2021 American Control Conference (ACC)
New Orleans, USA, May 25-28, 2021

Optimal Minimax Mobile Sensor Scheduling Over a Network

Samuel C. Pinto[1], Sean B. Andersson[1,2], Julien M. Hendrickx[3], and Christos G. Cassandras[2,4]
[1]Dept. of Mechanical Engineering, [2]Division of Systems Engineering,
[4]Dept. of Electrical and Computer Engineering
Boston University, Boston, MA 02215, USA
[3]ICTEAM Institute, UCLouvain, Louvain-la-Neuve 1348, Belgium
{samcerq,sanderss,cgc}@bu.edu, julien.hendrickx@uclouvain.be

***Abstract*— We investigate the problem of monitoring multiple targets using a single mobile sensor, with the goal of minimizing the maximum estimation error among all the targets over long time horizons. The sensor can move in a network-constrained structure, where it has to plan which targets to visit and for how long to dwell at each node. We prove that in an optimal observation time allocation, the peak uncertainty is the same among all the targets. By further restricting the agent policy to only visit each target once every cycle, we develop a scheme to optimize the agent's behavior that is significantly simpler computationally when compared to previous approaches for similar problems.**

I. INTRODUCTION

In many different applications, such as ocean temperature monitoring [1] or surveillance in smart cities [2], a mobile agent equipped with sensors is responsible for monitoring the state of a system at multiple points of interest. In this setting, a usual goal is to minimize some metric of uncertainty of the monitored variables over long time horizons. We call this problem "persistent monitoring" (PM). In terms of design, a common approach (see e.g. [3] and [4]) is to consider the agent policy as being composed of two interdependent tasks: searching for an optimal agent trajectory and planning for how long to observe each of the targets. In this work, we model the PM problem as one of an agent moving on a network [4] and seek a periodic trajectory that ensures each target is visited, while optimizing a given uncertainty metric over long time horizons. PM is closely related to the well known Traveling Salesman Problem (TSP), as both problems try to minimize the time spent traveling between different targets, and some connection between persistent surveillance and TSP have already been explored e.g. in [5]. The main distinction between these problems is that in PM one has to plan for how long to dwell at each target in order to minimize a metric of the targets state uncertainties (instead of simply the distance traveled).

In our previous work [6], [7], we investigated the multi-agent version of this problem, with the goal of finding locally optimal movement agent policies that minimized the mean squared estimation error. One major limitation of this previous work was that each step of the optimization (which was a gradient descent procedure) relied on the numerical integration of matrix ordinary differential equations in order to compute the gradients. Integrating these equations at each step of the iterative optimization process is a major computational burden and makes the approach in [6] unsuitable for settings with a large number of targets.

In the present paper, we constrain ourselves to considering a single agent in order to establish stronger results than in previous publications, with the goal of generalizing these stronger results to multi-agent settings in future work. Moreover, instead of minimizing the mean squared estimation error, we consider a different optimization metric, the minimization of the maximum uncertainty over different targets. In addition to simplifying the computational burden, this metric may be preferred in some settings. For example, when monitoring safety critical systems that cannot operate over a given threshold (for instance a maximum temperature), one wants to make sure that every target location does not exceed that "worst case" limit (as opposed to optimizing an "average" of the chance of violating it). Some examples of applications where a critical threshold on the state uncertainty should not be exceeded include monitoring of wildfire or faults in civil infrastructure systems using unmanned aerial vehicles [8], [9]. Considering this cost function, the main contribution of this paper is to show that, in an optimal schedule all the targets have the same peak uncertainty. A parallel can be drawn between this problem and resource allocation problems, where we often find results showing that all the users have the same utility in the optimal schedule (see e.g. [10]). In this paper, the resource to be allocated is the time for which each target is observed. We exploit this necessary condition and introduce a novel gradient-free and computationally inexpensive scheme to optimize the time spent at each target. For visiting sequences where each target is visited only once, this approach provably gives the optimal time that should be spent at each target for a fixed cycle period T. This is especially meaningful in scenarios where targets are heterogeneous and, in order to ensure a balanced operation of the entire network of targets, they must be observed for significantly different times.

This work was supported in part by NSF under grants ECCS-1931600, DMS-1664644, CNS-1645681, and CMMI-1562031, by ARPA-E under grant DE-AR0001282, by AFOSR under grant FA9550-19-1-0158, and by the MathWorks. The work of J. Hendrickx was supported by the "RevealFlight" Concerted Research Action (ARC) of the Federation Wallonie-Bruxelles, by the Incentive Grant for Scientific Research (MIS) "Learning from Pairwise Comparisons" of the F.R.S.-FNRS.

II. PROBLEM FORMULATION

We consider a set $\mathcal{V}$ composed of M nodes (which are interchangeably called targets) in a connected and undirected graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$. The set $\mathcal{E}$ is formed by edges $\{(i,j) : i,j \in \mathcal{V}\}$. Each edge has an associated cost $d_{i,j}$, representing the time the agent takes to travel between nodes i and j.

