Joint 48th IEEE Conference on Decision and Control and
28th Chinese Control Conference
Shanghai, P.R. China, December 16-18, 2009

Distributed coverage games for mobile visual sensors (I): Reaching the set of Nash equilibria

Minghui Zhu and Sonia Martínez

Abstract— **We formulate a coverage optimization problem for mobile visual sensor networks as a repeated multi-player game. Each visual sensor tries to optimize its own coverage while minimizing the processing cost. The rewards for the sensing are not prior information for the agents. We present a synchronous distributed learning algorithm where each sensor only remembers its own utility values and actions played during the last two time steps. The algorithm is proven to be convergent in probability to the set of (restricted) Nash equilibria from which none has incentive to unilaterally deviate.**

I. Introduction

A substantial body of research on sensor networks has concentrated on simple sensors that can collect scalar data; e.g. temperature, humidity or pressure data. Thus, a main objective is the design of algorithms that can lead to optimal collective sensing through efficient motion control and communication schemes. However, scalar measurements can be insufficient in many situations; e.g. in automated surveillance or traffic monitoring applications. In contrast, cameras can collect visual data that are rich in information, thus having tremendous potential for monitoring applications, but at the cost of a higher processing overhead.

Precisely, this paper, part I, and its companion, part II, aim to solve a coverage optimization problem taking into account part of the sensing/processing trade-off. Coverage optimization problems have mainly been formulated as cooperative problems where each sensor benefits from sensing the environment as a member of the group. However, sensing may also require expenditure; e.g. the energy consumed by image processing in visual networks. Because of this, we will endow each sensor with a utility function that quantifies this trade-off, formulating a coverage problem as a variation of congestion games in [22].

Literature review. In broad terms, the problem studied here is related to a bevy of sensor location and planning problems in the computational geometry, geometric optimization, and robotics literature. For example, different variations on the (combinatorial) Art Gallery problem include [21][24][27]. The objective here is how to find the optimum number of guards in a non-convex environment so that each point is visible from at least one guard. A related set of references for the deployment of mobile robots with omnidirectional cameras includes [9][8]. Unlike the Art Gallery classic algorithms, the latter papers only assume that robots have local knowledge of the environment and no recollection of the past. Other related references on robot deployment in convex environments include [4][13] for anisotropic and circular footprints.

The paper [1] is an excellent survey on multimedia sensor networks where the state of the art in algorithms, protocols, and hardware is surveyed, and open research issues are discussed in detail. As observed in [5], multimedia sensor networks enhance traditional surveillance systems by enlarging the view, enhancing the view and enabling multi-resolution views. The investigation of coverage problems for static visual sensor networks is conducted in [3][10][25].

Another set of relevant references to this paper comprise those on the use of game-theoretic tools to (i) solve static target assignment problems, and (ii) devise efficient and secure algorithms for communication networks. In [15], the authors present a game-theoretic analysis of a coverage optimization for static sensor networks. This problem is equivalent to the weapon-target assignment problem in [20] which is nondeterministic polynomial-time-complete. In general, the solution to assignment problems is hard from a combinatorial optimization viewpoint.

Game Theory and Learning in Games are used to analyze a variety of fundamental problems in; e.g. wireless communication networks and the Internet. An incomplete list of references includes [2] on power control, [23] on routing, and [26] on flow control. However, there has been limited research on how to employ Learning in Games to develop distributed algorithms for mobile sensor networks. One exception is the paper [14] where the authors establish a link between cooperative control problems (in particular, consensus problems) and games (in particular, potential games and weakly acyclic games).

Statement of contributions. The contributions of this part I paper pertain to both coverage optimization problems and Learning in Games. Compared with [12] and [13], this paper employs a more accurate sensing model and the results can be easily extended to include non-convex environments. Contrary to [12], we do not consider energy expenditure from sensor motions. Regarding Learning in Games, we extend the use of the payoff-based learning dynamics in [16][17]. In our problem, each agent is unable to access the utility values induced by alternative actions because motion and sensing capacities of each agent are limited and the reward is not a priori information to each agent. To tackle this challenge, we develop a distributed synchronous learning algorithm which only requires each sensor to remember its own utility values and actions played during the last two time steps. The algorithm is concisely described as follows: At each time

This work was supported in part by NSF Career Award CMS-0643673 and NSF IIS-0712746. The authors are with Department of Mechanical and Aerospace Engineering, University of California, San Diego, 9500 Gilman Dr, La Jolla CA, 92093, {mizhu,soniamd}@ucsd.edu

978-1-4244-3872-3/09/$25.00 ©2009 IEEE

step, each sensor repeatedly updates its action, either trying some new action or selecting the action which corresponds to a higher utility value in the most recent two time steps. This algorithm is shown to be convergent in probability to the set of (restricted) Nash equilibria from which no agent is willing to unilaterally deviate. The main advantages of our learning algorithm over those in [17], are its simplicity and stronger convergence properties; see also Remark 4.1

II. PROBLEM FORMULATION AND LEARNING ALGORITHM

Here, we first review some basic game-theoretic concepts; see, for example [7]. This will allow us to formulate subsequently an optimal coverage problem for mobile visual sensor networks as a repeated multi-player game. We then present an algorithm to solve the coverage game, and introduce notation used throughout the paper.

A. Background in Game Theory

A strategic game $\Gamma := \langle V, A, U\rangle$ has three components:

1. A set V enumerating players $i \in V := \{1, \cdots, N\}$.
2. An action set $A := \prod_{i=1}^{N} A_i$ is the space of all actions vectors, where $s_i \in A_i$ is the action of player i and an (multi-player) action $s \in A$ has components $s_1, \dots, s_N$.
3. The collection of utility functions U, where the utility function $u_i : A \to \mathbb{R}$ models player i's preferences over action profiles.

