

Communication-Free Distributed Coverage for Networked Systems

A. Yasin Yazıcıoğlu, Magnus Egerstedt, and Jeff S. Shamma

Abstract—In this paper, we present a communication-free algorithm for distributed coverage of an arbitrary network by a group of mobile agents with local sensing capabilities. The network is represented as a graph, and the agents are arbitrarily deployed on some nodes of the graph. Any node of the graph is covered if it is within the sensing range of at least one agent. The agents are mobile devices that aim to explore the graph and to optimize their locations in a decentralized fashion by relying only on their sensory inputs. We formulate this problem in a game theoretic setting and propose a communication-free learning algorithm for maximizing the coverage.

I. INTRODUCTION

In many networked systems, a typical task is to provide some service such as security or maintenance via some agents with limited capabilities (e.g., [1], [2], [3]). One way of achieving this task is to solve a locational optimization problem (e.g., [4], [5], [6], [7], [8]) and let each agent serve some part of the network around its assigned position. In the absence of a centralized mechanism, the agents are faced with a distributed coverage control problem, where their objective is to optimize their locations by following some decentralized controllers.

Distributed coverage control is widely studied on continuous domains (e.g., [9]–[10]). One possible approach is to employ potential fields to drive each agent away from the nearby agents and obstacles (e.g., [9], [11]). Alternatively, a prevailing approach introduced in [12] is to model the underlying locational optimization problem as a continuous p -median problem and to employ Lloyd’s algorithm [13]. As such, the agents are driven onto a local optimum, i.e. a centroidal Voronoi configuration, where each point in the space is assigned to the nearest agent, and each agent is located at the center of mass of its own region. Later on, this method was extended for agents with distance-limited sensing and communications (e.g., [14]) and limited power (e.g., [15]), as well as for heterogeneous agents covering non-convex regions (e.g., [16]). Also, the requirement of sensing density functions was relaxed by incorporating methods from adaptive control and learning (e.g., [10]).

In some studies, distributed coverage control was studied on discrete spaces represented as graphs (e.g., [17], [18], [19], [20]). One possible approach is to achieve a centroidal Voronoi partition of the graph via pairwise gossip algorithms (e.g., [17]) or via asynchronous greedy updates (e.g., [18]).

This work was supported by ONR project #N00014-09-1-0751.

A. Yasin Yazıcıoğlu is with the Laboratory for Information and Decision Systems, Massachusetts Institute of Technology, yasiny@mit.edu.

Magnus Egerstedt is with the School of Electrical and Computer Engineering, Georgia Institute of Technology, magnus@gatech.edu.

Jeff S. Shamma is with the School of Electrical and Computer Engineering, Georgia Institute of Technology, shamma@gatech.edu, and with King Abdullah University of Science and Technology (KAUST), jeff.shamma@kaust.edu.sa.

Alternatively, distributed coverage control on discrete spaces can be studied in a game theoretic framework (e.g., [19], [20]). Game theoretic methods have been used to solve many cooperative control problems such as vehicle-target assignment (e.g., [21]), dynamic vehicle routing (e.g. [22]), cooperative communication (e.g., [23]), and coverage optimization (e.g., [19], [20]). In [19], sensors with variable footprints achieve power-aware optimal coverage on a discretized space. In [20], a group of heterogeneous mobile agents are driven on a graph to maximize the number of covered nodes.

In this paper, we study a distributed coverage control problem on graphs in a game theoretic setting. In this problem, mobile agents are arbitrarily deployed on an unknown graph. Each agent is assumed to sense the local graph structure and the presence of other agents (if any) within its sensing range. Any node of the graph is covered if it is within the sensing range of at least one agent. The objective of the agents is to maximize the number of covered nodes by optimizing their locations on the graph. We present a game theoretic formulation for this coverage control problem. We particularly focus on a communication-free setting, where each agent should be driven via only its sensory inputs. In that case, the agents do not observe their exact utilities in the corresponding game. Accordingly, we propose a learning algorithm for driving the agent positions based on some estimated utilities. Using the proposed method, the agents maintain optimal coverage with an arbitrarily high probability as time goes to infinity.

The organization of this paper is as follows: Section II presents the distributed graph coverage problem. Section III sets up the game-theoretic formulation of the problem. Section IV presents a solution that requires some explicit communications among the agents. The proposed communication-free solution is presented in Section V. Some simulation results for the proposed method are presented in Section VI. Finally, Section VII concludes the paper.

II. DISTRIBUTED GRAPH COVERAGE

In this section, we present the distributed graph coverage (DGC) problem, where the goal is to maximize the number of covered nodes by driving the agents with limited sensing and mobility capabilities to optimal locations on a graph. First, some graph theory preliminaries are presented.

A. Graph Theory Concepts

An undirected graph $\mathcal{G} = (V, E)$ consists of a *node set* V and an *edge set* $E \subseteq V \times V$. For an undirected graph, the edge set consists of unordered node pairs (v, v') denoting that the nodes v and v' are adjacent.

A *path* is a sequence of nodes such that each node is adjacent to the preceding node in the sequence. For any two

nodes v and v' , the *distance* between the nodes $d(v, v')$ is the number of edges in a shortest path between v and v' . A graph is *connected* if the distance between any pair of nodes is finite.

The set of nodes containing a node v and all the nodes adjacent to v is called the (*closed*) *neighborhood* of v , and it is denoted as \mathcal{N}_v . For any $\delta \geq 0$, the δ -neighborhood of v , \mathcal{N}_v^δ , is the set of nodes that are at most δ away from v , i.e.

$$\mathcal{N}_v^\delta = \{v' \in V \mid d(v, v') \leq \delta\}. \quad (1)$$

For any $\mathcal{G} = (V, E)$, an *induced subgraph*, $\mathcal{G}[V_s]$, consists of the vertices, $V_s \subseteq V$, and the edges whose endpoints are both in V_s .

B. Problem Formulation

Consider a connected undirected graph, $\mathcal{G} = (V, E)$, and let $I = \{1, 2, \dots, m\}$ denote a set of m mobile agents arbitrarily deployed on some nodes of the graph. Let each agent have a sensing range, δ . We assume that each agent, i , can sense the subgraph induced by the nodes in $\mathcal{N}_{v_i}^\delta$ and the presence of other agents (if any) within its δ -neighborhood. As such, each $i \in I$ located at $v_i \in V$ covers all the nodes in $\mathcal{N}_{v_i}^\delta$. Any node of the graph is covered if it is included in the δ -neighborhood of at least one agent, and the set of covered nodes, $V_c \subseteq V$, is given as

$$V_c(v_1, \dots, v_m) = \bigcup_{i=1}^m \mathcal{N}_{v_i}^\delta. \quad (2)$$

The objective in the distributed graph coverage (DGC) problem is to have the agents update their positions over time in a distributed manner to maximize the number of covered nodes, i.e.

$$|V_c(v_1(t), \dots, v_m(t))|, \quad (3)$$

where each $v_i(t) \in V$ is the position of agent i at time t .