We also assume that each target has an internal state $\phi_i \in \mathbb{R}^{L_i}$ with dynamics

$$\dot{\phi}_i(t) = A_i\phi_i(t) + B_i u_i(t) + w_i(t), \quad i = 1, \dots, M, \quad (1)$$

where w_i are mutually independent, zero mean, white, Gaussian distributed processes with $E[w_i(t)w_i(t)'] = Q_i$ and Q_i is a positive definite matrix. Since we focus on the estimation problem, we do not discuss the design of the control $u_i(t)$, but we assume that the estimation algorithm has access to the values of $u_i(t)$.

The mobile agent can navigate through the graph $\mathcal{G}$. When the agent visits a given node i, it observes its internal state with a linear observation model given by

$$z_i(t) = H_i\phi_i(t) + v_i(t), \qquad v_i(t) \sim \mathcal{N}(0, R_i), \quad (2)$$

where v_i is also assumed to be white and statistically independent of v_j if $i \neq j$ and w_k, $\forall k$.

We can fully describe the behavior of the agent at its N visits by a vector of indices of the targets visited by the agent $\mathcal{Y} = [y_1, y_2, ..., y_N]$, $1 \leq y_k \leq M$, and the time spent by the agent at each of these visits $\mathcal{T} = [t_1, ..., t_N]$, $t_k \geq 0$. Note that even though in general $\mathcal{Y}$ and $\mathcal{T}$ can be infinitely long, in this paper we restrict ourselves to periodic policies (as will be discussed in the next subsection) and thus N is finite.

Considering models (1) and (2), the maximum likelihood estimator $\hat{\phi}_i(t)$ for the internal state of the targets is a Kalman-Bucy Filter with equations given by

$$\dot{\hat{\phi}}_i(t) = A_i\hat{\phi}_i(t) + B_i u_i(t) + \eta_i(t)\Omega_i(t)H_i'R_i^{-1}\left(z_i(t) - H_i\hat{\phi}_i(t)\right), \quad (3a)$$

$$\dot{\Omega}_i(t) = A_i\Omega_i(t) + \Omega_i(t)A_i' + Q_i - \eta_i(t)\Omega_i(t)G_i\Omega_i(t), \quad (3b)$$

where Ω_i is the covariance matrix of the estimator $\hat{\phi}_i(t)$, $G_i = H_i'R_i^{-1}H_i$ and $\eta_i(t) = 1$ if target i is observed at time t and $\eta_i(t) = 0$ otherwise.

The goal is to design an agent movement policy that minimizes the maximum (over time and over all the targets) of a weighted norm of the long term covariance matrix, i.e.,

$$J(\mathcal{Y}, \mathcal{T}) = \max_{i \in \{1,..,M\}} \limsup_{t \to \infty} g_i(\|\Omega_i(t)\|), \quad (4)$$

where the target-specific (possibly) nonlinear weighting function $g_i(\cdot)$ is a strictly increasing function with $g_i(0) = 0$ and $\lim_{x\to\infty} g_i(x) = \infty$, and $\|\cdot\|$ is a norm on the space of positive semi-definite matrices. A usual choice is to have $g_i(x) = x$ and $\|X\| = \text{tr}(X)$. An optimal strategy has an associated cost J^* given by

$$J^*(\mathcal{Y}^*, \mathcal{T}^*) = \min_{\mathcal{Y},\mathcal{T}} J(\mathcal{Y}, \mathcal{T}). \quad (5)$$

A. Periodic Policies

In this work we restrict ourselves to periodic policies. Note that in periodic trajectories, the number of parameters necessary to describe the behavior of the agent does not increase with the time horizon. They are therefore suitable for infinite horizon analysis. Moreover, if a target is visited in the period, it will be visited infinitely often, with inter-visit time upper bounded by the period. This notion fits very well into the paradigm of persistent monitoring, since one of the high level goals is to ensure this persistence of visits.

In order to analyze the infinite horizon behavior of the covariance matrix, we initially make the following natural assumption that ensures the uncertainty of the internal state ϕ_i can be finite over long time horizons.

Assumption 1: The pair (A_i, H_i) is detectable, $\forall\ i \in \{1, ..., M\}$.

We now revisit the following Proposition, initially introduced in [6]:

Proposition 1: If $\eta_i(t)$ is T-periodic and $\eta_i(t) > 0$ for t in some interval $[a, b] \in [0, T]$ with $b > a$, then, under Assumption 1, there exists a unique non-negative stabilizing T-periodic solution $\bar{\Omega}_i(t)$ to (3b).

The consequence of this proposition is that, in a periodic schedule, visiting a given target for any finite amount of time is enough to guarantee that the covariance will converge to a limit cycle that does not depend on the initial conditions. In this paper we will focus on cases where A_i is unstable, since it requires targets to be visited infinitely often to ensure a bounded cost, and thus fits well into the persistent monitoring framework. Note that in some real applications, such as the surveillance of wildfires or the expansion of cracks in civil infrastructure [8], [9], the dynamics of the state are often unstable. In addition, to avoid the degenerate case where Q_i is semidefinite such that covariance could potentially stay null in some direction, we assume that Q_i is positive definite. These assumptions are now formalized:

Assumption 2: A_i is not stable and Q_i is positive definite.