Denote by s_{-i} the action profile of all players other than i, and by $A_{-i} = \prod_{j\neq i} A_j$ the set of action profiles for all players except i. The concept of (pure) Nash equilibrium (NE, for short) is the most important one in Non-cooperative Game Theory [7] and is defined as follows.

Definition 2.1 (Nash equilibrium [7]): Consider the strategic game Γ. An action profile $s^* := (s_i^*, s_{-i}^*)$ is an NE of the game Γ if $\forall i \in V$ and $\forall s_i \in A_i$, it holds that $u_i(s^*) \geq u_i(s_i, s_{-i}^*)$. •

An action profile corresponding to an NE represents a scenario where no player has incentive to unilaterally deviate. Potential Games form an important class of strategic games where the change in a player's utility caused by a unilateral deviation can be measured by a potential function.

Definition 2.2 (Potential game [19]): The strategic game Γ is a potential game with potential function $\phi : A \to \mathbb{R}$ if for every $i \in V$, for every $s_{-i} \in A_{-i}$, and for every $s_i, s_i' \in A_i$, it holds that

$$\phi(s_i, s_{-i}) - \phi(s_i', s_{-i}) = u_i(s_i, s_{-i}) - u_i(s_i', s_{-i}). \quad (1)$$

•

In conventional Non-cooperative Game Theory, all the actions in A_i always can be selected by player i in response to other players' actions. However, in the context of motion coordination, the actions available to player i will often be restricted to a state-dependent subset of A_i. In particular, we denote by $F_i(s_i, s_{-i}) \subseteq A_i$ the set of feasible actions of player i when the action profile is $s := (s_i, s_{-i})$. We assume that $F_i(s_i, s_{-i}) \neq \emptyset$. Denote $F(s) := \prod_{i\in V} F_i(s) \subseteq A$, $\forall s \in A$ and $F := \cup\{F(s) \mid s \in A\}$. The introduction of F leads naturally to the notion of restricted strategic game $\Gamma_{\text{res}} := \langle V, A, U, F\rangle$, and the following associated concepts.

Definition 2.3 (Restricted Nash equilibrium): Consider the restricted strategic game Γ_{res}. An action profile s^* is a restricted NE of the game Γ_{res} if $\forall i \in V$ and $\forall s_i \in F_i(s_i^*, s_{-i}^*)$, it holds that $u_i(s^*) \geq u_i(s_i, s_{-i}^*)$. •

Definition 2.4 (Restricted potential game): The game Γ_{res} is a restricted potential game with potential function $\phi(s)$ if for every $i \in V$, every $s_{-i} \in A_{-i}$, and every $s_i \in A_i$, the equality (1) holds for every $s_i' \in F_i(s_i, s_{-i})$. •

Observe that if s^* is an NE of the strategic game Γ, then it is also a restricted NE of the restricted strategic game Γ_{res}. For any given strategic game, NE may not exist. However, the existence of NEs in potential games is guaranteed [19]. Hence, any restricted potential game has at least one restricted NE.

B. Coverage problem formulation

1) Mission space: We consider a convex 2-D mission space that is discretized into a (squared) lattice. We assume that each square of the lattice has unit dimensions. Each square will be labeled with the coordinate of its center $q = (q_x, q_y)$, where $q_x \in [q_{x_{\min}}, q_{x_{\max}}]$ and $q_y \in [q_{y_{\min}}, q_{y_{\max}}]$, for some integers $q_{x_{\min}}, q_{y_{\min}}, q_{x_{\max}}, q_{y_{\max}}$. Denote by $\mathcal{Q}$ the collection of all squares in the lattice.

We now define an associated location graph $\mathcal{G}_{\text{loc}} := (\mathcal{Q}, E_{\text{loc}})$ where $((q_x, q_y), (q_{x'}, q_{y'})) \in E_{\text{loc}}$ if and only if $|q_x - q_{x'}| + |q_y - q_{y'}| = 1$ for $(q_x, q_y), (q_{x'}, q_{y'}) \in \mathcal{Q}$. Note that the graph $\mathcal{G}_{\text{loc}}$ is undirected; i.e., $(q, q') \in E_{\text{loc}}$ if and only if $(q', q) \in E_{\text{loc}}$. The set of neighbors of q in E_{loc} is given by $\mathcal{N}_q^{\text{loc}} := \{q' \in \mathcal{Q} \setminus \{q\} \mid (q, q') \in E_{\text{loc}}\}$. We assume that the location graph $\mathcal{G}_{\text{loc}}$ is fixed and connected, and denote its diameter by D.

Agents are deployed in $\mathcal{Q}$ to detect certain events of interest. As agents move in $\mathcal{Q}$ and process measurements, they will assign a numerical value $W_q \geq 0$ to the events in each square (with center) $q \in \mathcal{Q}$. If $W_q = 0$, then there is no event of interest at the square q. The larger the value of W_q is, the more interest the set of events at the square q is of. Later, the amount W_q will be identified with a benefit of observing the point q. In this set-up, we assume the values W_q to be constant in time.

2) Modeling of the visual sensor nodes: Each mobile agent i is modeled as a point mass in $\mathcal{Q}$, with location $a_i := (x_i, y_i) \in \mathcal{Q}$. Each agent has mounted a pan-tilt-zoom camera, and can adjust its orientation and focal length.

The visual sensing range of a camera is directional, limited-range, and has a finite angle of view. Following a geometric simplification, we model the visual sensing region of agent i as an annulus sector in the 2-D plane; see Figure 1.