In order to achieve optimal coverage in a distributed fashion, the agents need some local rules to follow. In general, a rule is considered to be local if its execution by an agent requires only some information available within a small distance from the agent. In this paper, we consider a discrete time dynamics and we assume that each agent can either maintain its position or move to an adjacent node in the next time step, i.e.

$$d(v_i(t+1), v_i(t)) \leq 1, \quad \forall i \in \{1, 2, \dots, m\}. \quad (4)$$

In this setting, multiple agents can be located on the same node, and the only constraint on the agent movements is the locality constraint in (4).

C. Solution Approach

In the DGC problem, a group of mobile agents explore an unknown graph and aim to cover as many nodes as possible. As such, the underlying locational optimization problem is similar to the maximum coverage problem (e.g., [5], [6]). Such NP-hard problems are typically tackled by finding sufficiently good approximate solutions through fast algorithms (e.g. [24], [25], [26]). Similarly, in many distributed coverage control studies, a locational objective function is optimized by the agents aiming for the best local improvements (e.g., [12]-[18]).

Such a distributed greedy approach can be employed to solve the DGC problem. Accordingly, the agents may move locally on the graph to maximally improve their local coverage. In that case, the resulting performance would significantly depend on the graph structure and the initial configuration. This method may rapidly lead to a reasonable approximate solution if the agents start with a sufficiently good initial coverage or if the interaction graph satisfies some structural properties. However, it may also lead to arbitrarily poor configurations for arbitrary graphs and initial conditions. For instance, consider the scenario in Fig. 1, where 2 agents with sensing ranges $\delta = 1$ can achieve a globally optimal configuration in 2 time steps. In this example, the initial configuration would be stationary under a greedy approach since none of the agents can improve the coverage by moving to a neighboring node. Note that the performance in Fig. 1a would be arbitrarily poor for any arbitrarily large graph obtained by adding more leaf nodes attached to the unoccupied hub.

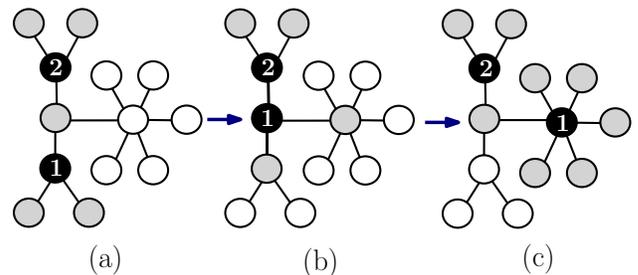


Fig. 1. A possible trajectory to a globally optimal configuration for two agents on a small graph. The agents have cover ranges $\delta = 1$, and they are initially located as in (a). The number of covered nodes (shown as non-white) is reduced in the intermediate step illustrated in (b) to reach the global optima shown in (c).

In order to ensure efficient coverage for arbitrary graphs and initial configurations, a solution method should occasionally allow for graph exploration at the expense of a better coverage. In this work, we present such a solution by approaching the problem from a game theoretic perspective. In particular, we map the DGC problem to a game, and we design a learning algorithm for the agents to follow in updating their actions.

III. GAME THEORETIC FORMULATION

In this section, a game-theoretic formulation of the DGC problem is presented. First, some game theory preliminaries are provided.

A. Game Theory Concepts

A finite *strategic game* $\Gamma = (I, A, U)$ consists of three components: (1) a set of m *players (agents)* $I = \{1, 2, \dots, m\}$, (2) an m -dimensional *action space* $A = A_1 \times A_2 \times \dots \times A_m$, where each A_i is the *action set* of player i , and (3) a set of *utility functions* $U = \{U_1, U_2, \dots, U_m\}$, where each $U_i : A \mapsto \mathfrak{R}$ is a mapping from the action space to real numbers.

For any *action profile* $a \in A$, let a_{-i} denote the actions of players other than i . Using this notation, an action profile a

can also be represented as $a = (a_i, a_{-i})$. An action profile $a^* \in A$ is called a *Nash equilibrium* if

$$U_i(a_i^*, a_{-i}^*) = \max_{a_i \in A_i} U_i(a_i, a_{-i}^*), \quad \forall i \in I. \quad (5)$$

A class of games that is widely utilized in cooperative control problems is the potential games. A game is called a potential game if there exists a *potential function*, $\phi : A \mapsto \mathbb{R}$, such that the change of a player's utility resulting from its unilateral deviation from an action profile equals the resulting change in ϕ . More precisely, for each player i , for every $a_i, a'_i \in A_i$, and for all $a_{-i} \in A_{-i}$,

$$U_i(a'_i, a_{-i}) - U_i(a_i, a_{-i}) = \phi(a'_i, a_{-i}) - \phi(a_i, a_{-i}). \quad (6)$$

When a cooperative control problem is mapped to a potential game, the game is designed such that its potential function captures the global objective (e.g., [27]).

B. DGC Game

In order to formulate the DGC problem in a game theoretic setting, we design a corresponding game, Γ_{DGC} , by defining the action space and the utility functions. More specifically, we design a potential game such that its potential function, $\phi(a)$, captures the global objective of the DGC problem, i.e.

$$\phi(a) = |V_c(a)|. \quad (7)$$

In the DGC problem, the coverage provided by each agent is determined by the position of the agent. Hence, the action of an agent can be defined as its position on the graph. Accordingly, each action set is equal to the node set of $\mathcal{G} = (V, E)$, i.e.

$$A_i = V, \quad \forall i \in I. \quad (8)$$

Then, the utilities should be designed such that $\phi(a)$ in (7) is indeed the potential function for the resulting game. To this end, we design the agent utilities as

$$\begin{aligned} U_i(a) &= |\mathcal{N}_{a_i}^\delta \setminus \bigcup_{j \neq i} \mathcal{N}_{a_j}^\delta|, \\ &= \sum_{v \in \mathcal{N}_{a_i}^\delta} u_i(v, a_{-i}), \end{aligned} \quad (9)$$

where, for every $v \in \mathcal{N}_{a_i}^\delta$, $u_i(v, a_{-i})$ is the partial utility agent i gathers by covering node v , and it is defined as

$$u_i(v, a_{-i}) = \begin{cases} 1 & \text{if } d(v, a_j) > \delta \quad \forall j \neq i, \\ 0 & \text{otherwise.} \end{cases} \quad (10)$$

In the resulting game, each agent gathers a utility equal to the number of nodes that are covered only by itself. Note that this utility is equal to the marginal contribution of the corresponding agent to the number of covered nodes.