Remark 1: The optimization approach we develop in this paper could be easily adapted to also handle targets with stable A_i. However, if A_i is allowed to be stable, it could be the case that the optimal action would be never to observe such a target, since its uncertainty would be bounded even if the agent never visits it. In this paper, due to space constraints, we mainly focus on the optimization of dwelling times $\mathcal{T}$ and by assuming that A_i is unstable, we have that the set of visited targets is fixed. In future work we plan to consider both stable and unstable A_i, along with criteria for the inclusion or exclusion of targets in the visiting sequence. A consequence of Assumption 2 is that every target should be visited in one period of the cycle, however the optimal visiting order is still to be determined.

III. PROPERTIES OF AN OPTIMAL POLICY

A. Target's Perspective of a Periodic Policy

We begin by defining some notations used in the remainder of this section. We recall that the goal is to optimize the

visiting sequence $\mathcal{Y}$ and dwelling times at the visited targets $\mathcal{T}$. One way to interpret the vectors $\mathcal{Y}$ and $\mathcal{T}$ is that they capture the agent perspective of the trajectory. Each individual target, however, does not need all this information in order to fully determine the behavior of its uncertainty. From the target's perspective, all the information necessary to characterize the evolution of its covariance matrix is to define, for a single cycle of the periodic schedule, how long the target is observed for and the time spent between consecutive observations. Without loss of generality, we assume that the length N of the vectors $\mathcal{Y}$ and $\mathcal{T}$ is one full cycle of the periodic trajectory.

However, from a given target's perspective, all that is necessary to fully specify its steady-state covariance matrix is to define for how long it was observed and how much time was spent between consecutive observations (i.e. the value of $\eta_i(t)$ in (3b)). This is illustrated in Fig. 1, where $t^k_{\text{on},i}$ is the amount of time that target i was observed when it was visited for the k-th time. On the other hand, $t^k_{\text{off},i}$ is the time spent between its k-th and $k+1$-st visits, with k ranging from 1 to N_i, where N_i is the total number of visits by the agent at target i in one cycle. The peak covariance values are denoted by $\overline{P}^k_i$.

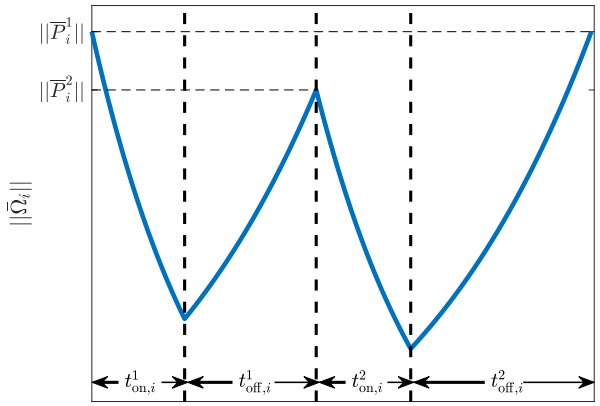


Fig. 1: Temporal evolution of the steady state covariance matrix and waiting/observation times.

We highlight that $t^k_{\text{on},i}$ and $t^k_{\text{off},i}$ can be easily computed from $\mathcal{Y}$ and $\mathcal{T}$. The explicit equations for this conversion are given in [11] but omitted here, since they are not essential for the understanding of the paper and involve complex notation. Analogously, the peaks $\overline{P}^k_i$ are defined as the steady state covariance $\bar{\Omega}_i(t')$ at a specific time t' computed from $\mathcal{Y}$ and $\mathcal{T}$, and details are given in [11].

B. Necessary Condition for Optimality

In this section, our main goal is to show that, for every visiting sequence $\mathcal{Y}$ that contains every target, an optimal allocation of dwelling times $\mathcal{T}$ must be such that $\limsup_{t\to\infty} g_i(\|\Omega_i(t)\|)$ is the same among all the targets. Towards this main goal of this section, first we claim the following auxiliary results.

Lemma 1: $\dot{\bar{\Omega}}_i(t) \prec 0$ if $\eta_i(t) = 1$ and $\dot{\bar{\Omega}}_i(t) \succ 0$ if $\eta_i(t) = 0$.

The interpretation of this lemma is very intuitive: the steady state covariance decreases under sensing and increases otherwise. The proof is a direct application of Theorem 2 in [12] and is omitted here for space reasons but given in [11].