The visual sensor footprint is completely characterized by the following parameters: the position of agent i, $a_i \in \mathcal{Q}$, the camera orientation, $\theta_i \in [0, 2\pi)$, the camera angle of view, $\alpha_i \in [\alpha_{\min}, \alpha_{\max}]$, and the shortest range (resp. longest range) between agent i and the nearest (resp. farthest) object that can be recognized from the image, $r_i^{\text{shrt}} \in [r_{\min}, r_{\max}]$ (resp. $r_i^{\text{lng}} \in [r_{\min}, r_{\max}]$). The parameters r_i^{shrt}, r_i^{lng}, α_i can be tuned by

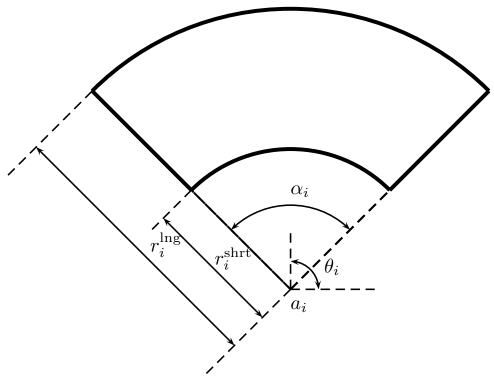

Fig. 1. Visual sensor footprint

changing the focal length FL_i of agent i's camera. In this way, $c_i := (\mathrm{FL}_i, \theta_i) \in [0, \mathrm{FL}_{\max}] \times [0, 2\pi)$ is the camera control vector of agent i. In what follows, we will assume that c_i takes values in a finite subset $\mathcal{C} \subset [0, \mathrm{FL}_{\max}] \times [0, 2\pi)$. An agent action is thus a vector $s_i := (a_i, c_i) \in \mathcal{A}_i := \mathcal{Q} \times \mathcal{C}$, and a multi-agent action is denoted by $s = (s_1, \ldots, s_N) \in \mathcal{A} := \Pi_{i=1}^N \mathcal{A}_i$.

Let $\mathcal{D}(a_i, c_i)$ be the visual sensor footprint of agent i. Now we can define a proximity sensing graph [**?**] $\mathcal{G}_{\text{sen}}(s) := (V, E_{\text{sen}}(s))$ as follows: the set of neighbors of agent i, $\mathcal{N}_i^{\text{sen}}(s)$, is given as $\mathcal{N}_i^{\text{sen}}(s) := \{j \in V \backslash \{i\} \mid \mathcal{D}(a_i, c_i) \cap \mathcal{D}(a_j, c_j) \cap \mathcal{Q} \neq \emptyset\}$.

Each agent is able to communicate with others to exchange information. We assume that the communication range of agents is $2r_{\max}$. This induces a $2r_{\max}$-disk communication graph $\mathcal{G}_{\text{comm}}(s) := (V, E_{\text{comm}}(s))$ as follows: the set of neighbors of agent i is given by $\mathcal{N}_i^{\text{comm}}(s) := \{j \in V \backslash \{i\} \mid (x_i - x_j)^2 + (y_i - y_j)^2 \leq (2r_{\max})^2\}$. Note that $\mathcal{G}_{\text{comm}}(s)$ is undirected and that $\mathcal{G}_{\text{sen}}(s) \subseteq \mathcal{G}_{\text{comm}}(s)$.

The motion of agents will be limited to a neighboring point in $\mathcal{G}_{\text{loc}}$ at each time step. Thus, an agent feasible action set will be given by $\mathcal{F}(a_i) := (\{a_i\} \cup \mathcal{N}_{a_i}^{\text{loc}}) \times \mathcal{C}$.

3) Coverage game: We now proceed to formulate a coverage optimization problem as a restricted strategic game. For each $q \in \mathcal{Q}$, we denote $n_q(s)$ as the cardinality of the set $\{k \in V \mid q \in \mathcal{D}(a_k, c_k) \cap \mathcal{Q}\}$; i.e., the number of agents which can observe the point q. The "profit" given by W_q will be equally shared by agents that can observe the point q. The benefit that agent i obtains through sensing is thus defined by $\sum_{q \in \mathcal{D}(a_i, c_i) \cap \mathcal{Q}} \frac{W_q}{n_q(s)}$.

On the other hand, and as argued in [18], the processing of visual data can incur a higher cost than that of communication. This is in contrast with scalar sensor networks, where the communication cost dominates. With this observation, we model the energy consumption of agent i by $f_i(c_i) := \frac{1}{2}\alpha_i((r_i^{\text{lng}})^2 - (r_i^{\text{shrt}})^2)$. This measure corresponds to the area of the visual sensor footprint and can serve to approximate the energy consumption or the cost incurred by image processing algorithms.

We will endow each agent with a utility function that aims to capture the above sensing/processing trade-off. In this way, we define a utility function for agent i by

$$u_i(s) = \sum_{q \in \mathcal{D}(a_i, c_i) \cap \mathcal{Q}} \frac{W_q}{n_q(s)} - f_i(c_i).$$

In our set-up, we assume that W_q is unknown to each agent i unless agent i senses the point q. Note that the utility function u_i is distributed over the visual sensing graph $\mathcal{G}_{\text{sen}}(s)$; i.e., u_i is only dependent on the points q within its sensing range $\mathcal{D}(a_i, c_i)$ and the actions of $\{i\} \cup \mathcal{N}_i^{\text{sen}}(s)$. With the set of utility functions $U_{\text{cov}} = \{u_i\}_{i \in V}$, and feasible action set $\mathcal{F}_{\text{cov}} = \Pi_{i=1}^N \bigcup_{a_i \in \mathcal{A}_i} \mathcal{F}(a_i)$, we now have all the ingredients to introduce the coverage game $\Gamma_{\text{cov}} := \langle V, \mathcal{A}, U_{\text{cov}}, \mathcal{F}_{\text{cov}} \rangle$. This game is a variation of the congestion games introduced in [22].