Lemma 3.1. *The utilities in (9) lead to a potential game $\Gamma_{\text{DGC}} = (P, A, U)$ with the potential function given in (7).*

Proof. Let $a_i = v_i$ and $a'_i = v'_i$ be two possible actions for agent i , and let a_{-i} denote the actions of other agents. Due to (2) and (7),

$$\phi(a) = \left| \bigcup_{i \in I} \mathcal{N}_{a_i}^\delta \right| \quad (11)$$

Using (9), for any agent i , (11) can be expanded as

$$\phi(a) = |\mathcal{N}_{a_i}^\delta \setminus \bigcup_{j \neq i} \mathcal{N}_{a_j}^\delta| + \left| \bigcup_{j \neq i} \mathcal{N}_{a_j}^\delta \right| = U_i(a_i, a_{-i}) + \left| \bigcup_{j \neq i} \mathcal{N}_{a_j}^\delta \right|. \quad (12)$$

Using (12) for any pair of actions a_i and a'_i ,

$$\phi(a'_i, a_{-i}) - \phi(a_i, a_{-i}) = U_i(a'_i, a_{-i}) - U_i(a_i, a_{-i}). \quad (13)$$

□

C. Learning

In game theoretic learning, starting from an arbitrary initial configuration, the agents repetitively play a game. At each step $t \in \{0, 1, 2, \dots\}$, each agent $i \in I$ plays an action $a_i(t)$ and receives some utility $U_i(a(t))$. In this setting, the agents update their actions in accordance with some learning algorithms. For the DGC problem, the learning process is desired to drive the agent positions to the set of configurations that maximize the number of covered nodes.

For potential games, there are many learning algorithms that provide convergence to a Nash equilibrium (e.g., [21], [27] and the references therein). While any potential maximizer is a Nash equilibrium, a potential game may also have some suboptimal Nash equilibria. In that case, a learning algorithm such as *log-linear learning* (LLL) can be used to drive the agents to the configurations that maximize the potential function $\phi(a)$ [28]. Essentially, LLL is a noisy best-response algorithm, and it induces a Markov chain over the action space with a unique limiting distribution, μ_ϵ^* , where ϵ denotes the noise parameter. As the noise parameter, ϵ , goes down to zero, the limiting distribution, μ_ϵ^* , has an arbitrarily large part of its mass over the set of potential maximizers [28]. However, LLL assumes that at any round each player i has access to all the actions in its action set A_i . In general, LLL may not provide potential maximization when the system evolves over constrained action sets, i.e. when each agent i is allowed to choose its next action from only a subset of actions. Note that this is indeed the case for the DGC problem, and each agent has to pick its the next action from the closed neighborhood of its current action a_i ,

$$A_i^c(a_i) = \mathcal{N}_{a_i} \quad \forall i \in I. \quad (14)$$

The issue of constrained action sets was addressed in [29], and a variant learning algorithm called *binary log-linear learning* (BLLL) was presented for such cases.

In learning algorithms, typically each agent is assumed to measure its current utility. For instance, in order to execute LLL or BLLL, the agents need to measure their utilities resulting from their current actions as well as the hypothetical utilities they may gather by unilaterally switching to some other actions. Alternatively, a payoff-based implementation may be utilized to avoid the necessity to compute the hypothetical utilities [29]. Note that, for Γ_{DGC} , even the computation of the current utility requires some *explicit communications* since the agents with overlapping coverage are not necessarily within the sensing range of each other. In general, such agents can be up to 2δ apart on the graph.

D. Stochastic Stability Concepts

Stochastic stability was introduced in game theory to explain how some equilibria may arise much more frequently than the others in the long run as the strategic decision makers update their actions in a stochastic setting with noise [28], [30], [31]. The concept of stochastic stability will be extensively used in the remainder of this paper. Hence, we provide some preliminaries prior to our derivations, and further details about the subject can be found in [28], [30], [31] and the references therein.

Definition (Regular Perturbed Markov Chain): Let P be the transition matrix of a discrete-time Markov chain over a finite state space \mathcal{X} . A perturbed Markov chain with the noise parameter ϵ is called a regular perturbed Markov chain if

- 1) P_ϵ is aperiodic and irreducible for $\epsilon > 0$,
- 2) $\lim_{\epsilon \rightarrow 0} P_\epsilon = P$,
- 3) For any $x, x^+ \in \mathcal{X}$ if $P_\epsilon(x, x^+) > 0$, then there exists $R(x, x^+) \geq 0$ such that

$$0 < \lim_{\epsilon \rightarrow 0^+} \frac{P_\epsilon(x, x^+)}{\epsilon^{R(x, x^+)}} < \infty, \quad (15)$$

where $R(x, x^+)$ is called the resistance of the transition from x to x^+ .

Any regular perturbed Markov chain, P_ϵ , has a unique limiting distribution, μ_ϵ^* , since it is aperiodic and irreducible. In a stochastic stability analysis, the objective is to investigate the behavior of μ_ϵ^* as the noise parameter ϵ goes down to zero. In a nutshell, a state is stochastically stable if the corresponding entry of the limiting distribution does not approach zero as the noise in the process becomes arbitrarily small.

Definition (Stochastically Stable State): Let P_ϵ denote a regular perturbed Markov chain over a state space, \mathcal{X} . Any state, $x \in \mathcal{X}$, is stochastically stable if

$$\lim_{\epsilon \rightarrow 0^+} \mu_\epsilon^*(x) > 0. \quad (16)$$

The stochastically stable states of any regular perturbed Markov chain, P_ϵ , can be characterized through a resistance tree analysis. For any state $x \in \mathcal{X}$ of a regular perturbed Markov chain, a spanning tree rooted at x can be constructed as a directed graph \mathcal{T}_x , where the nodes correspond to states, the edges correspond to feasible state transitions, and there is a unique directed path on \mathcal{T}_x from any state $x' \neq x$ to x . The resistance of such a tree, $R(\mathcal{T}_x)$, is defined as the sum of the resistances of its edges, where the resistance of each edge is given as in (15). A spanning tree \mathcal{T}_x^* is called a minimum resistance tree if $R(\mathcal{T}_x^*) \leq R(\mathcal{T}_x)$ for any \mathcal{T}_x , i.e. any spanning tree rooted at x has at least as much resistance as \mathcal{T}_x^* . The stochastic potential of a state, x , is defined as the total resistance of its minimum resistance tree, $R(\mathcal{T}_x^*)$. The following result characterizes the stochastically stable states through their stochastic potentials.

Lemma 3.2. [30] *Let P_ϵ be a regular perturbed Markov chain. A state $x \in \mathcal{X}$ is stochastically stable if and only if x is a recurrent state of the unperturbed chain, P_0 , with the minimum stochastic potential.*

IV. COVERAGE MAXIMIZATION

In this section, we will briefly show that if all the agents follow BLLL in a repetitive play of Γ_{DGC} , then the stochastically stable states are the coverage maximizers. A more detailed presentation of this approach can be found in [20]. As stated earlier, this solution requires some local communications among the agents. In the next section, we will present a novel learning algorithm that can achieve the same limiting behavior as this method without any explicit communications.

Algorithm I: Binary Log-linear Learning ([29])

- 1 : **initialization:** $\epsilon \in \mathbb{R}^+$ small, $a \in A$ arbitrary
- 2 : **repeat**
- 3 : Pick a random $i \in I$, and a random $a'_i \in A_i^c(a_i)$.
- 4 : Compute $\alpha = \epsilon^{-U_i(a(t))}$, $\beta = \epsilon^{-U_i(a'_i, a_{-i}(t))}$.
- 5 : With probability $\frac{\beta}{\alpha + \beta}$, set $a_i = a'_i$.
- 6 : **end repeat**

In BLLL, a single agent is randomly chosen at each time step. The selection of a single agent at each time step can be achieved (with a very high probability) without a centralized coordination by using methods such as the asynchronous time model proposed in [32]. The selected agent, assuming that all the other agents are stationary, updates its action depending on its current utility and the hypothetical utility it would receive by playing a random action in its constrained action set. This is illustrated in Fig. 2.