Lemma 2: If target i is visited for a non degenerate amount of time, then $\limsup_{t\to\infty} g_i(\|\Omega_i(t)\|) = \max_{1\le k\le N_i} g_i\left(\left\|\overline{P}^k_i\right\|\right)$

Proof: First, note that since $\Omega_i(t)$ converges to the bounded periodic function $\bar{\Omega}_i(t)$, $\limsup_{t\to\infty} g_i(\|\Omega_i(t)\|) = \max_{0\le t\le T} g_i\left(\left\|\bar{\Omega}_i(t)\right\|\right)$. For any time t for which $\exists\ \epsilon > 0$ such that $\eta_i(t+\epsilon) = 0$, Lemma 1 implies that $\bar{\Omega}_i(t+\epsilon) \succ \bar{\Omega}_i(t)$. Conversely, if $\exists\ \epsilon > 0\ \eta_i(t-\epsilon) = 1$, $\bar{\Omega}_i(t-\epsilon) \succ \bar{\Omega}_i(t)$.

Therefore, the maximum value of $g_i\left(\left\|\bar{\Omega}_i\right\|\right)$ can only occur in one of the instants when the target switches from not being observed ($\eta_i = 0$) to being observed ($\eta_i = 1$). The covariance at these instants is given by $\overline{P}^k_i$. ■

Using the results of the two lemmas just established, we now show how the upper peak values $\overline{P}_i$ vary with $t_{\text{on},i}$ and $t_{\text{off},i}$. Note that here we consider the steady state covariance as a function exclusively of $t_{\text{on},i}$ and $t_{\text{off},i}$, since these parameters fully define the steady state behavior of the covariance matrix. Additionally, the proofs of the next two propositions are technical and are omitted here; they can be found in [11]. Their interpretations, however, are intuitive: when a given target is observed for a longer time, its peak uncertainty will be lower. Conversely, if the time between observations increases, then the peak uncertainties will be higher.

Proposition 2: $\frac{\partial \overline{P}^k_i}{\partial t^m_{\text{on},i}} \prec 0$.

Proposition 3: $\frac{\partial \overline{P}^k_i}{\partial t^m_{\text{off},i}} \succ 0$.

Now that we have established some properties of the variation of the peaks of the steady state covariance matrix, we can explore the main result of this section. The next proposition can be interpreted analogously to resource allocation problems, where different targets are competing for the same resource $t^k_{\text{on},i}$. Therefore, an equilibrium (in the minimax sense) is reached when all the targets have the same utility. However, unlike typical resource allocation problems, here the total resource $\sum_{i=1}^{M}\sum_{k=1}^{N_i} t^k_{\text{on},i}$ is not fixed. The reason why the total resource does not go to infinity is that increasing $t^k_{\text{on},i}$ to one target has an adverse effect to all other targets.

Proposition 4: Assuming a fixed sequence of visiting targets $\mathcal{Y}$, in an optimal allocation of visiting times $\mathcal{T}$ according to the cost (4), the following must hold:

$$\limsup_{t\to\infty} g_i(\|\Omega_i(t)\|) = \limsup_{t\to\infty} g_j(\|\Omega_j(t)\|),$$

with $i,\ j \in \{1, ..., M\}$.

Proof: We prove by contradiction, showing that if the property given in the proposition does not hold, then there is a way to re-balance the observation times that is guaranteed to improve the performance. Suppose that for some target i

$$g_i\left(\left\|\overline{P}^{\max}_i(t^{1:N_i}_{\text{on},i}, t^{1:N_i}_{\text{off},i})\right\|\right) < g_j\left(\left\|\overline{P}^{\max}_j(t^{1:N_j}_{\text{on},j}, t^{1:N_j}_{\text{off},j})\right\|\right), \quad (6)$$

where the upper index $\max$ indicates that among all the peaks P_i^k, we pick the value of k that yields the highest $g_i(\|P_i^k\|)$. We now propose to decrease the amount of time target i is observed by ϵ and thus the waiting time between observations for all the other targets will also decrease. This updated policy generates a new set of observation times for target i, ($\tilde{t}_{\text{on},i}^{1:N_i}$), and updated waiting times between visits for all the other targets, ($\tilde{t}_{\text{off},j}^{1:N_j}$), while maintaining $t_{\text{off},i}^{1:N_i}$ and $t_{\text{on},j}^{1:N_j}$ constant. There exists $\epsilon > 0$ such that $\tilde{t}_{\text{on},i}^{k} = t_{\text{on},i}^{k} - \epsilon$ for some $k \in \{1, .., N_i\}$ and $\tilde{t}_{\text{off},j}^{m} < t_{\text{off},j}^{m}$ for some $m \in \{1, ..., N_j\}$. Using Props. 2 and 3 we get that

$$\overline{P}_i^{\max}(\tilde{t}_{\text{on},i}^{1:N_i}, t_{\text{off},i}^{1:N_i}) \succ \overline{P}_i^{\max}(t_{\text{on},i}^{1:N_i}, t_{\text{off},i}^{1:N_i}), \tag{7}$$

$$\overline{P}_j^{\max}(t_{\text{on},j}^{1:N_j}, \tilde{t}_{\text{off},j}^{1:N_j}) \prec \overline{P}_j^{\max}(t_{\text{on},j}^{1:N_j}, t_{\text{off},j}^{1:N_j}). \tag{8}$$