Lemma 2.1: The coverage game Γ_{cov} is a restricted potential game with potential function

$$\phi(s) = \sum_{q \in \mathcal{Q}} \sum_{\ell=1}^{n_q(s)} \frac{W_q}{\ell} - \sum_{i=1}^{N} f_i(c_i).$$

We denote by $\mathcal{E}(\Gamma_{\text{cov}})$ the set of restricted NEs of Γ_{cov}. It is worthy to mention that $\mathcal{E}(\Gamma_{\text{cov}}) \neq \emptyset$ due to the fact of Γ_{cov} being a restricted potential game.

Remark 2.1: The assumptions of our problem formulation admit several extensions. For example, it is straightforward to extend our results to non-convex 3-D spaces. This is because the results that follow can also handle other shapes of the sensor footprint; e.g., a complete disk, a subset of the annulus sector. On the other hand, note that the coverage problem can be interpreted as a target assignment problem – here, the value $W_q \geq 0$ would be associated with the value of a target located at the point q.

C. Inhomogeneous synchronous learning algorithm

A number of learning rules, e.g. better (or best) reply dynamics and adaptive play, have been proposed to reach Nash equilibria in potential games. In these algorithms, each player must have access to the utility values induced by alternative actions. In our problem, however, this information is unaccessible because motion and sensing capacities of each agent are limited and W_q is not priori information to each agent. To tackle this challenge, we present a distributed learning algorithm, say the Inhomogeneous Synchronous Learning (ISL) Algorithm, which only requires each sensor to remember its own utility values and actions played during the last two time steps. The ISL algorithm will be shown to be convergent to the set $\mathcal{E}(\Gamma_{\text{cov}})$ in probability. Let $\tau_i(t) = \min \operatorname{argmax}_{v \in \{t, t-1\}} u_i(s(v))$, for all $t \geq 1$ and $i \in V$. The main steps of the ISL Algorithm are the following:

1: [Initialization] At $t = 0$, all agents are uniformly placed in $\mathcal{Q}$. Each agent i uniformly chooses its camera control vector c_i from the set $\mathcal{C}$, communicates with agents in $\mathcal{N}_i^{\text{sen}}(s(0))$, and computes $u_i(s(0))$. At $t = 1$, all the agents keep their actions.

2: [Update] At each time $t \geq 2$, each agent i updates its state according to the following rules:

- Agent i chooses the exploration rate $\epsilon(t) = t^{-\frac{1}{N(D+1)}}$ and compute $s_i(\tau_i(t))$.
- With probability $\epsilon(t)$, agent i experiments, and chooses the temporary action $s_i^{\text{tp}} := (a_i^{\text{tp}}, c_i^{\text{tp}})$ uniformly from the set $\mathcal{F}(a_i(t)) \setminus \{s_i(\tau_i(t))\}$.
- With probability $1-\epsilon(t)$, agent i does not experiment, and sets $s_i^{\text{tp}} = s_i(\tau_i(t))$.
- After s_i^{tp} is chosen, agent i moves to the position a_i^{tp} and sets the camera control vector to c_i^{tp}.

3: [Communication and computation] At position a_i^{tp}, agent i communicates with agents in $\mathcal{N}_i^{\text{sen}}(s_i^{\text{tp}}, s_{-i}^{\text{tp}})$, and computes $u_i(s_i^{\text{tp}}, s_{-i}^{\text{tp}})$ and $\mathcal{F}(a_i^{\text{tp}})$.

4: Repeat Step 2 and 3.

Remark 2.2: A variation of the previous algorithm corresponds to $\epsilon(t) = \epsilon \in (0, \frac{1}{2}]$ constant for all $t \geq 2$. If this is the case, we will refer to the algorithm as the Homogeneous Synchronous Learning (HSL, for short) Algorithm. Later, the convergence analysis of the ISL will be based on the analysis of the HSL.

D. Notation

In the following, we will use the Landau symbol, O, as in $O(\epsilon^k)$, for some $k \geq 0$. This implies that $0 < \lim_{\epsilon\to 0^+} \frac{O(\epsilon^k)}{\epsilon^k} < \infty$. We denote by $\text{diag}(\mathcal{A}) := \{(s,s) \in \mathcal{A}^2 \mid s \in \mathcal{A}\}$ and $\text{diag}(\mathcal{E}(\Gamma_{\text{cov}})) := \{(s,s) \in \mathcal{A}^2 \mid s \in \mathcal{E}(\Gamma_{\text{cov}})\}$.

Consider $a, a' \in \mathcal{Q}^N$ where $a_i \neq a_i'$ and $a_{-i} = a_{-i}'$ for some $i \in V$. The transition $a \to a'$ is feasible if and only if $(a_i, a_i') \in E_{\text{loc}}$. A feasible path from a to a' consisting of multiple feasible transitions is denoted by $a \Rightarrow a'$. Let $\diamond a := \{a' \in \mathcal{Q} \mid a \Rightarrow a'\}$ be the reachable set from a.

Let $s = (a,c), s' = (a',c') \in \mathcal{A}$ where $a_i \neq a_i'$ and $a_{-i} = a_{-i}'$ for some $i \in V$. The transition $s \to s'$ is feasible if and only if $s_i' \in \mathcal{F}(a)$. A feasible path from s to s' consisting of multiple feasible transitions is denoted by $s \Rightarrow s'$. Finally, $\diamond s := \{s' \in \mathcal{A} \mid s \Rightarrow s'\}$ will be the reachable set from s.

III. Preliminaries to convergence analysis

For the sake of a self-contained exposition, we include here some background in the Theory of Resistance Trees [28]. This section also includes a sufficient condition on the convergence of a class of time-inhomogeneous Markov chains that will be used in the general algorithm proof later.