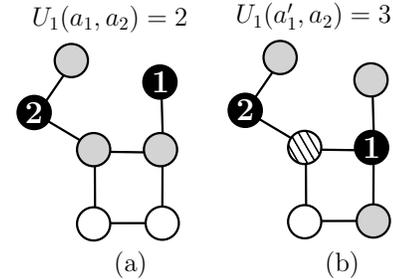


Fig. 2. An illustration of the BLLL algorithm, where two agents with $\delta = 1$ are located as in (a) and Agent 1 is updating its action. Agent 1 randomly picks a candidate action, $a'_1 \in A_i^c(a_i)$, as in (b). Its next action is picked from $\{a_1, a'_1\}$ with probabilities depending on the corresponding utilities. For the configuration in (b), the tiled node is not providing any utility to either of the agents since it is covered by both of them.

In [29], it was shown that BLLL can be used to achieve potential maximization if the constrained action sets satisfy Properties 1 and 2 provided below.

Property 1 (Reachability) For any agent $i \in I$ and any action pair $a_i^0, a_i^k \in A_i$, there exists a sequence of actions $\{a_i^0, a_i^1, \dots, a_i^k\}$ such that $a_i^r \in A_i^c(a_i^{r-1})$ for all $r \in \{1, 2, \dots, k\}$.

Property 2 (Reversability) For any agent $i \in I$ and any action pair $a_i, a'_i \in A_i$,

$$a'_i \in A_i^c(a_i) \Leftrightarrow a_i \in A_i^c(a'_i).$$

Theorem 4.1. [29] Consider any finite potential game and constrained action sets satisfying Properties 1 and 2. If all players adhere to BLLL, then the stochastically stable states are the set of potential maximizers.

In light of Theorem 4.1, the agents can maximize the coverage by following the BLLL algorithm in a repetitive play of Γ_{DGC} , if the constrained action sets given in (14) satisfy Properties 1 and 2. Lemma 4.2 shows that the constrained action sets indeed satisfy these properties if the graph to be covered is connected.

Lemma 4.2. The constrained action sets in (14) satisfy Properties 1 and 2 if the graph $\mathcal{G} = (V, E)$ is connected.

Proof. If the graph is connected, then there exists a finite-length path $\{v^0, \dots, v^k\}$ between any pair of nodes $v^0, v^k \in V$, and Property 1 is satisfied. Furthermore, for undirected graphs, $d(v, v') = d(v', v)$. Hence, Property 2 is also satisfied. \square

Theorem 4.3. Let $\mathcal{G} = (V, E)$ be connected graph, and let all agents follow BLLL in a repetitive play of Γ_{DGC} with the constrained action sets in (14). Then the stochastically stable states are the maximizers of $|V_c(a)|$.

Proof. If $\mathcal{G} = (V, E)$ is connected, then the constrained action sets in (14) satisfy Properties 1 and 2 due to Lemma 4.2. Hence, in light of Theorem 4.1, if all agents follow BLLL in a repetitive play of Γ_{DGC} , the stochastically stable states are the potential maximizers. Due to (7), those are the configurations maximizing the number of covered nodes, $|V_c(a)|$. \square

V. COMMUNICATION-FREE COVERAGE MAXIMIZATION

In the DGC problem, the sensory inputs do not reveal which of the nodes within the sensing range of an agent is covered also by some other agents. However, each agent can sense if any other agent is also covering its current position as illustrated in Fig. 3. Hence, each agent i observes the partial utility, $u_i(a_i, a_{-i})$ in (10), via its sensory input.

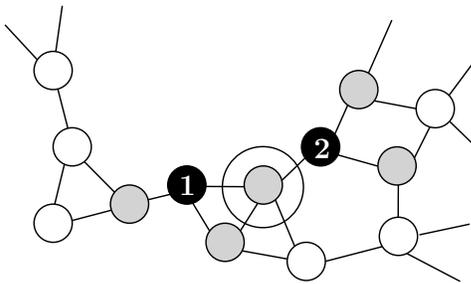


Fig. 3. Distributed graph coverage by agents with sensing ranges $\delta = 1$. Agents 1 and 2 do not observe that the encircled node is covered by both of them. However, each of them knows that its current position is covered only by itself since no other agent is within its sensing range.

Since the exact utilities in Γ_{DGC} are not measurable without explicit communications, the agents need to update their actions based on some estimated utilities. Assuming that the nearby agents will remain stationary for a sufficient amount of time, each agent i can construct an estimated utility by

visiting each $v \in \mathcal{N}_{a_i}^\delta$ and combining the sampled $u_i(v, a_{-i})$. Note that the resulting estimation will not necessarily be equal to the actual utility since multiple agents may be moving simultaneously as illustrated in Fig. 4. However, if the probability of having simultaneously moving agents is sufficiently small, then false estimations will be sufficiently rare for the agents to achieve the desired limiting behavior by following a noisy best-response based on the estimated utilities. The proposed communication-free algorithm is based on this approach, and it achieves the desired limiting behavior by making the erroneous utility estimations sufficiently rare without requiring any scheduling or coordination among the agents.

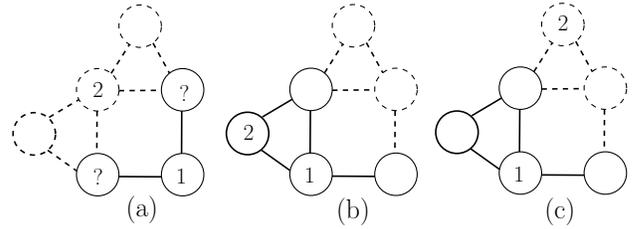


Fig. 4. Two agents with sensing ranges $\delta = 1$ are located on a graph as in (a). Part of the graph that is not sensed by agent 1 is dashed in the figures. Agent 1 can estimate its utility in (a) by sampling the partial utilities from the nodes in its sensing range. If agent 2 is stationary in the meantime, then the resulting estimation will be true. However, if agent 2 is also moving, then the sampled partial utilities may be true as in (b) or false as in (c).

In the remainder of this section, we present the proposed communication-free coverage maximization algorithm (CFCM) and an analysis of the corresponding dynamics.

A. CFCM Algorithm

The proposed algorithm has two parameters to be set. The first parameter, $\epsilon \in \mathbb{R}^+$, is the noise in the agent decisions when choosing between the candidate actions based on the corresponding estimated utilities. The second parameter, $r \in \mathbb{R}^+$, sets the likelihood of each agent to update its action. As it will be shown later in this section, the desired global behavior emerges when r is sufficiently large and ϵ is small.