Using the fact that both the norm and the derivative are continuous and strictly increasing, we can always pick an ϵ small enough such that the new peak of target i is lower or equal to the new peak of j, i.e.,

$$g_j\left(\left\|\overline{P}_j^{\max}(t_{\text{on},j}^{1:N_j}, \tilde{t}_{\text{off},j}^{1:N_j})\right\|\right) < g_j\left(\left\|\overline{P}_j^{\max}(t_{\text{on},j}^{1:N_j}, t_{\text{off},j}^{1:N_j})\right\|\right), \tag{9}$$

$$g_i\left(\left\|\overline{P}_i^{\max}(\tilde{t}_{\text{on},i}^{1:N_i}, t_{\text{off},i}^{1:N_i})\right\|\right) \leq g_j\left(\left\|\overline{P}_j^{\max}(t_{\text{on},j}^{1:N_j}, \tilde{t}_{\text{off},j}^{1:N_j})\right\|\right). \tag{10}$$

Since under the updated policy all the peaks $\overline{P}_j^m$, $1 \leq m \leq N_j$, $1 \leq j \leq M$, are lower for all the targets except target i, we recall Lemma 2 and conclude that this updated policy has a lower cost than the previous one. Therefore, the previous policy cannot be optimal, which proves the proposition. ∎

This proposition gives a necessary condition for optimally allocating observation times. Moreover, its constructive proof also gives insight on how to locally optimize the dwelling times for a fixed sequence. However, in general this property is not sufficient for optimally determining how much time the agent should spend at each target. In the remainder of this paper, we will restrict ourselves to a specific set of visiting sequences $\mathcal{Y}$ where the optimality property in Prop. 4 can indeed be exploited to optimize the dwelling times at each target.

IV. Optimal Monitoring with Single Visit per Cycle

In this section, we restrict ourselves to investigating visiting sequences $\mathcal{Y}$ for which each target is visited only once in a cycle. By making this assumption, we are able to prove that the optimal one with respect to the cost (4) is the Traveling Salesman Problem (TSP) cycle, which was intuitively expected. Additionally, we present a practical algorithm for optimizing for how long each target should be observed. As a side note, in this section we will omit the upper index of $t_{\text{on},i}$ and $t_{\text{off},i}$, since $N_i = 1$.

Proposition 5: Among all the visiting sequences where each target is visited only once at each cycle, for any allocation of dwelling times T , the TSP solution is the one that minimizes the cost (4).

Proof: We recall that $t_{\text{off},i}$ is completely defined by the amount of time the agent dwells in other targets plus the time it spends traveling between targets. If every target is visited only once, the TSP cycle is the one with the least amount of traveling time. Consequently, for fixed $t_{\text{on},i}$, the values of $t_{\text{off},i}$, $\forall i$, are minimized. This fact along with Proposition 3 implies that for any set of observation times $\mathcal{T}$, the TSP tour will yield the lowest cost. ∎

Remark 2: The problem of computing the TSP tour is NP-hard. However, efficient sub-optimal solutions are available (see e.g. [13]). The approach we will discuss for optimizing the dwelling times does not rely on having the optimal TSP tour, and indeed can handle any tour as long as every target is visited exactly once. Hence we can use a suboptimal solution when the computation burden prohibits the optimal one.

Having discussed the optimal visiting sequence in this particular subset of admissible $\mathcal{Y}$, we now approach the question of how to adjust the visitation time in a computationally simple way. The main idea behind our approach is to exploit the property that the upper peaks must coincide in an optimal solution. We develop a procedure to balance the observation times at each target, such that the peak uncertainties will coincide after convergence. This iterative scheme, which updates the all the target's observation time $t_{\text{on},i}$ at each iteration, is given by:

$$t_{on,i}[k+1] = t_{on,i}[k] + k_p \log\left(\frac{g_i\left(\|\overline{P}_i\|\right)}{g_{\text{avg}}}\right), \tag{11}$$

where $g_{\text{avg}} = \left(\prod_{j=1}^{M} g_j\left(\|\overline{P}_j\|\right)\right)^{\frac{1}{M}}$ and k_p is a small positive constant. It can be interpreted as a "consensus" algorithm on the peak uncertainties $\overline{P}_i$, and thus its structure is very similar to geometric mean consensus algorithms [14]. The expression (11) does not require the computation of gradients, which makes it computationally much less demanding than gradient-based approaches.

Remark 3: At each iteration of (11), it is necessary to compute $\overline{P}_i$ and g_{avg}. For computing $\overline{P}_i$, we use the algorithm described in [15], which converges quadratically and is numerically stable. Moreover, we note that g_{avg} can be computed distributively if $\overline{P}_i$ is computed locally by each target and we assume that targets communicate among themselves using the same network connection structure that the agent uses to move (graph $\mathcal{G}$). Since this network is connected, consensus protocols can be employed at each iteration in order to have each target compute g_{avg} separately.