A. Background in the Theory of Resistance Trees

Let P^0 be the transition matrix of the time-homogeneous Markov chain $\{\mathcal{P}_t^0\}$ on a finite state space X. And let P^ϵ be the transition matrix of a *perturbed Markov chain*, say $\{\mathcal{P}_t^\epsilon\}$. With probability $1-\epsilon$, the process $\{\mathcal{P}_t^\epsilon\}$ evolves according to P^0, while with probability ϵ, the transitions do not follow P^0.

A family of stochastic processes $\{\mathcal{P}_t^\epsilon\}$ is called a *regular perturbation* of $\{\mathcal{P}_t^0\}$ if the following holds $\forall x, y \in X$:

(A1) For some $\varsigma > 0$, the Markov chain $\{\mathcal{P}_t^\epsilon\}$ is irreducible and aperiodic for all $\epsilon \in (0, \varsigma]$.

(A2) $\lim_{\epsilon\to 0^+} P_{xy}^\epsilon = P_{xy}^0$.

(A3) If $P_{xy}^\epsilon > 0$ for some ϵ, then there exists a real number $\chi(x \to y) \geq 0$ such that $\lim_{\epsilon\to 0^+} P_{xy}^\epsilon / \epsilon^{\chi(x\to y)} \in (0,\infty)$.

In (A3), the nonnegative real number $\chi(x \to y)$ is called the resistance of the transition from x to y.

Let $H_1, H_2, \cdots, H_J$ be the recurrent communication classes of the Markov chain $\{\mathcal{P}_t^0\}$. Note that within each class H_ℓ, there is a path of zero resistance from every state to every other. Given any two distinct recurrence classes H_ℓ and H_k, consider all paths which start from H_ℓ and end at H_k. Denote $\chi_{\ell k}$ by the least resistance among all such paths.

Now define a complete directed graph $\mathcal{G}$ where there is one vertex ℓ for each recurrent class H_ℓ, and the resistance on the edge (ℓ, k) is $\chi_{\ell k}$. An *ℓ-tree* on $\mathcal{G}$ is a spanning tree such that from every vertex $k \neq \ell$, there is a unique path from k to ℓ. Denote by $G(\ell)$ the set of all *ℓ-trees* on $\mathcal{G}$. The resistance of an *ℓ-tree* is the sum of the resistances of its edges. The *stochastic potential* of the recurrent class H_ℓ is the least resistance among all *ℓ-trees* in $G(\ell)$.

Theorem 3.1 ([28]): Let $\{\mathcal{P}_t^\epsilon\}$ be a regular perturbation of $\{\mathcal{P}_t^0\}$, and for each $\epsilon > 0$, let $\mu(\epsilon)$ be the unique stationary distribution of $\{\mathcal{P}_t^\epsilon\}$. Then $\lim_{\epsilon\to 0^+} \mu(\epsilon)$ exists and the limiting distribution $\mu(0)$ is a stationary distribution of $\{\mathcal{P}_t^0\}$. The stochastically stable states (i.e., the support of $\mu(0)$) are precisely those states contained in the recurrence classes with minimum stochastic potential.

B. Convergence of a class of time-inhomogeneous Markov chains

Here we derive sufficient conditions for a class of time-inhomogeneous Markov chains to converge. The main references include [6] and [11].

Consider now a time-inhomogeneous Markov chain $\{\mathcal{P}_t\}$ on a finite state space X with transition matrix $P^{\epsilon(t)}$ where $\epsilon(t) \in (0, \varsigma]$ for some $\varsigma > 0$. Let P^ϵ be the transition matrix if $\epsilon(t)$ is a constant $\epsilon(t) \in (0, \varsigma]$ for all $t \geq 1$. Denote by $\{\mathcal{P}_t^\epsilon\}$ the time-homogeneous Markov chain corresponding to P^ϵ.

Proposition 3.1: Assume that, $\{\mathcal{P}_t^\epsilon\}$ is a regular perturbation of $\{\mathcal{P}_t^0\}$. The time-inhomogeneous Markov chain $\{\mathcal{P}_t\}$ is strongly ergodic if the following conditions hold:

(C1) The Markov chain $\{\mathcal{P}_t\}$ is weakly ergodic.

(C2) $\epsilon(t) > 0$ and is strictly decreasing.

(C3) If $P_{xy}^{\epsilon(t)} > 0$, then $P_{xy}^{\epsilon(t)} = \alpha_{xy}(\epsilon(t))/\beta_{xy}(\epsilon(t))$ for some polynomials $\alpha_{xy}(\epsilon(t))$ and $\beta_{xy}(\epsilon(t))$ in $\epsilon(t)$.

Proof: We omit the proof due to the space limit. ■

Remark 3.1: In Proposition 3.1, (C3) can be replaced by the following. (C3') If $P_{xy}^{\epsilon(t)} > 0$, then $P_{xy}^{\epsilon(t)} = \alpha_{xy}(\epsilon(t))/\beta_{xy}(\epsilon(t))$ where $\alpha_{xy}(\epsilon(t))$ and $\beta_{xy}(\epsilon(t))$ are smooth at the origin. Following the same lines in Proposition 3.1, one can complete the proof by utilizing the Taylor expansions of $\alpha_{xy}(\epsilon(t))$ and $\beta_{xy}(\epsilon(t))$ at the origin.

IV. Convergence analysis of the ISL Algorithm

In this section, we analyze the convergence properties of the ISL Algorithm to the set $\mathcal{E}(\Gamma_{\text{cov}})$ by appealing to the results in Section III. Relevant papers include [16][17] on payoff-based learning algorithms.