In CFCM, each agent i is either stationary or experimenting. Each stationary agent repeats its current action in the next time step with a high probability, $1 - \epsilon^r$, or starts an experiment with probability ϵ^r . An experiment involves comparing its current action, a_i^1 , to an alternative randomly picked from its constrained action set, $a_i^2 \in A_i^c(a_i^1)$, where $A_i^c(a_i^1)$ is the local neighborhood of a_i^1 as given in (14). In this aspect, the agent behavior is similar to the payoff-based BLLL in [29]. However, since the agents receive only some partial utilities, $u_i(a_i, a_{-i})$, an experiment consists of visiting all the nodes in $\mathcal{N}_{a_i^1}^\delta \cup \mathcal{N}_{a_i^2}^\delta$ to see which of those nodes are also covered by some other agents. We refer to the corresponding path to be traversed as an experiment path between a_i^1 and a_i^2 .

Definition (Experiment Path): Let δ be the sensing range of the agents. For any a_i^1 and $a_i^2 \in A_i^c(a_i^1)$, a finite path, $\{a_i^1, \dots, a_i^2\}$, is an experiment path if it traverses $\mathcal{N}_{a_i^1}^\delta \cup \mathcal{N}_{a_i^2}^\delta$.

For any a_i^1 and $a_i^2 \in A_i^c(a_i^1)$, an experiment path can be obtained locally by utilizing methods such as depth-first search or breadth-first search (e.g., [33]). In the CFCM algorithm, an experiment path between a_i^1 and a_i^2 is denoted as $\mathcal{E}(a_i^1, a_i^2)$. During an experiment, the agent traverses its experiment path to construct the estimated utilities, \hat{U}_i^1 and \hat{U}_i^2 , from the sampled partial utilities. For simplicity, a partial utility from a node is sampled only at the last visit to that node during the experiment. As such, if it is the agent's last visit of the current position, a_i , and the agent does not sense any other agent within δ , then the utility estimations corresponding to the candidate actions within δ from a_i are incremented by 1. Once the experiment path is traversed, the agent randomly chooses between the two candidate actions based on the estimated utilities, \hat{U}_i^1 and \hat{U}_i^2 . At the next time step, the agent becomes stationary at its chosen action until it starts a new experiment.

Algorithm II: Communication-free Coverage Maximization (CFCM)

```

1 : initialization:  $\epsilon \in \mathfrak{R}^+$  (small),  $r \in \mathfrak{R}^+$ ,  $a_i \in A_i$  arbitrary,
    $S_i = \{a_i\}$ ,  $k_i = 1$ ,  $\hat{U}_i^1 = \hat{U}_i^2 = 0$ .
2 : repeat
3 :    $a_i = S_i(k_i)$ ,  $a_i^1 = S_i(1)$ ,  $a_i^2 = S_i(|S_i|)$ .
4 :   if ( $|S_i| = 1$ )
5 :     Generate a random (uniform)  $\gamma \in [0, 1]$ .
6 :     if ( $\gamma \leq \epsilon^r$ )
7 :        $a_i^2$  is randomly (uniform) chosen over  $A_i^c(a_i^1)$ .
8 :        $S_i = \mathcal{E}(a_i^1, a_i^2)$ .
9 :     end if
10 :  else
11 :    if ( $k_i \geq k$ ,  $\forall k \in \{k \mid S_i(k) = a_i\}$ )
12 :       $\hat{U}_i^1 = \hat{U}_i^1 + u_i(a_i, a_{-i})$ , if  $a_i \in \mathcal{N}_{a_i^1}^\delta$ .
13 :       $\hat{U}_i^2 = \hat{U}_i^2 + u_i(a_i, a_{-i})$ , if  $a_i \in \mathcal{N}_{a_i^2}^\delta$ .
14 :    end if
15 :    if ( $k_i = |S_i|$ )
16 :       $\alpha = \epsilon^{-\hat{U}_i^1}$ ,  $\beta = \epsilon^{-\hat{U}_i^2}$ .
17 :       $S_i = \begin{cases} \{a_i^1\} & \text{w.p. } \frac{\alpha}{\alpha+\beta}, \\ \{a_i^2\} & \text{otherwise.} \end{cases}$ 
18 :       $k_i = 1$ ,  $\hat{U}_i^1 = \hat{U}_i^2 = 0$ .
19 :    else
20 :       $k_i = k_i + 1$ .
21 :    end if
22 :  end if
23 : end repeat

```

For the CFCM algorithm, the state of any agent i can be defined as

$$x_i = [S_i \quad k_i \quad \hat{U}_i^1 \quad \hat{U}_i^2], \quad (17)$$

where S_i is a sequence of actions, which is either a singleton (stationary) or an experiment path (experimenting), $k_i \in \{1, \dots, |S_i|\}$ is an index variable denoting which action in S_i is currently taken by the agent, and \hat{U}_i^1, \hat{U}_i^2 are the estimations for $U_i(a_i^1, a_{-i})$ and $U_i(a_i^2, a_{-i})$, respectively. In this representation, the current action, a_i , and the candidate actions, a_i^1 and a_i^2 , are given as

$$a_i = S_i(k_i), \quad a_i^1 = S_i(1), \quad a_i^2 = S_i(|S_i|), \quad (18)$$

where $S_i(k_i)$ denotes the k_i^{th} element in S_i , and $|S_i|$ denotes the length of S_i . The CFCM algorithm is memoryless since the state of every agent in the next time step is independent of its past trajectory. As such, if all agents follow the CFCM algorithm, then a Markov chain is induced over the state space, \mathcal{X} , where each $x \in \mathcal{X}$ is the global state obtained by concatenating the states of all agents, i.e.

$$x = [x_1, x_2, \dots, x_m]. \quad (19)$$

In the remainder of this section, the limiting behavior of the resulting Markov chain will be inspected through a stochastic stability analysis.

B. Limiting Behavior

For any $x \in \mathcal{X}$, the agents can be grouped into two distinct sets consisting of the stationary agents, $I_s(x)$, and the experimenting agents, $I_e(x)$, as

$$I_s(x) = \{i \in I \mid |S_i| = 1\}, \quad (20)$$

$$I_e(x) = I \setminus I_s(x). \quad (21)$$

Using these sets, for any feasible transition, $x \rightarrow x^+$, the agents can be grouped into 4 disjoint sets based on the transition of their individual states:

$$I_{ss}(x, x^+) = I_s(x) \cap I_s(x^+), \quad (22)$$

$$I_{se}(x, x^+) = I_s(x) \cap I_e(x^+), \quad (23)$$

$$I_{ee}(x, x^+) = I_e(x) \cap I_e(x^+), \quad (24)$$

$$I_{es}(x, x^+) = I_e(x) \cap I_s(x^+), \quad (25)$$

where $I_{ss}(x, x^+)$ are the agents that remain stationary, $I_{se}(x, x^+)$ are the ones starting to experiment, $I_e(x)$ are the experimenting agents that have not completed moving along their experiment paths, and $I_{es}(x, x^+)$ are the agents that have completed traversing their experiment paths and choose between their candidate actions.