In order to simplify the convergence analysis of this update law, we abstract it with the following differential equation:

$$\frac{d}{dr} t_{\text{on},i} = k_p \log\left(\frac{g_i\left(\|\overline{P}_i\|\right)}{g_{\text{avg}}}\right). \tag{12}$$

Note that this version of the update law considers continuous parameter variation, i.e., the auxiliary variable r that should be understood as the continuous time equivalent of "iteration index" and does not carry any "time" interpretation. Moreover we assume that, at $r = 0$, all the targets were observed

for a strictly positive amount of time. We now prove its convergence:

Proposition 6: Under the update law (12), $\lim_{r\to\infty} \max_{i,j} |g_i(\|P_i\|) - g_j(\|P_j\|)| = 0$.

Proof:

$$\frac{d\overline{P}_i}{dr} = \frac{\partial \overline{P}_i}{\partial t_{\text{on},i}} \frac{dt_{\text{on},i}}{dr} + \frac{\partial \overline{P}_i}{\partial t_{\text{off},i}} \frac{dt_{\text{off},i}}{dr}. \tag{13}$$

However, note that the sum of observation times among different targets is constant because

$$\begin{aligned} \sum_{i=1}^{M} \frac{dt_{\text{on},i}}{dr} &= k_p \log \left(\frac{\prod_{j=1}^{M} g_j \left(\left\| \overline{P}_j \right\| \right)}{g_{\text{avg}}} \right) \\ &= k_p M \log \left(\frac{g_{\text{avg}}}{g_{\text{avg}}} \right) = 0, \end{aligned} \tag{14}$$

therefore under this control law the period T is constant. Since $T = t_{\text{on},i} + t_{\text{off},i}$, we have $\frac{dt_{\text{on},i}}{dr} = -\frac{dt_{\text{off},i}}{dr}$ and

$$\frac{d\overline{P}_i}{dr} = \left(\frac{\partial \overline{P}_i}{\partial t_{\text{on},i}} - \frac{\partial \overline{P}_i}{\partial t_{\text{off},i}} \right) k_p \log \left(\frac{g_i \left(\left\| \overline{P}_i \right\| \right)}{g_{\text{avg}}} \right). \tag{15}$$

Consequently, if $g_i(\left\|\overline{P}_i\right\|) > g_{\text{avg}}$, $\frac{d\overline{P}_i}{dr} \prec 0$. Conversely, if $g_i(\left\|\overline{P}_i\right\|) < g_{\text{avg}}$, then $\frac{d\overline{P}_i}{dr} \succ 0$. Note that

$$\max_{i,j} |g_i(\left\|\overline{P}_i\right\|) - g_j(\left\|\overline{P}_j\right\|)| = \max_i g_i(\left\|\overline{P}_i\right\|) - \min_j g_j(\left\|\overline{P}_j\right\|). \tag{16}$$

Therefore since g_{avg} is the geometric mean (i.e. its value is lower than the maximum and higher than the minimimum $g_j(\left\|\overline{P}_j\right\|)$), we get that

$$\frac{d}{dr} \max_i g_i(\left\|\overline{P}_i\right\|) < 0, \ \frac{d}{dr} \min_j g_j(\left\|\overline{P}_j\right\|) > 0. \tag{17}$$

Hence

$$\frac{d}{dr} \max_{i,j} |g_i(\left\|\overline{P}_i\right\|) - g_j(\left\|\overline{P}_j\right\|)| < 0, \tag{18}$$

which proves the proposition. ∎

Remark 4: In the proof of Prop. 6, we see that $\frac{d}{dr} \max_i g_i(\left\|\overline{P}_i\right\|) < 0$. Therefore, this update law always reduces the cost defined in (4). This also implies that if the targets do not have the same peak value $g_i\left(\left\|\overline{P}_i\right\|\right)$, then the cost can be reduced by applying (12).

Now we show that the value achieved by update law (12) is unique, i.e. does not depend on the the initial observation time distribution.

Lemma 3: For a given period T and a fixed visiting sequence $\mathcal{Y}$ where each target is visited once, there is a unique observation time distribution $\mathcal{T}$ such that $g_i(\left\|\overline{P}_i\right\|) = g_j(\left\|\overline{P}_j\right\|)$, $\forall\ i,\ j$.

Proof: Suppose there are two different costs g_{con} and g'_{con} such that all targets have the same peak value. Without loss of generality, we assume $g_{con} < g'_{con}$. Since the period T is the same, $t_{\text{on},i} > t'_{\text{on},i}$ and $t_{\text{off},i} < t'_{\text{off},i}$ $\forall i$, due to Props. 2 and 3. However, since the period is the same, we must have $\sum_i t_{\text{on},i} = \sum_i t'_{\text{on},i}$, which yields a contradiction. ∎

Finally, we give a specialization of Prop. 4 to the particular case we discuss in this section.

