Asymptotic Analysis for Greedy Initialization of Threshold-Based Distributed Optimization of Persistent Monitoring on Graphs [⋆]

Shirantha Welikala [*] **Christos G. Cassandras** [*]

[*] *Division of Systems Engineering, Boston University, Brookline, MA 02446*
USA (e-mail: shiran27@bu.edu and cgc@bu.edu).

Abstract: We consider the optimal multi-agent persistent monitoring problem defined for a team of agents on a set of nodes (targets) interconnected according to a fixed graph topology. The objective is to minimize a measure of mean overall node state uncertainty evaluated over a finite time interval. In prior work, a class of distributed threshold-based parametric controllers has been proposed where agent dwell times at nodes and transitions from one node to the next are controlled by enforcing thresholds on the respective node uncertainties. Under such a threshold policy, on-line gradient-based techniques are then used to determine optimal threshold values. However, due to the non-convexity of the problem, this approach leads to often poor local optima highly dependent on the initial thresholds used. To overcome this initialization challenge, in this paper, the asymptotic steady-state behavior of the agent-target system is extensively analyzed for a single-agent system and dense graphs. Based on the obtained theoretical results, a computationally efficient off-line greedy technique is developed to systematically generate initial thresholds. Extensive numerical results show that the initial thresholds obtained lead to significantly better results than the locally optimal solutions known to date.

1. INTRODUCTION

A *persistent monitoring* problem arises when a dynamically changing environment needs to be monitored by a set of mobile agents. In contrast to applications like sweep coverage and patrolling (Huynh et al., 2010) where every point in the environment is equally valued for agents to monitor, in many applications (Zhou et al., 2018, 2019) the motion of the agents must be focused around a finite set of "points of interest" which can be viewed as "data sources" or "targets" to be monitored consistently. This particular problem setting can be seen in applications of surveillance systems, environmental sensing, data collecting and also in energy management (Leahy et al., 2016; Trevathan and Johnstone, 2018; Meng et al., 2019).

The persistent monitoring problem considered in this paper is also focused on a finite number of known data sources (also called "targets") located in a two-dimensional (2D) environment. In this setting, the goal of the agent team is to sense and collect information from each target to reduce an "uncertainty metric" associated with the target state. The behavior of a target uncertainty metric is such that it increases while no agent is present in the vicinity of the target and decreases when the target is being sensed by one or more agents in its vicinity. The underlying global objective is to minimize an overall measure of target uncertainties through controlling the agent trajectories.

Persistent monitoring in 1D environments has been addressed in (Zhou et al., 2018) by formulating an optimal control problem and showing that it can be reduced to a parametric optimization problem. This enables the use of Infinitesimal Perturbation Analysis (IPA) (Cassandras et al., 2010) to determine the gradients of the objective function with respect to the parameters and use gradient descent to determine their optimal values. For 2D environments, (Khazaeni and Cassandras, 2018) constrain agents to follow certain families of parametric trajectories (e.g., elliptical, Lissajous, Fourier) and use IPA to obtain an optimal solution within these families. However, as first pointed out in (Zhou et al., 2019), limiting the agent trajectories to such forms can lead to poor local optima.

To overcome the challenges mentioned above, a graph topology is adopted in (Zhou et al., 2019) where the targets and the feasible inter-target agent trajectories are abstracted as graph nodes and edges respectively. This abstraction has the added advantage of accounting for physical obstacles that might be present in the environment by constructing the graph accordingly. In this paradigm, an agent trajectory is fully characterized by a *sequence of nodes* to be visited and an associated *dwell time* to be spent at each visited node in the sequence. Therefore, the controller which optimizes a given objective should yield such a (target, dwell-time) sequence for all agents. Due to the complexity of this optimization problem, a class of distributed threshold-based parametric controllers is introduced in (Zhou et al., 2019) which characterizes agent transitions from one node to the next based on enforcing thresholds on respective node uncertainties. This parameterization enables the use of IPA to find optimal thresholds in an on-line manner using gradient descent. However, due to the non-convexity of the associated objective function, this often results in a poor local optimum highly dependent on the initial thresholds selected, which in (Zhou et al., 2019) are generated randomly.

This paper improves the solution approach proposed in (Zhou et al., 2019) by determining a set of high-performing thresholds to initialize the IPA-based gradient descent process so that an improved set of (still locally) optimal thresholds is subsequently obtained. This is accomplished by analyzing the

[⋆] Supported in part by NSF under grants ECCS-1931600, DMS-1664644, CNS-1645681, by AFOSR under grant FA9550-19-1-0158, by ARPA-E's NEXTCAR program under grant DE-AR0000796 and by the MathWorks.

asymptotic behavior of persistent monitoring systems on graphs where agents are constrained to follow periodic and non-overlapping sequences of nodes, also called "target-cycles". Based on the obtained theoretical results, a computationally efficient off-line greedy approach is proposed to construct a high-performing set of agent trajectories on the given graph. Next, we translate the constructed trajectories into a set of thresholds which we use to initialize the IPA-based gradient descent. Extensive simulation results show that this initialization technique improves performance by large margins.