The agents in $I_{es}(x, x^+)$ can be further partitioned as the ones choosing their first candidate action and the ones that choose their second candidate action, i.e.

$$I_{es}^1(x, x^+) = \{i \in I_{es}(x, x^+) \mid a_i^+ = a_i^1\}, \quad (26)$$

$$I_{es}^2(x, x^+) = \{i \in I_{es}(x, x^+) \mid a_i^+ = a_i^2\}. \quad (27)$$

Note that the agents in $I_{es}(x, x^+)$ do not necessarily choose the action resulting in the higher estimated utility. For each $i \in I$, let $\hat{U}_i^* = \max\{\hat{U}_i^1, \hat{U}_i^2\}$. Then, the amount of estimated utility that is denied in the transition $x \rightarrow x^+$ is given as

$$\Delta_i(x_i, x_i^+) = \begin{cases} \hat{U}_i^* - \hat{U}_i^1 & \text{if } i \in I_{es}^1(x, x^+), \\ \hat{U}_i^* - \hat{U}_i^2 & \text{if } i \in I_{es}^2(x, x^+), \\ 0 & \text{otherwise.} \end{cases} \quad (28)$$

Next, we show that the CFCM algorithm induces a regular perturbed Markov chain, where the resistance of any feasible transition depends on the estimated utilities denied by the agents becoming stationary and the number of agents starting new experiments.

Lemma 5.1. *Let $\mathcal{G} = (V, E)$ be connected graph. If all agents employ the CFCM algorithm, then a regular perturbed Markov chain is induced over \mathcal{X} , and the resistance of any feasible transition, $x \rightarrow x^+$, is*

$$R(x, x^+) = r|I_{se}(x, x^+)| + \sum_{i \in I_{es}(x, x^+)} \Delta_i(x_i, x_i^+). \quad (29)$$

Proof. Let P_ϵ denote the transition matrix of the Markov chain induced by the CFCM algorithm. For $\epsilon > 0$, any all-stationary state can be reached from any other all-stationary state through a sequence of experiments, given $\mathcal{G} = (V, E)$ is connected. Furthermore, any state that is not all-stationary lies on a feasible path between two all-stationary states. Hence, P_ϵ is irreducible. Furthermore, since the stationary agents remain stationary with probability $1 - \epsilon^r$, aperiodicity immediately follows from the resulting self-loops at all-stationary states.

The probability any feasible transition from x to x^+ , given in P_ϵ , is the joint probability of state transitions of individual agents. Note that for any agent, $i \in I_{ee}(x)$, the transition from x_i to x_i^+ does not have any randomness. Hence, the probability of transition from x to x^+ is

$$P_\epsilon(x, x^+) = \Pr[I_{es}^1(x, x^+)] \Pr[I_{es}^2(x, x^+)] \Pr[I_{ss}(x, x^+)] \Pr[I_{se}(x, x^+)], \quad (30)$$

where each term on the right side of (30) denote the joint probability of state transitions for the agents in the corresponding subset, and they are given as

$$\Pr[I_{es}^1(x, x^+)] = \prod_{i \in I_{es}^1(x, x^+)} \frac{\epsilon^{-\hat{U}_i^1}}{\epsilon^{-\hat{U}_i^1} + \epsilon^{-\hat{U}_i^2}}, \quad (31)$$

$$\Pr[I_{es}^2(x, x^+)] = \prod_{i \in I_{es}^2(x, x^+)} \frac{\epsilon^{-\hat{U}_i^2}}{\epsilon^{-\hat{U}_i^1} + \epsilon^{-\hat{U}_i^2}}, \quad (32)$$

$$\Pr[I_{ss}(x, x^+)] = \prod_{i \in I_{ss}(x, x^+)} (1 - \epsilon^r), \quad (33)$$

$$\Pr[I_{se}(x, x^+)] = \prod_{i \in I_{se}(x, x^+)} \frac{\epsilon^r}{|A_i^c(a_i^1)|} \Pr[S_i^+; a_i^1, a_i^2], \quad (34)$$

where $\Pr[S_i^+; a_i^1, a_i^2]$ is the probability of having S_i^+ as the experiment path for an agent comparing a_i^1 and a_i^2 . $\Pr[S_i^+; a_i^1, a_i^2]$ depends on the function $\mathcal{E}(a_i^1, a_i^2)$, and it is independent of ϵ . Plugging (31)-(34) into (30), one can verify that the resistance $R(x, x^+)$ given in (29) satisfies

$$0 < \lim_{\epsilon \rightarrow 0^+} \frac{P_\epsilon(x, x^+)}{\epsilon^{R(x, x^+)}} < \infty. \quad (35)$$

□

Since the CFCM algorithm induces a regular perturbed Markov chain, the stochastically stable states are the recurrent states of the unperturbed chain with the minimum stochastic potential, as given in Lemma 3.2. Note that if $\epsilon = 0$, then no agent starts an experiment. In that case, the set of recurrent states, \mathcal{X}_R^0 , consists of the all-stationary states. All the other states, where at least one agent is experimenting, form the set of transient states, \mathcal{X}_T^0 , i.e.

$$\mathcal{X}_R^0 = \{x \mid I_s(x) = I\}, \quad (36)$$

$$\mathcal{X}_T^0 = \mathcal{X} \setminus \mathcal{X}_R^0. \quad (37)$$

The stochastic potentials of the states in \mathcal{X}_R^0 are determined by the resistances of the feasible transitions. Note that the parameter r in the CFCM algorithm has a direct influence on the resistances as given in (29). We will show that, for any connected graph \mathcal{G} , if r is sufficiently large, then the states in \mathcal{X}_R^0 with the minimum stochastic potential are the coverage maximizers. To provide a sufficient value of r , first we relate the structure of the graph to the maximum amount of estimated utility that can be denied by an agent in any feasible transition under the CFCM algorithm.

Lemma 5.2. *Let all agents follow the CFCM algorithm to cover a connected graph, $\mathcal{G} = (V, E)$, and let $\nu(\mathcal{G})$ be*

$$\nu(\mathcal{G}) = \max_{(v, v') \in E} |\mathcal{N}_v^\delta \setminus \mathcal{N}_{v'}^\delta|. \quad (38)$$

Then, for any feasible transition $x \rightarrow x^+$,

$$\nu(\mathcal{G}) \geq \max_{i \in I} \Delta_i(x_i, x_i^+). \quad (39)$$

Proof. Let $x \rightarrow x^+$ be a feasible transition. For any $i \in I_s(x)$, $\hat{U}_i^1 = \hat{U}_i^2 = 0$. On the other hand, for any $i \in I_e(x)$, the sampled partial utilities from the nodes $\mathcal{N}_{a_i^1}^\delta \cap \mathcal{N}_{a_i^2}^\delta$, contribute equally to both \hat{U}_i^1 and \hat{U}_i^2 . Hence,

$$\max\{|\mathcal{N}_{a_i^1}^\delta \setminus \mathcal{N}_{a_i^2}^\delta|, |\mathcal{N}_{a_i^2}^\delta \setminus \mathcal{N}_{a_i^1}^\delta|\} \geq |\hat{U}_i^1 - \hat{U}_i^2|, \quad \forall i \in I. \quad (40)$$

In light of (40) and (28),

$$\max\{|\mathcal{N}_{a_i^1}^\delta \setminus \mathcal{N}_{a_i^2}^\delta|, |\mathcal{N}_{a_i^2}^\delta \setminus \mathcal{N}_{a_i^1}^\delta|\} \geq \Delta_i(x_i, x_i^+), \quad \forall i \in I. \quad (41)$$

Since $(a_i^1, a_i^2) \in E$ for any $i \in I_e(x)$, (38) implies

$$\nu(\mathcal{G}) \geq \max\{|\mathcal{N}_{a_i^1}^\delta \setminus \mathcal{N}_{a_i^2}^\delta|, |\mathcal{N}_{a_i^2}^\delta \setminus \mathcal{N}_{a_i^1}^\delta|\}, \quad \forall i \in I. \quad (42)$$

Finally, (41) and (42) together imply (39). □

Next, we show that $r > \nu(\mathcal{G})$ is a sufficient condition to ensure that the paths between the states in \mathcal{X}_R^0 on a minimum resistance tree consist of unilateral experimentations. As such, if r is sufficiently large, then simultaneous experimentations become sufficiently rare as ϵ goes down to zero.