Proposition 7: For a fixed visiting sequence $\mathcal{Y}$ and a given cycle period T, the observation times under the update law (12) converge to the optimal observation time allocation $\mathcal{T}$ with respect to the cost function in (4).

Proof: For any set of observation times with period T, update law (12) always reduces the cost while maintaining T constant if $g_i(\left\|\overline{P}_i\right\|)$ is not the same for all the targets. Since there is a unique way such that every target has the same peak (and the update law (12) ensures convergence to it), then the observation times after convergence of (12) have to be optimal, otherwise the update law would be able to improve them. ∎

We have shown a simple way to optimize the observation times, given a fixed period T. However, we have not addressed the problem of choosing an adequate T. For choosing an optimized period, we use golden ratio search [16], that finds the global optimum in unimodal functions and local optima in a generic single variable function.

V. Simulation Results

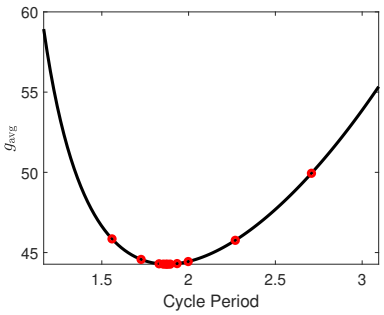
We simulated the model described in (1) and (2), with parameters indicated in Table I. Note that we assigned colors to targets in order to identify them in the figures. For simplicity, the internal states of the targets were assumed to be scalars. The target locations were drawn from a uniform distribution in $[0, 0.5] \times [0, 0.5]$ and are shown in Fig. 3b. The graph was assumed to be fully connected with edge costs being the Euclidean distance between two targets. For the definition of the optimization goal as in (4), we used $g_i(\xi) = \xi$, $\forall i$, and $\|\Gamma\| = |\Gamma|$.

TABLE I: Simulation parameters.

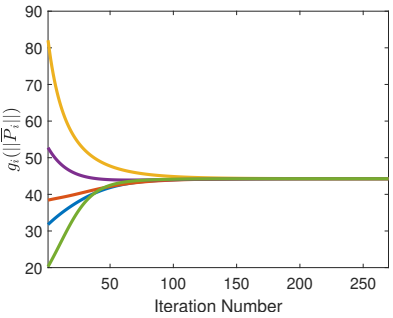
Target	1	2	3	4	5
Color	blue	red	yellow	purple	green
A_i	0.3487	0.1915	0.4612	0.2951	0.1110
Q_i	1.1924	1.2597	0.8808	1.7925	0.4363
R_i	2.3140	7.1456	4.2031	5.2866	7.5314

For the visiting order, we used the optimal TSP tour. Note that the number of targets in this example is small enough that computing the TSP tour is computationally feasible. Then, the golden ratio search was deployed, using k_p in (11) set to 10^{-2} and $[T_{\min}, T_{\max}] = [0.1t_{\text{travel}}, 3t_{\text{travel}}]$, where t_{travel} is the total travel time in one period of the TSP cycle.

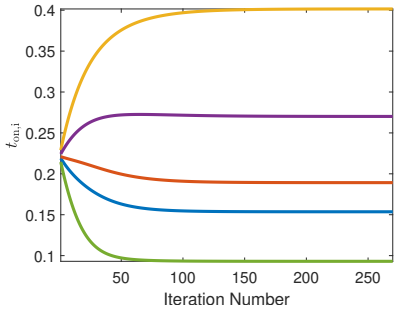
The results are shown in Figs. 2 and 3. In particular, Fig. 2 shows the evolution of the norm of the steady-state covariance matrix over one complete period of the agent trajectory, with the optimal distribution of observation times. In Fig. 2, details of the optimization process are highlighted. In Fig. 2a we can see how the uncertainty behaved as a function of the period (after balancing the observation times among targets by running (11) until convergence). Moreover, this figure also highlights the fact that the golden ratio search scheme was able to efficiently converge to a local minimum. Figs. 2b and 2c show how the observation time and the peak covariance varied while using the update law (12). Initially, all targets were visited for the same amount of time. However, as the iteration number increases, some targets are



(a) Peak value after balance among targets, as a function of cycle period.

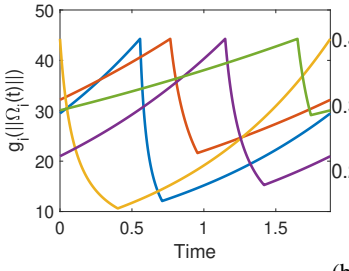


(b) Peak uncertainty at the optimal period.

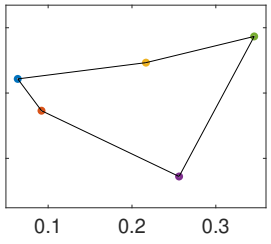


(c) Observation time at the optimal period.

Fig. 2: Results of simulating the described algorithm. In (a), the balanced peak uncertainty, as a function of the total cycle period. The red dots mark the values of T that were explored by the golden ratio search. In (b)-(c), we show the the evolution of the peak uncertainty and the observation time for each target at the optimal period.