2. PROBLEM FORMULATION

Consider a two-dimensional mission space with M targets (nodes) in the set $\mathscr{T} = \{1,2,\dots,M\}$ and N agents in the set $\mathscr{A} = \{1,2,\dots,N\}$. Each target $i \in \mathscr{T}$ is located at a fixed position $X_i \in \mathbb{R}^2$. Each agent $a \in \mathscr{A}$ is allowed to move in the mission space, therefore, its trajectory is denoted by $\{s_a(t) \in \mathbb{R}^2, t \geq 0\}$. Target locations and initial agent locations are pre-specified. Each target $i \in \mathscr{T}$ has an associated *uncertainty state* $R_i(t) \in \mathbb{R}$ with the following properties: (i) $R_i(t)$ increases at a rate A_i when no agent is visiting it, (ii) $R_i(t)$ decreases at a rate $B_i N_i(t) - A_i$ where B_i is the uncertainty removal (data collection) rate by an agent and $N_i(t) = \sum_{a=1}^{N} \mathbf{1}_{\{s_a(t)=X_i\}}$ is the number of agents present at target i at time t, and (iii) $R_i(t) \geq 0,\ \forall t \geq 0$. The target uncertainty state dynamics for any $i \in \mathscr{T}$ are

$$\dot{R}_i(t) = \begin{cases} 0 & \text{if } R_i(t) = 0 \text{ and } A_i \leq B_i N_i(t), \\ A_i - B_i N_i(t) & \text{otherwise,} \end{cases} \tag{1}$$

with A_i,B_i and $R_i(0)$ values pre-specified. As pointed out in (Zhou et al., 2019), this target uncertainty model has an attractive queueing system interpretation where A_i, and $B_i N_i(t)$ can be thought of as an arrival rate and a controllable service rate respectively for each target (i.e., a node) in a queueing network.

In some persistent monitoring models (Zhou et al., 2018), each agent $a \in \mathscr{A}$ is assumed to have a finite sensing range $r_a > 0$ which allows it to decrease $R_i(t)$ whenever $\|s_a(t) - X_i\| \leq r_a$. However, we follow the approach used in (Zhou et al., 2019) where $r_a = 0$ is assumed and $N_i(t)$ is used to replace the joint detection probability of a target $i \in \mathscr{T}$. As we will see, our analysis does not depend on the dynamic model of the agents.

The objective of this persistent monitoring system is to minimize a measure of *mean system uncertainty* J_T (evaluated over a finite time horizon T), where

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

by controlling agent motion.

Graph Topology: We embed a directed graph topology $\mathscr{G} = (\mathscr{V}, \mathscr{E})$ into the 2D mission space where the graph vertices represent the targets ($\mathscr{V} = \{1,2,\dots,M\} = \mathscr{T}$), and the graph edges ($\mathscr{E} \subseteq \{(i,j) : i,j \in \mathscr{V}\}$) represent inter-target trajectory segments (which may be curvilinear paths with arbitrary shapes so as to account for potential obstacles in the mission space) available for agents to travel. It is assumed that each edge $(i,j) \in \mathscr{E}$ has an associated fixed time value $\rho_{ij} \in \mathbb{R}_{\geq 0}$ which represents the amount of time that an agent has to spend to travel from target i to j. In general, ρ_{ij} may depend on the agent dynamic model, target locations and any obstacles that may be present in the mission space. Also, we assume that if $(i,j) \in \mathscr{E}$, then, $(j,i) \in \mathscr{E}$ with $\rho_{ij} = \rho_{ji}$ (for simplicity only and no loss of generality). In this paradigm, the *neighbor set* $\mathscr{N}_i$ of target i is defined as $\mathscr{N}_i = \{j : (i,j) \in \mathscr{E}\}$.

Under the assumed target dynamics in (1) and the agent sensing capabilities, it is intuitive that to minimize the objective J_T in (2) each agent may dwell (i.e., remain stationary) only at each target that it visits in its trajectory. Further, according to the embedded target topology $\mathscr{G}$ which constrains the agent motion, when an agent $a \in \mathscr{A}$ leaves a target $i \in \mathscr{V}$ after finishing its dwelling period, its next target would be some $j \in \mathscr{N}_i$. It is clear that in order to minimize the objective J_T, the predefined ρ_{ij} value should correspond to the minimum time an agent may take to travel between targets i and j. This *dwell-travel* approach is intended to minimize the agent time spent outside of targets When an agent $a \in \mathscr{A}$ arrives at a target $i \in \mathscr{V}$ it has to determine a *dwell time* $\tau_i^a \in \mathbb{R}_{\geq 0}$ and a *next visit target* $v_i^a \in \mathscr{N}_i$. After these two decisions are made, the same process repeats at the chosen next target. Therefore, the optimal approach to control the set of agents which minimizes the objective J_T in (2) can be determined in the form of a set of optimal dwelling time and next visit target sequences. This is a challenging task even for the simplest problem configurations due to the nature of the search space.

Threshold based control policy: Similar to the framework proposed in (Zhou et al., 2019), we introduce a Threshold-based Control Policy (TCP). Under this TCP, each agent a bases its decision sequence by adhering to a set of pre-specified parameters denoted by $\Theta^a \in \mathbb{R}^{M \times M}$ which serve as thresholds on the target uncertainties. Note that the $(i,j)^{\text{th}}$ parameter in the Θ^a matrix is denoted by $\theta_{ij}^a \in \mathbb{R}_{\geq 0}\ \forall i,j \in \mathscr{V}$.

Let us denote the set of neighbors of a target i which violate their thresholds (called *active neighbors*) when agent a is residing in i at time t by $\mathscr{N}_i{}^a(t) \subseteq \mathscr{N}_i$ where

$$\mathscr{N}_i{}^a(t) \triangleq \{j : R_j(t) > \theta_{ij}^a,\ j \in \mathscr{N}_i\}. \tag{3}$$