Definition (Unilateral Experimentation Path): A feasible sequence of states, $\mathcal{P} = \{x^1, x^2, \dots, x^n\}$, is a unilateral experimentation path if $x^1, x^n \in \mathcal{X}_R^0$, $x^2, \dots, x^{n-1} \in \mathcal{X}_T^0$ and for all $1 \leq p \leq n - 1$

$$|I_{se}(x^p, x^{p+1})| = \begin{cases} 1 & \text{if } p = 1, \\ 0 & \text{otherwise.} \end{cases} \quad (43)$$

Lemma 5.3. *Let \mathcal{T}^* be a minimum resistance tree, and let $x \rightarrow x^+ \in \mathcal{T}^*$. If $x \in \mathcal{X}_R^0$, then $|I_{se}(x, x^+)| = 1$.*

Proof. Since $x \in \mathcal{X}_R^0$, $|I_{se}(x, x^+)| > 0$, as otherwise, $x^+ = x$ and $x \rightarrow x^+$ cannot be contained in a tree. Assume that $|I_{se}(x, x^+)| > 1$. Then, choose an arbitrary $i \in I_{se}(x, x^+)$

to define an $\tilde{x}^+ \neq x$ as

$$\tilde{x}_j^+ = \begin{cases} x_j^+ & \text{if } j \neq i, \\ x_i & \text{otherwise.} \end{cases} \quad (44)$$

Note that $x \rightarrow \tilde{x}^+$ is a feasible transition, and $R(x, \tilde{x}^+) = R(x, x^+) - r(|I_{se}(x, x^+)| - 1)$. Replacing $x \rightarrow x^+$ with $x \rightarrow \tilde{x}^+$ would give an alternative tree with a smaller resistance, which contradicts with \mathcal{T} being a minimum resistance tree. \square

Lemma 5.4. *Let $x \rightarrow x^+ \in \mathcal{T}^*$ for some minimum resistance tree \mathcal{T}^* . If $x \in \mathcal{X}_T^0$ and $r > \nu(\mathcal{G})$, then $|I_{se}(x, x^+)| < |I_e(x)|$.*

Proof. Since $x \in \mathcal{X}_T^0$, $I_{se}(x, \tilde{x}^+) = \emptyset$ doesn't imply $\tilde{x}^+ = x$. Hence, there exists an $\tilde{x}^+ \neq x$ such that $x \rightarrow \tilde{x}^+$ is feasible and $I_{se}(x, \tilde{x}^+) = \emptyset$. For any such \tilde{x}^+ , we have

$$R(x, \tilde{x}^+) - R(x, x^+) \leq -r|I_{se}(x, x^+)| + |I_{es}(x, \tilde{x}^+)|\nu(\mathcal{G}). \quad (45)$$

Note that $|I_{es}(x, \tilde{x}^+)| \leq |I_e(x)|$. Hence, given $r > \nu(\mathcal{G})$, the right side of (45) is negative for any $|I_{se}(x, x^+)| \geq |I_e(x)|$. In that case, replacing $x \rightarrow x^+$ with $x \rightarrow \tilde{x}^+$ would give an alternative tree with a smaller resistance, which contradicts with \mathcal{T} being a minimum resistance tree. Consequently, $|I_{se}(x, x^+)| < |I_e(x)|$. \square

Lemma 5.5. *Let $r > \nu(\mathcal{G})$, and let $\mathcal{P} = \{x^1, x^2, \dots, x^n\}$ be a sequence of states, where $x^1, x^n \in \mathcal{X}_R^0$ and $x^2, \dots, x^{n-1} \in \mathcal{X}_T^0$. If $\mathcal{P} \in \mathcal{T}^*$ for some minimum resistance tree \mathcal{T}^* , then \mathcal{P} is a unilateral experimentation path.*

Proof. Since $x^1 \in \mathcal{X}_R^0$, from Lemma 5.3, we have $|I_{se}(x^1, x^2)| = 1$ leading to $|I_e(x^2)| = 1$. Furthermore, for $r > \nu(\mathcal{G})$, from Lemma 5.4, we have $|I_{se}(x^2, x^3)| = 0$. Hence, we have $|I_e(x^3)| \leq 1$. Using Lemma 5.4 recursively along \mathcal{P} we obtain

$$|I_{se}(x^p, x^{p+1})| = \begin{cases} 1 & \text{if } p = 1, \\ 0 & \text{otherwise.} \end{cases} \quad (46)$$

Hence, \mathcal{P} is a unilateral experimentation path. \square

Lemma 5.6. *If $\mathcal{P} = \{x^1, x^2, \dots, x^n\}$ be a unilateral experimentation path, then*

$$R(\mathcal{P}) = \sum_{p=1}^{n-1} R(x^p, x^{p+1}) = r + \max\{\phi(x^n), \phi(x^1)\} - \phi(x^n). \quad (47)$$

Proof. Since $\mathcal{P} = \{x^1, x^2, \dots, x^n\}$ be a unilateral experimentation path, for $x^p, x^{p+1} \in \mathcal{X}_T^0$, we have

$$|I_{se}(x^p, x^{p+1})| = |I_{es}(x^p, x^{p+1})| = 0. \quad (48)$$

Hence, such transitions have zero resistance, resulting in

$$R(\mathcal{P}) = R(x^1, x^2) + R(x^{n-1}, x^n). \quad (49)$$

Note that, since $x^1 \in \mathcal{X}_R^0$ and \mathcal{P} is a unilateral experimentation path, we have $R(x^1, x^2) = r$ and $R(x^{n-1}, x^n) = \Delta_i(x_i^{n-1}, x_i^n)$, where $i \in I$ is the unique experimenting agent. Since all the other agents are stationary, i.e. a_{-i} is constant

along \mathcal{P} , the estimated utilities satisfy

$$(\hat{U}_i^1)^{n-1} = \sum_{v \in \mathcal{N}_{a_i^1}^\delta} u(v, a_{-i}) = U_i(a_i^1, a_{-i}), \quad (50)$$

$$(\hat{U}_i^2)^{n-1} = \sum_{v \in \mathcal{N}_{a_i^2}^\delta} u(v, a_{-i}) = U_i(a_i^2, a_{-i}). \quad (51)$$