(a) Covariance over one period.



(b) Agent trajectory (black) and targets location (colored).

Fig. 3: Results after optimizing the trajectory.

observed for more time than others. As expected, in the final iteration, all the peak covariances have converged to the same value. In Fig. 3, we can see the agent trajectory and that the targets indeed achieve the same steady state peak uncertainty.

VI. Conclusion and Future Work

In this paper, we have derived a necessary condition for the optimal allocation of observation times of targets in a network being monitored by a mobile agent. In particular, for the set of policies where each target is observed only once, we were able to claim that the optimal visiting sequence is the TSP tour and we derived a computationally cheap and derivative free procedure to optimally allocate visiting times among different targets. This allows for scalable optimization of the monitoring performance considering heterogeneous targets. In future work, we want to investigate whether we can find a similar procedure to allocate observation times for trajectories where targets are allowed to be visited multiple times. Moreover, we intend to develop greedy schemes to efficiently search over different visiting sequences. Finally, we also intend to extend this approach to multi-agent scenarios.

References

[1] X. Lan and M. Schwager, "Planning Periodic Persistent Monitoring Trajectories for Sensing Robots in Gaussian Random Fields," in *IEEE Int. Conf. on Robotics and Automation*. IEEE, 2013, pp. 2415–2420.

[2] T. Alam, G. M. Reis, L. Bobadilla, and R. N. Smith, "A Data-Driven Deployment Approach for Persistent Monitoring in Aquatic Environments," in *IEEE International Conference on Robotic Computing*, 2018, pp. 147–154.

[3] C. G. Cassandras, X. Lin, and X. Ding, "An Optimal Control Approach to the Multi-agent Persistent Monitoring Problem," *IEEE Transactions on Automatic Control*, vol. 58, no. 4, pp. 947–961, 2013.

[4] S. Welikala and C. G. Cassandras, "Asymptotic Analysis for Greedy Initialization of Threshold-Based Distributed Optimization of Persistent Monitoring on Graphs," in *Proc. of 21st IFAC WC*, 2020.

[5] F. Pasqualetti, A. Franchi, and F. Bullo, "On cooperative patrolling: Optimal trajectories, complexity analysis, and approximation algorithms," *IEEE Transac. on Robotics*, vol. 28, no. 3, pp. 592–606, 2012.

[6] S. C. Pinto, S. B. Andersson, J. M. Hendrickx, and C. G. Cassandras, "Multi-Agent Persistent Monitoring of Targets with Uncertain States," in *Arxiv*. Available online, 2020.

[7] ——, "Multi-Agent Infinite Horizon Persistent Monitoring of Targets with Uncertain States in Multi-Dimensional Environments," in *IFAC World Congress*, 2020.

[8] Z. Lin, H. H. Liu, and M. Wotton, "Kalman Filter-based Large-scale Wildfire Monitoring with a System of UAVs," *IEEE Transactions on Industrial Electronics*, vol. 66, no. 1, pp. 606–615, 2018.

[9] H. Shakhatreh, A. H. Sawalmeh, A. Al-Fuqaha, Z. Dou, E. Almaita, I. Khalil, N. S. Othman, A. Khreishah, and M. Guizani, "Unmanned Aerial Vehicles (UAVs): A Survey on Civil Applications and Key Research Challenges," *Ieee Access*, vol. 7, pp. 48 572–48 634, 2019.

[10] H. Luss, "On Equitable Resource Allocation Problems: A Lexicographic Minimax Approach," *Operations Research*, vol. 47, no. 3, pp. 361–378, 1999.

[11] S. C. Pinto, S. B. Andersson, J. M. Hendrickx, and C. G. Cassandras, "Optimal Minimax Mobile Sensor Scheduling Over a Network," in *ArXiv 2009.11386*. Available online, 2020.

[12] L. Dieci and T. Eirola, "Preserving Monotonicity in the Numerical Solution of Riccati Differential Equations," *Numerische Mathematik*, vol. 74, no. 1, pp. 35–47, 1996.

[13] J. Kirk, "Traveling Salesman Problem - Genetic Algorithm," 2020. [Online]. Available: https://www.mathworks.com/matlabcentral/fileexchange/13680-traveling-salesman-problem-genetic-algorithm

[14] F. Bullo, *Lectures on Network Systems*, 1st ed. Kindle Direct Publishing, 2020, with contributions by J. Cortes, F. Dorfler, and S. Martinez. [Online]. Available: http://motion.me.ucsb.edu/book-lns

[15] E.-W. Chu, H.-Y. Fan, W.-W. Lin, and C.-S. Wang, "Structure-preserving Algorithms for Periodic Discrete-time Algebraic Riccati Equations," *International Journal of Control*, vol. 77, no. 8, pp. 767–788, 2004.

[16] J. Kiefer, "Sequential Minimax Search for a Maximum," *Proceedings of the American Mathematical Society*, vol. 4, no. 3, pp. 502–506, 1953.