When agent a arrives at target i at time $t = t'$, the dwell time τ_i^a spent at target i is determined by: (i) the diagonal element θ_{ii}^a based on the threshold satisfaction condition $R_i(t) < \theta_{ii}^a$, and, (ii) the active neighbor existence condition $|\mathscr{N}_i{}^a(t)| > 0$ at $t = t' + \tau_i^a$ (where $|\cdot|$ is the cardinality operator). Subsequently, agent a's next visit target v_i^a is chosen from the set of active targets $\mathscr{N}_i{}^a(t) \subseteq \mathscr{N}_i$ using the off-diagonal thresholds $\{\theta_{iv}^a : v \in \mathscr{N}_i{}^a(t)\}$ at $t = t' + \tau_i^a$. Formally,

$$\begin{aligned} \tau_i^a &:= \operatorname*{arginf}_{\tau \geq 0} \left\{ [R_i(t'+\tau) < \theta_{ii}^a] \ \&\ [|\mathscr{N}_i{}^a(t'+\tau)| > 0] \right\}, \\ v_i^a &:= \operatorname*{argmax}_{v \in \mathscr{N}_i{}^a(t'+\tau_i^a)} \left\{ R_v(t'+\tau_i^a) - \theta_{iv}^a \right\}. \end{aligned} \tag{4}$$

These update equations define the (dwell time, next target) sequences of agents under the TCP. The first condition in the τ_i^a expression in (4) ensures that agent a will dwell at target i until at least its own uncertainty $R_i(t)$ drops below θ_{ii}^a; the second condition ensures that there exists at least one neighbor $v \in \mathscr{N}_i$ whose uncertainty $R_v(t)$ has increased beyond the threshold θ_{iv}^a. According to the v_i^a expression in (4), v_i^a is the neighboring target of i chosen from the set $\mathscr{N}_i{}^a(t' + \tau_i^a) \subseteq \mathscr{N}_i$ with the largest threshold violation.

It is important to point out that under this TCP based on (3) and (4), we limit agents from using non-neighboring target state information. This enables each agent to operate in a *distributed* manner using only the state information obtained from the target where it currently resides and from its neighboring targets. An example target topology and an agent threshold matrix are shown in Fig. 1. By convention, when edges are missing in the graph, the respective off-diagonal entries become irrelevant.

Discrete event system view: Under the described TCP, the behavior of the persistent monitoring system is fully defined by the set of agent decision sequences

$$\mathscr{U}(\Theta) = \{(\tau^a_{i(l)}(\Theta^a), v^a_{i(l)}(\Theta^a)) : a \in \mathscr{A},\ i(l) \in \mathscr{V},\ l = 1, 2, \dots\}.$$

Here, $\Theta \in \mathbb{R}^{M\times M\times N}$ is the collection of all agent threshold matrices and $i(l) \in \mathscr{V}$ is the l^{th} target visited by agent a. Based on (4), the persistent monitoring system is a discrete event system (DES) (Cassandras and Lafortune, 2010) where the underlying *event set* consists of: (i) All possible agent arrivals/departures at/from targets, (ii) The instances where a target uncertainty reaches 0 from above, and (iii) The 'start' and 'end' events triggered respectively at $t = 0$ and $t = T$. The *output* of this DES can be considered as a vector $\bar{R}^k = [R_i(t^k)]_{i\in\mathscr{V}} \in \mathbb{R}^M$ evaluated at all event times $\{t^k : k \in \{0, 1, \dots, K\}\}$ with $t^0 = 0$ and $t^K = T$. Note that under a given TCP Θ, the output trajectory of this DES is fully defined by the event times implied by (4).

Considering the dependence of both the state trajectory and the output trajectory on the chosen set of parameters Θ, the performance metric J_T in (2) depends on the parameters Θ. Therefore, within the TCP class of agent controllers, we aim at determining an Optimal TCP (OTCP) Θ^* such that

$$\Theta^* = \underset{\Theta\geq\mathbf{0}}{\operatorname{argmin}}\ J_T(\Theta) = \frac{1}{T}\sum_{i=1}^{M}\sum_{k=0}^{K}\int_{t^k}^{t^{k+1}} R_i(t)dt. \tag{5}$$

Differentiating the cost $J_T(\Theta)$ w.r.t. parameters Θ gives $\nabla J_T(\Theta)$ $= \frac{1}{T}\sum_{i=1}^{M}\sum_{k=0}^{K}\int_{t^k}^{t^{k+1}}\nabla R_i(t)dt$. where $\nabla \equiv \frac{\partial}{\partial\Theta}$. It should be noted that even though event times are dependent on the TCP Θ, when taking the derivative (Flanders, 1973), the effect of it gets canceled out since we have fixed $t^0 = 0$ and $t^K = T$ (Cassandras et al., 2010). Further, using the linear behavior of the target uncertainty dynamics in (1), and the way that we have designed our event space, following Lemma 1 in (Zhou et al., 2019) we can easily show that $\nabla R_i(t) = \nabla R_i(t^k)\ \forall t \in (t^k, t^{k+1}]$. Therefore, the gradient $\nabla J_T(\Theta)$ becomes a simple summation: $\nabla J_T(\Theta) = \frac{1}{T}\sum_{i=1}^{M}\sum_{k=0}^{K}\nabla R_i(t^k)(t^{k+1} - t^k)$. In (Zhou et al., 2019), where the class of TCP controllers was introduced, the use of Infinitesimal Perturbation Analysis (IPA) (Cassandras et al., 2010) is extensively discussed so as to evaluate $\nabla R_i(t^k)$ (hence, $\nabla J_T(\Theta)$) on-line and in a distributed manner. This enables the use of a gradient descent algorithm:

$$\Theta^{(l+1)} = \left[\Theta^{(l)} - \beta^{(l)}\nabla J_T(\Theta^{(l)})\right]^+, \tag{6}$$

to update the TCP Θ iteratively. In (6), the projection operator $[\cdot]^+ = \max\{0, \cdot\}$ is used. The step size $\beta^{(l)}$ is selected such that it diminishes according to the standard conditions $\sum_{l=1}^{\infty}\beta^{(l)} = \infty$ and $\lim_{l\to\infty}\beta^{(l)} = 0$ (Bertsekas, 2016).

