., Liu, H., Tian, Y., and Sun, C. (2020). Reinforcement Learning Based Two-Level Control Framework of Uav Swarm for Cooperative Persistent Surveillance in an Unknown Urban Area. *Aerospace Sci. Technology.* 98, 105671. doi:10.1016/j.ast.2019.105671

Mackin, E., and Patterson, S. (2018). Submodular Optimization for Consensus Networks with Noise-Corrupted Leaders. *IEEE Trans. Automatic Control.* 64, 3054–3059. doi:10.1109/TAC.2018.2874306

Maini, P., Tokekar, P., and Sujit, P. (2020). Visibility-based Persistent Monitoring of Piecewise Linear Features on a Terrain Using Multiple Aerial and Ground Robots. *IEEE Trans. Automation Sci. Eng.* 18 (4), 1692–1704. doi:10.1109/TASE.2020.3014949

Mi, X., Zou, Y., and Li, S. (2018). Event-triggered Mpc Design for Distributed Systems Toward Global Performance. *Int. J. Robust Nonlinear Control.* 28, 1474–1495. doi:10.1002/rnc.3969

Nigam, N. (2014). The Multiple Unmanned Air Vehicle Persistent Surveillance Problem: A Review. *Machines.* 2, 13–72. doi:10.3390/machines2010013

Ostertag, M., Atanasov, N., and Rosing, T. (2019). "Robust Velocity Control for Minimum Steady State Uncertainty in Persistent Monitoring Applications,"2019 American Control Conference (ACC), 2501–2508. doi:10.23919/acc.2019.8814376

Pinto, S. C., Andersson, S. B., Hendrickx, J. M., and Cassandras, C. G. (2020). "Optimal Periodic Multi-Agent Persistent Monitoring of a Finite Set of Targets with Uncertain States," in 2020 American Control Conference (ACC) (IEEE), 5207–5212. doi:10.23919/acc45564.2020.9147536

Rajkumar, R., Lee, I., Sha, L., and Stankovic, J. (2010). "Cyber-physical Systems: The Next Computing Revolution," in Design automation conference (IEEE), 731–736.

Rezazadeh, N., and Kia, S. S. (2021). A Sub-modular Receding Horizon Solution for Mobile Multi-Agent Persistent Monitoring. *Automatica.* 127, 109460. doi:10.1016/j.automatica.2020.109460

Smith, S. L., Schwager, M., and Rus, D. (2011). "Persistent Monitoring of Changing Environments Using a Robot with Limited Range Sensing," in 2011 IEEE International Conference on Robotics and Automation (IEEE), 5448–5455. doi:10.1109/icra.2011.5980251

Wang, T., Huang, P., and Dong, G. (2020). Modeling and Path Planning for Persistent Surveillance by Unmanned Ground Vehicle. *IEEE Trans. Automation Sci. Eng.* 18 (4), 1615–1625. doi:10.1109/TASE.2020.3013288

Welikala, S., and Cassandras, C. G. (2021). Event-driven Receding Horizon Control for Distributed Persistent Monitoring in Network Systems. *Automatica.* 127, 109519. doi:10.1016/j.automatica.2021.109519

Yu, J., Karaman, S., and Rus, D. (2015). Persistent Monitoring of Events with Stochastic Arrivals at Multiple Stations. *IEEE Trans. Robot.* 31, 521–535. doi:10.1109/tro.2015.2409453

Zeng, J., and Liu, J. (2015). Distributed Moving Horizon State Estimation: Simultaneously Handling Communication Delays and Data Losses. *Syst. Control. Lett.* 75, 56–68. doi:10.1016/j.sysconle.2014.11.007

Zhao, M., Wei, Y., and Xiao, J. (2018). "An Optimal Control Approach to the Multi-Agent Persistent Monitoring with Weight Coefficient," in 2018 Chinese Automation Congress (CAC), 358–363. doi:10.1109/cac.2018.8623586

Zhou, N., Cassandras, C. G., Yu, X., and Andersson, S. B. (2020). The price of Decentralization: Event-Driven Optimization for Multi-Agent Persistent Monitoring Tasks. *IEEE Trans. Control. Netw. Syst.* 8 (2), 976–986. doi:10.1109/TCNS.2020.3047314

Zhou, N., Yu, X., Andersson, S. B., and Cassandras, C. G. (2018). Optimal Event-Driven Multiagent Persistent Monitoring of a Finite Set of Data Sources. *IEEE Trans. Automat. Contr.* 63, 4204–4217. doi:10.1109/tac.2018.2829469

Zou, Y., Su, X., Li, S., Niu, Y., and Li, D. (2019). Event-triggered Distributed Predictive Control for Asynchronous Coordination of Multi-Agent Systems. *Automatica.* 99, 92–98. doi:10.1016/j.automatica.2018.10.019