A. Convergence analysis of the associated HSL Algorithm

We first utilize Proposition 3.1 to characterize the convergence properties of the associated HSL algorithm. This is essential for the analysis of the ISL Algorithm.

Denote the space $\mathcal{B} := \{(s, s') \in \mathcal{A} \times \mathcal{A} \mid s'_i \in \mathcal{F}(a_i),\ \forall i \in V\}$. Observe that $z(t) := (s(t-1), s(t))$ in the HSL Algorithm constitutes a time-homogeneous Markov chain $\{\mathcal{P}^\epsilon_t\}$ on the space $\mathcal{B}$. Consider $z, z' \in \mathcal{B}$. A feasible path from z to z' consisting of multiple feasible transitions of $\{\mathcal{P}^\epsilon_t\}$ is denoted by $z \Rightarrow z'$. The reachable set from z is denoted as $\diamond z := \{z' \in \mathcal{B} \mid z \Rightarrow z'\}$.

Lemma 4.1: $\{\mathcal{P}^\epsilon_t\}$ is a regular perturbation of $\{\mathcal{P}^0_t\}$.

Proof: We omit the proof due to the space limit. ■

Lemma 4.2: For any $(s^0, s^0) \in \mathrm{diag}(\mathcal{A}) \setminus \mathrm{diag}(\mathcal{E}(\Gamma_{\mathrm{cov}}))$, there is a finite sequence of transitions from (s^0, s^0) to some $(s^*, s^*) \in \mathrm{diag}(\mathcal{E}(\Gamma_{\mathrm{cov}}))$ that satisfies

$$\mathcal{L} := (s^0, s^0) \overset{O(\epsilon)}{\rightarrow} (s^0, s^1) \overset{O(1)}{\rightarrow} (s^1, s^1) \overset{O(\epsilon)}{\rightarrow} (s^1, s^2) \overset{O(1)}{\rightarrow} (s^2, s^2) \overset{O(\epsilon)}{\rightarrow} \cdots \overset{O(\epsilon)}{\rightarrow} (s^{k-1}, s^k) \overset{O(1)}{\rightarrow} (s^k, s^k)$$

where $(s^k, s^k) = (s^*, s^*)$ for some $k \geq 1$.

Proof: We omit the proof due to the space limit. ■

A direct result of Lemma 4.1 is that for each ϵ, there exists a unique stationary distribution of $\{\mathcal{P}^\epsilon_t\}$, say $\mu(\epsilon)$. We now proceed to utilize Theorem 3.1 to characterize $\lim_{\epsilon\to 0^+} \mu(\epsilon)$.

Proposition 4.1: Consider the regular perturbation $\{\mathcal{P}^\epsilon_t\}$ of $\{\mathcal{P}^0_t\}$. Then $\lim_{\epsilon\to 0^+} \mu(\epsilon)$ exists and the limiting distribution $\mu(0)$ is a stationary distribution of $\{\mathcal{P}^0_t\}$. Furthermore, the stochastically stable states (i.e., the support of $\mu(0)$) are contained in the set $\mathrm{diag}(\mathcal{E}(\Gamma_{\mathrm{cov}}))$.

Proof: The only candidates for the stochastically stable states are the recurrent communication classes of the unperturbed Markov chain that corresponds to the HSL Algorithm with $\epsilon = 0$. Thus the recurrent communication classes are the states in the set $\mathrm{diag}(\mathcal{A}) \subset \mathcal{B}$. Denote by $T_{\min}$ the minimum resistance tree and by h_v the root of $T_{\min}$. Each edge of $T_{\min}$ has resistance $0, 1, 2, \ldots$ corresponding to the transition probability $O(1), O(\epsilon), O(\epsilon^2), \ldots$. The state z' is the *successor* of the state z if and only if $(z, z') \in T_{\min}$. We will construct $T_{\min}$ over states in the set $\mathcal{B}$ (rather than $\mathrm{diag}(\mathcal{A})$) with the restriction that all the edges leaving the states in $\mathcal{B} \setminus \mathrm{diag}(\mathcal{A})$ have resistance 0. This difference will not affect the set of the stochastically stable states. The remainder of the proofs will be based on the following four claims. Due to the space limit, we omit the details here.

Claim 1: For any $(s^0, s^1) \in \mathcal{B} \setminus \mathrm{diag}(\mathcal{A})$, there is a finite path

$$\mathcal{L}' := (s^0, s^1) \overset{O(1)}{\rightarrow} (s^1, s^2) \overset{O(1)}{\rightarrow} (s^2, s^2)$$

where $s^2_i = s_i^{\tau_i(0,1)}$ for all $i \in V$.

Claim 2: The root h_v belongs to the set $\mathrm{diag}(\mathcal{A})$.

Claim 3: Pick any $s^* \in \mathcal{E}(\Gamma_{\mathrm{cov}})$ and consider $z := (s^*, s^*)$, $z' := (s^*, \tilde{s})$ where $\tilde{s} \neq s^*$. If $(z, z') \in T_{\min}$, then the resistance of the edge (z, z') is some $k \geq 2$.

Claim 4: The root h_v belongs to the set $\mathrm{diag}(\mathcal{E}(\Gamma_{\mathrm{cov}}))$.

Proof of Proposition 4.1: It follows from Claim 4 that the states in $\mathrm{diag}(\mathcal{E}(\Gamma_{\mathrm{cov}}))$ have minimum stochastic potential. Since Lemma 4.1 shows that Markov chain $\{\mathcal{P}^\epsilon_t\}$ is a regularly perturbed Markov process, Proposition 4.1 is a direct result of Theorem 3.1. ■

B. Convergence analysis of ISL Algorithm

Observe that $z(t) := (s(t-1), s(t))$ in the ISL Algorithm constitutes a time-inhomogeneous Markov chain $\{\mathcal{P}_t\}$ on the space $\mathcal{B}$. We will combine Proposition 4.1 and Proposition 3.1 to characterize the convergence property of the ISL Algorithm.