Initialization $\Theta^{(0)}$: In (Zhou et al., 2019) a randomly generated set of values is used to initialize thresholds $\Theta^{(0)}$ for (6). Due to the non-convexity of the objective function in (5), we expect that the resulting value of Θ when (6) converges depends on $\Theta^{(0)}$. Therefore, identifying well-performing initial thresholds will generally provide significant improvements over the local minimum resulting from randomly selected ones. Motivated by this idea, we first investigate the structural and behavioral properties of the underlying system (under a few additional constraints). That knowledge is then used to construct a candidate for $\Theta^{(0)}$.

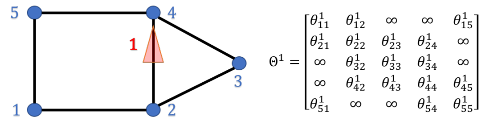

5
4
1
3
1
2
$\Theta^1 = \begin{bmatrix} \theta^1_{11} & \theta^1_{12} & \infty & \infty & \theta^1_{15} \\ \theta^1_{21} & \theta^1_{22} & \theta^1_{23} & \theta^1_{24} & \infty \\ \infty & \theta^1_{32} & \theta^1_{33} & \theta^1_{34} & \infty \\ \infty & \theta^1_{42} & \theta^1_{43} & \theta^1_{44} & \theta^1_{45} \\ \theta^1_{51} & \infty & \infty & \theta^1_{54} & \theta^1_{55} \end{bmatrix}$

Fig. 1. An example target topology with five targets and one agent with its threshold parameters.

3. ASYMPTOTIC ANALYSIS

It is proved in (Zhou et al., 2019) that in a single-agent persistent monitoring system it is optimal to make the target uncertainty $R_i(t) = 0$ on each visit of agent a at target i. In other words, in the OTCP, $\theta^a_{ii} = 0$. Moreover, experimental results in (Zhou et al., 2019) provide some intuition about better performing agent behaviors: (i) after a brief initial transient phase, each agent converges to a (steady-state) periodic behavior where it cycles across a fixed subset of targets, and, (ii) in this steady state, agents do not tend to share targets with other agents.

We now aim to use the aforementioned observations to efficiently (and in an off-line manner) construct better performing (favorable) agent trajectories which can then be used to derive a favorable candidate for $\Theta^{(0)}$. This can provide a systematic initialization to the gradient descent scheme (6) so that it eventually converges to a Θ^* with much better performance compared to the random initialization approach used in (Zhou et al., 2019). Such favorable agent trajectories take the form of a *target-cycle* on the given graph. Therefore, we need to construct a set of target-cycles (one per agent) in the given graph topology. In this paper, we limit ourselves to single-agent persistent monitoring problems on sufficiently dense target topologies (generalizations are given in (Welikala and Cassandras, 2019)). More precisely, we consider a given target topology $\mathscr{G} = (\mathscr{V}, \mathscr{E})$ to be a 'sufficiently dense' graph, if $\mathscr{G}$ is *bi-triangular*.

Definition 1. A directed graph $\mathscr{G} = (\mathscr{V}, \mathscr{E})$ with $|\mathscr{V}| > 3$ is bi-triangular if for all $(i, j) \in \mathscr{E}$ there exist $k, l \in \mathscr{V}$ such that $(i, k), (k, j) \in \mathscr{E}$, $(i, l), (l, j) \in \mathscr{E}$, and $k \neq l$.

Assumption 1. (i). Only one agent is available, and (ii). The given target topology $\mathscr{G} = (\mathscr{V}, \mathscr{E})$ is bi-triangular.

In (Welikala and Cassandras, 2019), this assumption is relaxed with more general theoretical results developed based on the analysis in this paper. This enables eventually solving persistent monitoring problems with *multiple agents* on *sparse graphs*. Under Assumption 1, we only search for a single target-cycle (agent trajectory) in the given graph $\mathscr{G}$. Moreover, exploiting the assumed dense nature of the given graph, we propose an iterative greedy scheme to construct a high-performing target-cycle. This constructed agent trajectory is then transformed to a TCP $\Theta^{(0)}$ for the subsequent use in gradient descent (6) to obtain an OTCP Θ^*.

Analysis of an unconstrained target-cycle: A *target-cycle* is a finite sequence of targets selected from the given graph $\mathscr{G}$ such that the corresponding sequence of edges also exist in $\mathscr{E}$. The latter condition enables an agent to traverse such a target-cycle. An *unconstrained target-cycle* is a target-cycle with no target on it being repeated. We define the set $\mathscr{C}$ to include all possible unconstrained target-cycles on the graph $\mathscr{G}$. A generic element of $\mathscr{C}$ (i.e., a generic target-cycle) is denoted by $\Xi_i = \{i_1, i_2, \dots, i_m\} \subseteq \mathscr{V}$, where $i_j \in \mathscr{V},\ j \in \{1, 2, \dots, m\}$ and $m = |\Xi_i| \leq M$. The corresponding sequence of edges is denoted by $\xi_i = \{(i_m, i_1), (i_1, i_2), \dots, (i_{m-1}, i_m)\} \subseteq \mathscr{E}$. Note that ξ_i is fully defined by Ξ_i, and, vice versa.

Since we are interested in greedily constructing a target-cycle with a high-performing mean system uncertainty value (i.e., J_T

in (2)), we need to have an assessment criterion for any given arbitrary target-cycle. Thus, we define the *steady-state mean cycle uncertainty* metric $J_{ss}(\Xi_i)$:

$$J_{ss}(\Xi_i) = \lim_{T\to\infty} \frac{1}{T} \int_0^T \sum_{j\in\Xi_i} R_j(t)dt. \tag{7}$$

We now present an off-line technique to evaluate $J_{ss}(\Xi_i)$. For notational convenience, Ξ_i and its targets are relabeled as $\Xi = \{1,2,\dots,n,n+1,\dots,m\}$ by dropping the subscript i (see Fig. 2). We make the following assumption regarding an agent's behavior on a target-cycle.

Assumption 2. After visiting a target $n \in \Xi$, the agent will leave it if and only if the target uncertainty reaches zero.

Here, the 'only if' component follows from the aforementioned result in (Zhou et al., 2019) that it is optimal to make the target uncertainty $R_i(t) = 0$ whenever the agent visits target i. The 'if' component restricts agent decisions by assuming the existence of an active neighbor to i as soon as $R_i(t) = 0$ occurs in (4). Nonetheless, recalling that our focus is only on initializing (6), this potential sub-optimality will be compensated by the eventual use of (6).

A *tour* on the target-cycle Ξ (shown in Fig. 2) starts/ends when the agent leaves the last target m to reach target 1. When agent a is in its k^{th} tour on Ξ, the dwell time spent at a target n is denoted by $\tau^a_{n,k}$ and the travel time spent on an edge $(n-1,n) \in \mathscr{E}$ is (by definition) $\rho_{(n-1)n}$. However, we respectively use $\tau_{n,k}$ and ρ_n to represent these two quantities as it does not introduce any ambiguity (note that $\rho_1 = \rho_{m1}$). Observe that travel times ρ_n are independent of k due to the embedded graph topology.

Under this notation, the trajectory of the target uncertainty $R_n(t)$ where $n \in \Xi$ is shown in Fig. 3. Using the triangle labeled XYZ (in Fig. 3), the dynamics of target n's dwell time $\tau_{n,k}$ (w.r.t. k) can be obtained as

$$(B_n - A_n)\tau_{n,k+1} = A_n\left(\sum_{i=n+1}^{m}\left[\rho_i + \tau_{i,k}\right] + \sum_{i=1}^{n-1}\left[\rho_i + \tau_{i,k+1}\right] + \rho_n\right).$$

Defining $\alpha_n \triangleq \frac{B_n - A_n}{A_n}$, and setting $\rho_\Xi = \sum_{i=1}^{m}\rho_i$ to be the total cycle travel time, the above relationship can be simplified as

$$-\sum_{i=1}^{n-1}\tau_{i,k+1} + \alpha_n\tau_{n,k+1} = \rho_\Xi + \sum_{i=n+1}^{m}\tau_{i,k}. \tag{8}$$

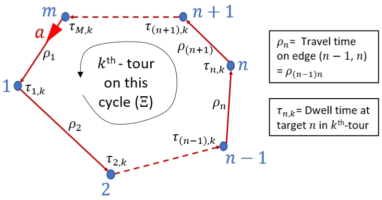

m
$n+1$
a
$\tau_{M,k}$
$\tau_{(n+1),k}$
ρ_1
$\rho_{(n+1)}$
k^{th} - tour on this cycle (Ξ)
$\tau_{n,k}$
n
1
$\tau_{1,k}$
ρ_n
ρ_2
$\tau_{(n-1),k}$
$\tau_{2,k}$
$n-1$
2
ρ_n= Travel time on edge $(n-1, n)$ = $\rho_{(n-1)n}$
$\tau_{n,k}$= Dwell time at target n in k^{th}-tour

Fig. 2. A generic single agent unconstrained target-cycle Ξ.

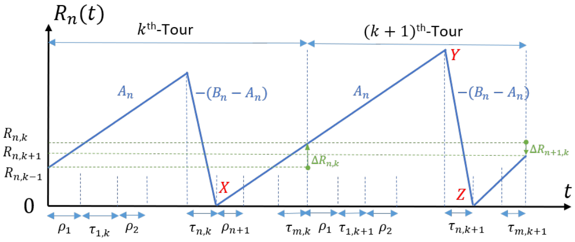


Fig. 3. Trajectories of target uncertainties during agent tours.

Now, (8) can be written for all $n \in \{1,2,\dots,m\}$ in a compact form using the vectors $\bar{\tau}_k = [\tau_{1,k},\tau_{2,k},\dots,\tau_{m,k}]^T$, $\bar{\alpha} = [\alpha_1,\alpha_2,\dots,\alpha_m]^T$ and $\bar{1}_m = [1,1,\dots,1]^T \in \mathbb{R}^m$, as follows:

$$\Delta_1\bar{\tau}_{k+1} = \Delta_2\bar{\tau}_k + \bar{1}_m\rho_\Xi, \tag{9}$$

where $\Delta_2 \in \mathbb{R}^{m\times m}$ is the strictly upper triangular matrix with all non-zero elements being 1, and $\Delta_1 = diag(\bar{\alpha}) - \Delta_2^T$. The expression in (9) has the form of an affine linear system and describes the evolution of the agent's dwell times at all the targets on the target-cycle Ξ over the number k of tours completed. To get an explicit expression for $J_{ss}(\Xi)$ in (7), we first need to establish the following two lemmas (due to space limitations, all proofs are provided in Welikala and Cassandras (2019)).

Lemma 1. When $\sum_{i=1}^{m}\frac{A_i}{B_i} < 1$, the system of equations given in (9) has a feasible equilibrium point $\bar{\tau}_{eq}$ (reached at $k = k_{eq}$),

$$\bar{\tau}_{eq} = \left(\frac{\bar{\beta}}{1-\bar{1}_m^T\bar{\beta}}\right)\rho_\Xi, \quad \text{i.e., } \tau_{n,k_{eq}} = \left(\frac{\beta_n}{1-\sum_{i=1}^{m}\beta_i}\right)\rho_\Xi, \tag{10}$$

for all $n \in \Xi$ with $\beta_n \triangleq \frac{A_n}{B_n}$ and $\bar{\beta} = [\beta_1,\beta_2,\dots,\beta_m]^T$.

In order to establish the stability properties of $\bar{\tau}_{eq}$ given by Lemma 1, we need make the following assumption.