Theorem 4.1: Consider the Markov chain $\{\mathcal{P}_t\}$ induced by the ISL Algorithm. It holds that $\lim_{t\to\infty} \mathbb{P}(z(t) \in \mathrm{diag}(\mathcal{E}(\Gamma_{\mathrm{cov}}))) = 1$.

Proof: Denote by $P^{\epsilon(t)}$ the transition matrix of $\{\mathcal{P}_t\}$. It is obvious that (C2) in Proposition 3.1 holds. The probability of the feasible transition $z^1 \rightarrow z^2$ is given by

$$P^{\epsilon(t)}_{z^1 z^2} = \prod_{i\in\Lambda_1} (1 - \epsilon(t)) \times \prod_{j\in\Lambda_2} \frac{\epsilon(t)}{|\mathcal{F}(a^1_i)| - 1}.$$

Then (C3) in Proposition 3.1 is satisfied. We now proceed to verify weak ergodicity by using Theorem V.3.2 in [11]. Observe that $|\mathcal{F}(a^1_i)| \leq 5|\mathcal{C}|$. Since $\epsilon(t)$ is strictly decreasing, there is $t_0 \geq 1$ such that t_0 is the first time when $1 - \epsilon(t) \geq \frac{\epsilon(t)}{5|\mathcal{C}|-1}$. Then for all $t \geq t_0$, it holds that

$$P^{\epsilon(t)}_{z^1 z^2} \geq (\frac{\epsilon(t)}{5|\mathcal{C}| - 1})^N.$$

Denote $P(m, n) := \prod_{t=m}^{n-1} P^{\epsilon(t)}$, $0 \leq m < n$. Pick any $z \in \mathcal{B}$ and let $u_z \in \mathcal{B}$ be such that $P_{u_z z}(t, t+D+1) = \min_{x\in\mathcal{B}} P_{xz}(t, t+D+1)$. Consequently, it follows that for all $t \geq t_0$,

$$\begin{aligned}
&\min_{x\in\mathcal{B}} P_{xz}(t, t+D+1)\\
&= \sum_{i_1\in\mathcal{B}} \cdots \sum_{i_D\in\in\mathcal{B}} P^{\epsilon(t)}_{u_z i_1} \cdots P^{\epsilon(t+D-1)}_{i_{D-1} i_D} P^{\epsilon(t+D)}_{i_D z}\\
&\geq P^{\epsilon(t)}_{u_z i_1} \cdots P^{\epsilon(t+D-1)}_{i_{D-1} i_D} P^{\epsilon(t+D)}_{i_D z}\\
&\geq \prod_{i=0}^{D} (\frac{\epsilon(t+i)}{5|\mathcal{C}| - 1})^N\\
&\geq (\frac{\epsilon(t)}{5|\mathcal{C}| - 1})^{(D+1)N}
\end{aligned}$$

where in the last inequality we use the fact that $\epsilon(t)$ is strictly decreasing. Then we have

$$\begin{aligned}
&1 - \lambda(P(t, t+D+1))\\
&= \min_{x,y\in\mathcal{B}} \sum_{z\in\mathcal{B}} \min\{P_{xz}(t, t+D+1), P_{yz}(t, t+D+1)\}\\
&\geq \sum_{z\in\mathcal{B}} \min_{x\in\mathcal{B}} P_{xz}(t, t+D+1)\\
&\geq \sum_{z\in\mathcal{B}} P_{u_z z}(t, t+D+1)\\
&\geq |\mathcal{B}|(\frac{\epsilon(t)}{5|\mathcal{C}| - 1})^{(D+1)N}.
\end{aligned}$$

Choose $k_i := (D+1)i$ and let i_0 be the smallest integer such that $(D+1)i_0 \geq t_0$. Then it holds that

$$\begin{aligned}
&\sum_{i=0}^{\infty}(1-\lambda(P(k_i,k_{i+1}))) \\
&\geq \sum_{i=i_0}^{\infty}(1-\lambda(P(k_i,k_{i+1}))) \\
&\geq |\mathcal{B}| \sum_{i=i_0}^{\infty}(\frac{\epsilon((D+1)i)}{5|\mathcal{C}|-1})^{(D+1)N} \\
&= \frac{|\mathcal{B}|}{(5|\mathcal{C}|-1)^{(D+1)N}} \sum_{i=i_0}^{\infty} \frac{1}{(D+1)i} = \infty.
\end{aligned}$$

Hence, the weak ergodicity follows from Theorem V.3.2 in [11]. The strong ergodicity follows directly from Proposition 3.1. It follows from Theorem V.4.3 in [11] that the limiting distribution is $\mu^* = \lim_{t\to\infty}\mu^t$. Note that $\lim_{t\to\infty}\mu^t = \lim_{t\to\infty}\mu(\epsilon(t)) = \mu(0)$ and Proposition 4.1 shows that the support of $\mu(0)$ is contained in the set $\mathrm{diag}(\mathcal{E}(\Gamma_{\mathrm{cov}}))$. Hence, the support of μ^* is contained in the set $\mathrm{diag}(\mathcal{E}(\Gamma_{\mathrm{cov}}))$. ■

Remark 4.1: Compared with the payoff-based learning algorithms in [17], the dynamics of the ISL algorithm involves less variables and is simpler. Furthermore, the algorithms in [17] converge to the set of Nash equilibria with sufficiently large probability by choosing a sufficiently small exploration rate in advance, and the induced evolution is a time-homogeneous Markov chain. In contrast, our ISL algorithm employs a diminishing exploration rate. This leads to the evolution of the ISL algorithm being a time-inhomogeneous Markov chain and a stronger convergence property of reaching the set of Nash equilibria in probability.