Assumption 3. The matrix $\Delta_1^{-1}\Delta_2$ is Schur stable.

We point out that all the eigenvalues of Δ_2 are located at the origin as it is a strictly upper triangular matrix. Further, since Δ_1 is a lower triangular matrix with diagonal elements being $\{\alpha_i : i \in \Xi\}$, the eigenvalues of Δ_1^{-1} are located at $\{\frac{1}{\alpha_i} : i \in \Xi\}$. Using the definition of $\alpha_i (= \frac{B_i - A_i}{A_i})$, it is easy to show that $|\frac{1}{\alpha_i}| < 1 \iff 0 \leq \frac{A_i}{B_i} < \frac{1}{2}$, which is less restrictive than the condition $\sum_{i=1}^{m}\frac{A_i}{B_i} < 1$ required in Lemma 1. Therefore, it seems reasonable to conjecture that the eigenvalues of $\Delta_1^{-1}\Delta_2$ are located within the unit circle; however, to date, we have not provided a formal proof to the statement in Assumption 3.

Lemma 2. Under Assumption 3, the equilibrium point $\bar{\tau}_{eq}$ given in Lemma 1 for the affine linear system (9) is globally asymptotically stable.

We now present our main theorem regarding the mean steady state cycle uncertainty in (7) achieved for the persistent monitoring system shown in Fig. 2.

Theorem 1. Under Assumptions 2 and 3 with $\sum_{i=1}^{m}\frac{A_i}{B_i} < 1$, the single agent unconstrained target-cycle achieves a steady state mean cycle uncertainty value

$$J_{ss}(\Xi) = \frac{1}{2}(\bar{B} - \bar{A})^T\bar{\tau}_{eq}, \tag{11}$$

where $\bar{B} = [B_1,B_2,\dots,B_m]^T$, $\bar{A} = [A_1,A_2,\dots,A_m]^T$, and $\bar{\tau}_{eq}$ is given in (10).

Theorem 1 provides a computationally effective means of assessing off-line simple persistent monitoring configurations (like the one in Fig. 2). Hence its use in constructing high-performing target-cycles on a given $\mathscr{G}$ is discussed next.

4. GREEDY TARGET-CYCLE CONSTRUCTION

If $|\mathscr{C}|$ is small, Theorem 1 can be used to directly identify the best performing (steady state) target-cycle (via brute-force)

$$\Xi^* = \arg\min_{\Xi\in\mathscr{C}} J_{ss}(\Xi). \tag{12}$$

However, such a brute-force approach is computationally expensive and eventually prohibitive as $|\mathscr{C}|$ grows exponentially with respect to the number of targets/edges in the graph $\mathscr{G}$ (i.e., w.r.t. $|\mathscr{V}|$ or $|E|$). As a computationally efficient alternative, we propose to construct a sub-optimal target-cycle (say $\Xi^{\#}$) as a candidate solution to (12) by following a specifically designed greedy scheme. In this scheme, each iteration search expands the current target-cycle (say Ξ) by adding an unvisited target $i \in \mathscr{V} \backslash \Xi$ to it ($\cdot\backslash\cdot$ represents the set subtraction operator). Subsequently, we transform the constructed $\Xi^{\#}$ to a set of corresponding TCP parameters which we then use as a candidate for $\Theta^{(0)}$ in the gradient descent scheme (6). Therefore, getting the optimal target-cycle Ξ^* is not a necessity at this stage when compared to the importance of keeping efficient the overall off-line process of getting a candidate $\Theta^{(0)}$.

Let us define the finite horizon version of $J_{ss}(\Xi_i)$ defined in (7) as the finite horizon mean cycle uncertainty $J_T(\Xi_i)$:

$$J_T(\Xi_i) = \frac{1}{T}\int_0^T \sum_{j\in\Xi_i} R_j(t)dt. \qquad (13)$$

Note that this $J_T(\Xi_i)$ metric is equivalent to the mean system uncertainty metric J_T defined in (2), when $\Xi_i = \mathscr{V}$.

Contribution of a neglected target: A *neglected target* is defined as a target which is not visited by any agent during the entire time period $[0,T]$. Given the finite horizon nature of the objective function J_T in (2), if one or more targets are located remotely relative to the rest, then neglecting such remote targets might be preferable than to visiting them. The following lemma quantifies the contribution of such neglected targets to the objective J_T.

Lemma 3. The contribution of a neglected target $i \in \mathscr{V}$ to the mean system uncertainty J_T (defined in (2)) is $\left(R_{i,0} + \frac{A_i T}{2}\right)$.

Assumption 4. For any target-cycle $\Xi \in \mathscr{C}$, the difference between the steady state mean cycle uncertainty $J_{ss}(\Xi)$ (defined in (7)) and the finite horizon mean cycle uncertainty $J_T(\Xi)$ (defined in (13)) is bounded by some finite constant $K_e \in \mathbb{R}_{\geq 0}$, i.e., $|J_{ss}(\Xi) - J_T(\Xi)| \leq K_e$.

In the greedy scheme, we use the $J_{ss}(\cdot)$ metric defined in (7) to compare the performance of different target-cycles because it can be efficiently evaluated using Theorem 1. However, since we consider a finite horizon objective as in (2), the $J_T(\cdot)$ metric defined in (13) is more appropriate in evaluating target-cycle performances. The above assumption states that $J_T(\cdot)$ will always lie within $J_{ss}(\cdot) \pm K_e$ and we observe that: (i) K_e is small when the steady state tour duration T_Ξ and the finite horizon T are such that: $T \gg T_\Xi$, and (ii) K_e is small if the dynamics of the steady state error of (9) are faster (i.e., when $0 \simeq \frac{A_i}{B_i} \ll 1$).

Target-cycle expansion operation: Recall that we used the notation $\Xi_i = \{i_1, i_2, \ldots, i_m\}$ to represent a target-cycle and $\xi_i = \{(i_m, i_1), (i_1, i_2), \ldots, (i_{m-1}, i_m)\}$ to represent the corresponding sequence of edges in Ξ_i. Omitting the subscript i (for notational convenience), consider a target-cycle $\Xi = \{1, 2, \ldots, m\}$ with its respective edges sequence $\xi = \{(m,1), (1,2), \ldots, (m-1,m)\}$. In order to expand Ξ so that it includes one more target i selected from the set of neglected targets $\mathscr{V}\backslash\Xi$, we have to: (i) remove one edge (say $(j,l) \in \xi$) chosen from ξ and replace it with two new consecutive edges, and (ii) place target i inside Ξ according to this edge replacement. These two operations define a new (expanded) target-cycle Ξ' (and ξ') as shown in Fig. 4. The following theorem quantifies the estimated gain in the objective function due to such a target-cycle expansion in terms of $J_{ss}(\cdot)$ given in Theorem 1. This gain due to an added target i is formally referred to as *marginal gain* and is denoted by $\Delta J_T(i|\xi, e)$ where e is the edge removed from ξ to add the target $i \in \mathscr{V}\backslash\Xi$ to Ξ.

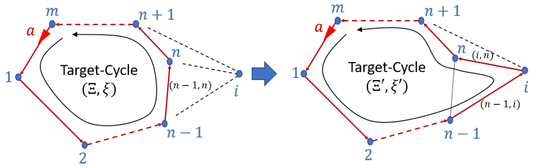


Fig. 4. A basic target-cycle expanding operation.

Theorem 2. Under Assumptions 1, 2 and 4, the estimated gain (reduction) in the objective function J_T due to the basic target-cycle expansion operation (shown in Fig. 4) is

$$\Delta J_T(i|\xi, (n-1, n)) = \left(R_{i,0} + \frac{A_i T}{2}\right) + J_{ss}(\Xi) - J_{ss}(\Xi'), \quad (14)$$

where Ξ' is the expanded target-cycle. The associated estimation error of this term is $\pm 2K_e$.

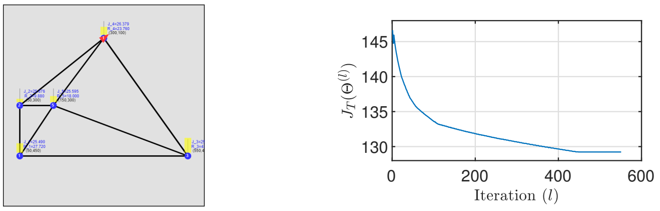
Greedy algorithm: The greedy target-cycle construction scheme starts by conducting a brute-force search to determine the best target-cycle of length 2 (over $\mathscr{G}$). The obtained solution is then iteratively expanded by adding external targets. In each of such iterations, the optimum target to be added (and the edge to be removed) is determined via conducting a brute-force search - where the objective function takes the form (14). More details are provided in (Welikala and Cassandras, 2019).

5. SIMULATION RESULTS

The proposed persistent monitoring solution (including the improvements detailed in (Welikala and Cassandras, 2019)) and the solution given in (Zhou et al., 2019) have been implemented in a JavaScript web simulator which is available at http://www.bu.edu/codes/simulations/shiran27/PersistentMonitoring/.

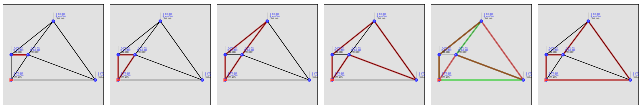
A large collection of simulation results obtained under different problem settings can be found in (Welikala and Cassandras, 2019). Here, due to space limitations, we limit ourselves to the configuration shown in Fig. 5(a) where blue circles represent targets while black lines represent available edges that agents can take to travel between targets. Red triangles and yellow vertical bars respectively indicate agent locations and target uncertainty levels. However, since both of these quantities ($s_a(t)$ and $R_i(t)$) are time-varying, we show their state only at the terminal time $t = T$ when the best threshold matrix Θ^* (obtained from (6)) is used in the TCP. For comparison purposes, Θ^* was determined using two methods: (i) The randomized initialization method proposed in (Zhou et al., 2019) and (ii) The greedy initialization method proposed in this paper.

The target parameters were chosen as $A_i = 1$, $B_i = 10$ and $R_i(0) = 0.5$, $\forall i \in \mathscr{V}$. All targets have been placed inside a 600×600 mission space and their location coordinates are specified in each problem configuration figure. The time horizon was taken as $T = 500$. The first-order agent model (Zhou et al., 2019) was used (limiting each agent's maximum speed to 50 units per second). The initial locations of agents were chosen such that they are uniformly distributed over the targets at $t = 0$. In cases where the initial TCP $\Theta^{(0)}$ is randomly generated (i.e., to replicate (Zhou et al., 2019)), each finite element in $\Theta^{(0)}$ is chosen from a uniform distribution over $[0, 10]$. Finally, in (6), diminishing step sizes $\beta^{(l)} = \frac{0.25}{\sqrt{l}}$ were used.



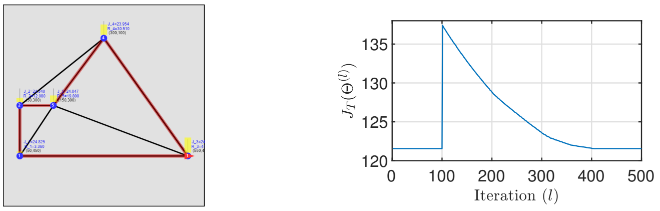
(a) Config. at $t = T$. (b) Cost vs iterations plot.

Fig. 5. SASE1: Starting with a random $\Theta^{(0)}$, converged to a OTCP with cost $J_T = 129.2$.