V. Conclusion

We have formulated a coverage optimization problem as a restricted potential game. We have proposed a synchronous distributed learning algorithm for this coverage game and shown that this algorithm asymptotically converges to the set of restricted NEs in probability. In our companion paper [29], we present an asynchronous distributed learning algorithm which converges in probability to the set of global optima of certain coverage performance metric.

References

[1] I.F. Akyildiz, T. Melodia, and K. Chowdhury. Wireless multimedia sensor networks: a survey. *IEEE Wireless Communications Magazine*, 14(6):32–39, 2007.

[2] T. Alpcan, T. Basar, and S. Dey. A power control game based on outage probabilities for multicell wireless data networks. *IEEE Transactions on Wireless Communications*, 5(4):890–899, 2006.

[3] F. Bullo, J. Cortés, and S. Martínez. *Distributed Control of Robotic Networks*. Applied Mathematics Series. Princeton University Press, 2009. Available at http://www.coordinationbook.info.

[4] K.Y. Chow, K.S. Lui, and E.Y. Lam. Maximizing angle coverage in visual sensor networks. In *Proc. of IEEE International Conference on Communications*, pages 3516–3521, June 2007.

[5] J. Cortés, S. Martínez, and F. Bullo. Spatially-distributed coverage optimization and control with limited-range interactions. *ESAIM: Control, Optimisation & Calculus of Variations*, 11:691–719, 2005.

[6] R. Cucchiara. Multimedia surveillance systems. In *Proc. of the third ACM international workshop on Video surveillance and sensor networks*, pages 3–10, 2005.

[7] M. Freidlin and A. Wentzell. *Random perturbations of dynamical systems*. New York: Springer Verlag, 1984.

[8] D. Fudenberg and J. Tirole. *Game theory*. The MIT press, 1991.

[9] A. Ganguli, J. Cortés, and F. Bullo. Visibility-based multi-agent deployment in orthogonal environments. In *American Control Conference*, pages 3426–3431, New York, July 2007.

[10] A. Ganguli, J. Cortés, and F. Bullo. Multirobot rendezvous with visibility sensors in nonconvex environments. *IEEE Transactions on Robotics*, 25(2):340–352, 2009.

[11] E. Hörster and R. Lienhart. On the optimal placement of multiple visual sensors. In *Proc. of the third ACM international workshop on Video surveillance and sensor networks*, pages 111–120, 2006.

[12] D. Isaacson and R. Madsen. *Markov chains*. Wiley, 1976.

[13] A. Kwok and S. Martínez. Deployment algorithms for a power-constrained mobile sensor network. *International Journal of Robust and Nonlinear Control*, 2009. In press.

[14] K. Laventall and J. Cortés. Coverage control by multi-robot networks with limited-range anisotropic sensory. *International Journal of Control*, 2009. To appear.

[15] J. R. Marden, G. Arslan, and J. S. Shamma. Connections between cooperative control and potential games. *IEEE Transactions on Systems, Man and Cybernetics, Part B: Cybernetics*, 2008. submitted.

[16] J. R. Marden and A. Wierman. Distributed welfare games. *Operations Research*, 2008. submitted.

[17] J.R. Marden and J.S. Shamma. Revisiting log-linear learning : asynchrony, completeness and payoff-based omplementation. *Games and Economic Behavior*, 2008. submitted.

[18] J.R. Marden, H.P. Young, G. Arslan, and J.S. Shamma. Payoff based dynamics for multi-player weakly acyclic games. *SIAM Journal on Control and Optimization*, 2006. submitted.

[19] C.B. Margi, V. Petkov, K. Obraczka, and R. Manduchi. Characterizing energy consumption in a visual sensor network testbed. In *Proc. of 2nd International Conference on Testbeds and Research Infrastructures for the Development of Networks and Communities*, pages 332–339, March 2006.

[20] D. Monderer and L. Shapley. Potential games. *Games and Economic Behavior*, 14:124–143, 1996.

[21] R. A. Murphey. Target-based weapon target assignment problems. In P. M. Pardalos and L. S. Pitsoulis, editors, *Nonlinear Assignment Problems: Algorithms and Applications*, pages 39–53. Kluwer Academic Publishers, 1999.

[22] J. O'Rourke. *Art Gallery Theorems and Algorithms*. Oxford University Press, 1987.

[23] R.W. Rosenthal. A class of games possseeing pure strategy nash equilibria. *International Journal of Game Theory*, 2:65–67, 1973.

[24] T. Roughgarden. *Selfish routing and the price of anarchy*. MIT press, 2005.

[25] T. C. Shermer. Recent results in art galleries. *Proceedings of the IEEE*, 80(9):1384–1399, 1992.

[26] S. Soro and W. B. Heinzelman. On the coverage problem in video based wireless sensor networks. In *Proc. of 2nd International Conference on Broadband Networks*, pages 932–939, 2005.

[27] A. Tang and L. Andrew. Game theory for heterogeneous flow control. In *Proc. of 42nd Annual Conference on Information Sciences and Systems*, pages 52–56, December 2008.

[28] J. Urrutia. Art gallery and illumination problems. In J. R. Sack and J. Urrutia, editors, *Handbook of Computational Geometry*, pages 973–1027. North-Holland, 2000.

[29] H.P. Young. The evolution of conventions. *Econometrica*, 61:57–84, Juanary 1993.

[30] M. Zhu and S. Martínez. Distributed coverage games for mobile visual sensors (ii): Reaching the set of global optima. In *Proc. of the 48th IEEE Conf. on Decision and Control and 28th Chinese Control Conference*, Shanghai, China, December 2009. To appear.