(a) Step 1 (b) Step 2 (c) Step 3 (d) $\Xi^{\#}$ (e) 2-Opt (f) Ξ^{R}

Fig. 6. Greedy target-cycle construction iterations and a refinement step for the target topology of SASE1.



(a) Config. at $t = T$. (b) Cost vs iterations plot.

Fig. 7. SASE1: Starting with $\Theta^{(0)}$ given by the proposed greedy initialization scheme, at $l = 100, \Theta^{(l)}$ is randomly perturbed and converges back to the same (initial) objective function value $J_T = 121.6$ (improvement $= +7.6$).

With random initialization of $\Theta^{(0)}$ *(Zhou et al., 2019):* Figure 5(b) shows the evolution of $J_T(\Theta^{(l)})$ when $\Theta^{(l)}$ is updated according to (6) using gradients $\nabla J_T(\Theta^{(l)})$ given by the IPA method as described in (Zhou et al., 2019).

With proposed greedy initialization of $\Theta^{(0)}$***:*** Figures 6 (a)-(d) show the intermediate target-cycles generated by the greedy target-cycle construction process when applied to the single-agent simulation example (referred to as SASE1) in Fig. 5(a). The target-cycle shown as a red contour in Fig. 6(d) (labeled $\Xi^{\#}$) corresponds to $J_{ss}(\Xi^{\#}) = 135.7$. A subsequently identified refinement step and its improved result (labeled Ξ^{R}) are shown in Figs. 6(e) and (f) respectively, where $J_{ss}(\Xi^{R}) = 128.7$.

Next, the identified target-cycle Ξ^{R} is converted to the respective TCP $\Theta^{(0)}$. Figure 7(b) shows that this target-cycle results in $J_T(\Theta^{(0)}) = 121.6$ which cannot be further improved using (6). To ensure that this is (at least locally) optimal, after 100 iterations (at $l = 100$) (6), the running $\Theta^{(l)}$ was randomly perturbed. It can be seen in Fig. 7(b) that $\Theta^{(l)}$ converged back to the same objective function value $J_T(\Theta^{(l)}) = 121.6$. It is important to note that this solution is better than the best TCP Θ^{*} obtained with random initialization of $\Theta^{(0)}$ (shown in Fig 5) by 5.88%. For a compilation of such results, we direct the reader to Tab. 1, Fig. 39 and Fig. 40 in (Welikala and Cassandras, 2019).

All simulation examples (including those in (Welikala and Cassandras, 2019)) were evaluated on an Intel® Core™ i7-7800 CPU 3.20 GHz Processor with a 32 GB RAM. The average execution time observed for the greedy initialization technique to generate $\Theta^{(0)}$ was $4.67s$. In contrast, when initialized with a randomly chosen $\Theta^{(0)}$, the average execution time observed for the subsequent convergence of (6) was $82.2s$.

6. CONCLUSION

We have considered the optimal multi-agent persistent monitoring problem on a set of targets interconnected according to a fixed graph topology. We have adopted a class of distributed threshold-based parametric controllers where IPA can be used to determine optimal threshold parameters in an on-line manner using gradient descent. Due to the non-convex nature of the problem, optimal thresholds given by gradient descent highly depend on the initial thresholds. To address this issue, the asymptotic steady-state behavior of the persistent monitoring system is studied for a single agent and bi-triangular target configurations, leading to a computationally efficient and effective threshold initialization scheme. Ongoing work is directed at extending the analysis to multi-agent persistent monitoring and arbitrary, potentially sparse, graphs.

REFERENCES

Bertsekas, D.P. (2016). *Nonlinear Programming*. Athena Scientific.

Cassandras, C.G. and Lafortune, S. (2010). *Introduction to Discrete Event Systems*. Springer Publishing Company, Inc., 2nd edition.

Cassandras, C.G., Wardi, Y., Panayiotou, C.G., and Yao, C. (2010). Perturbation Analysis and Optimization of Stochastic Hybrid Systems. *European Journal of Control*, 16(6), 642–661.

Flanders, H. (1973). Differentiation Under the Integral Sign. *The American Mathematical Monthly*, 80(6), 615–627.

Huynh, V.A., Enright, J.J., and Frazzoli, E. (2010). Persistent patrol with limited-range on-board sensors. In *49th IEEE Conf. on Decision and Control*, 7661–7668.

Khazaeni, Y. and Cassandras, C.G. (2018). Event-driven trajectory optimization for data harvesting in multiagent systems. *IEEE Trans. on Control of Network Systems*, 5(3), 1335–1348.

Leahy, K., Zhou, D., Vasile, C.I., Oikonomopoulos, K., Schwager, M., and Belta, C. (2016). Persistent surveillance for unmanned aerial vehicles subject to charging and temporal logic constraints. *Autonomous Robots*, 40(8), 1363–1378.

Meng, X., Houshmand, A., and Cassandras, C.G. (2019). Multi-Agent Coverage Control with Energy Depletion and Repletion. In *58th IEEE Conf. on Decision and Control*, 2101–2106.

Trevathan, J. and Johnstone, R. (2018). Smart environmental monitoring and assessment technologies (Semat)a new paradigm for low-cost, remote aquatic environmental monitoring. *Sensors (Switzerland)*, 18(7).

Welikala, S. and Cassandras, C.G. (2019). Asymptotic Analysis for Greedy Initialization of Threshold-Based Distributed Optimization of Persistent Monitoring on Graphs. URL `http://arxiv.org/abs/1911.02658`.

Zhou, N., Cassandras, C.G., Yu, X., and Andersson, S.B. (2019). Optimal Threshold-Based Distributed Control Policies for Persistent Monitoring on Graphs. In *American Control Conference*, 2030–2035.

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. on Automatic Control*, 63(12), 4204–4217.