Plugging (50) and (51) into (28) we obtain

$$\Delta_i(x_i^{n-1}, x_i^n) = \max\{U_i(x^n), U_i(x^1)\} - U_i(x^n). \quad (52)$$

Since Γ_{DGC} is a potential game, from (52) we obtain

$$\Delta_i(x_i^{n-1}, x_i^n) = \max\{\phi(x^n), \phi(x^1)\} - \phi(x^n). \quad (53)$$

\square

Lemma 5.7. *Let $r > \nu(\mathcal{G})$, and let \mathcal{T}_x^* and $\mathcal{T}_{x'}^*$ be minimum resistance trees rooted at some $x, x' \in \mathcal{X}_R^0$. Then,*

$$R(\mathcal{T}_x^*) \leq R(\mathcal{T}_{x'}^*) \Rightarrow \phi(x) \geq \phi(x'). \quad (54)$$

Proof. For $r > \nu(\mathcal{G})$, in light of Lemma 5.5, the paths between the states in \mathcal{X}_R^0 on a minimum resistance tree consist of unilateral experimentations. Let $x_R^0 \in \mathcal{X}_R^0$, and let $\mathcal{T}_{x_R^0}^*$ be a minimum resistance tree rooted at x_R^0 . Let $x_R^n \in \mathcal{X}_R^0$ be a state such that $R(\mathcal{T}_{x_R^0}^*) \leq R(\mathcal{T}_{x_R^n}^*)$ and the unique path, $\mathcal{P} \in \mathcal{T}_{x_R^0}^*$, from x_R^n to x_R^0 consists of n unilateral experimentations, i.e.

$$R(\mathcal{P}) = \sum_{k=1}^n R(\mathcal{P}_k), \quad (55)$$

where \mathcal{P}_k is the unilateral experimentation starting at x_R^{n-k+1} and ending at x_R^{n-k} . Note that, for each such \mathcal{P}_k , there exists a feasible unilateral experimentation path \mathcal{P}'_k in the reversed direction, starting at x_R^{n-k} and ending at x_R^{n-k+1} . Replacing each \mathcal{P}_k with \mathcal{P}'_k , one can construct a tree rooted at $\mathcal{T}_{x_R^n}^*$. Note that the resistances of these trees satisfy

$$\begin{aligned} R(\mathcal{T}_{x_R^n}^*) - R(\mathcal{T}_{x_R^0}^*) &= \sum_{k=1}^n (R(\mathcal{P}'_k) - R(\mathcal{P}_k)) \\ &= \sum_{k=1}^n (\phi(x_R^{n-k}) - \phi(x_R^{n-k+1})) \\ &= \phi(x_R^0) - \phi(x_R^n). \end{aligned} \quad (56)$$

Note that by definition $R(\mathcal{T}_{x_R^n}^*) \leq R(\mathcal{T}_{x_R^0}^*)$. Hence, if $R(\mathcal{T}_{x_R^0}^*) \leq R(\mathcal{T}_{x_R^n}^*)$, then $R(\mathcal{T}_{x_R^0}^*) \leq R(\mathcal{T}_{x_R^n}^*)$ for any $\mathcal{T}_{x_R^n}^*$. Plugging this into (56), we obtain $\phi(x_R^0) \geq \phi(x_R^n)$. \square

Theorem 5.8. *Let $\mathcal{G} = (V, E)$ be a connected graph, let all agents follow the CFM algorithm with $r > \nu(\mathcal{G})$, and let x be a stochastically stable state of the resulting Markov chain. Then, $I_s(x) = I$ and $|V_c(x)| \geq |V_c(x')|$ for all $x' \in \mathcal{X}$.*

Proof. Let x be a stochastically stable state of the resulting Markov chain. Due to Lemma 3.2, $x \in \mathcal{X}_R^0$ and $R(\mathcal{T}_x^*) \leq R(\mathcal{T}_{x'}^*)$ for all $x' \in \mathcal{X}_R^0$. Hence, due to (36), $I_s(x) = I$. In light of Lemma 5.7, if $r > \nu(\mathcal{G})$, then $R(\mathcal{T}_x^*) \leq R(\mathcal{T}_{x'}^*)$ implies $\phi(x) \geq \phi(x')$ for all $x' \in \mathcal{X}_R^0$. Since $\phi(x) = |V_c(x)|$ for Γ_{DGC} and $V_c(x)$ only depends on the current positions of

the agents on the graph, $\phi(x) \geq \phi(x')$ for all $x' \in \mathcal{X}_R^0$ implies $|V_c(x)| \geq |V_c(x')|$ for all $x' \in \mathcal{X}$. \square

Theorem 5.8 indicates that if all agents follow the CFCM algorithm with sufficiently large r , then the stochastically stable states are all-stationary states maximizing the number of covered nodes. As such, the agents asymptotically maintain maximum coverage with an arbitrarily high probability for arbitrarily small values of the noise parameter ϵ .

C. Adaptive Implementation

As presented in the previous section, a sufficient condition for r to guarantee the stochastic stability of the coverage maximizers is that $r > \nu(\mathcal{G})$. Accordingly, an appropriate r can be determined if $\nu(\mathcal{G})$ or any graph property that defines an upper bound of $\nu(\mathcal{G})$ is available as prior information. When no prior information about the target graph is available, one possible way of achieving the same limiting behavior is to let each agent i have an individual r_i that is updated based on the local observations as the agent moves on the graph. In this section, we present such an adaptive implementation of the CFCM algorithm that achieves the desired limiting behavior by having every agent i eventually obtain a stationary $r_i = \beta + \nu(\mathcal{G})$ for some $\beta > 0$.

Algorithm III: Adaptive CFCM

```

1 : initialization:  $\epsilon \in \mathfrak{R}^+$  (small),  $\beta \in \mathfrak{R}^+$ ,  $a_i \in A_i$  arbitrary,
    $S_i = \{a_i\}$ ,  $k_i = 1$ ,  $\hat{U}_i^1 = \hat{U}_i^2 = 0$ ,  $r_i = \beta$ .
2 : repeat
3 :    $a_i = S_i(k_i)$ ,  $a_i^1 = S_i(1)$ ,  $a_i^2 = S_i(|S_i|)$ .
4 :   if ( $|S_i| = 1$ )
5 :     Generate a random (uniform)  $\gamma \in [0, 1]$ .
6 :     if ( $\gamma \leq \epsilon^{r_i}$ )
7 :        $a_i^2$  is randomly (uniform) chosen over  $A_i^c(a_i^1)$ .
8 :        $r_i = \max(r_i, \beta + |\mathcal{N}_{a_i^1}^\delta \setminus \mathcal{N}_{a_i^2}^\delta|, \beta + |\mathcal{N}_{a_i^2}^\delta \setminus \mathcal{N}_{a_i^1}^\delta|)$ .
9 :        $S_i = \mathcal{E}(a_i^1, a_i^2)$ .
10 :    end if
11 :   else
12 :     if ( $k_i \geq k$ ,  $\forall k \in \{k \mid S_i(k) = a_i\}$ )
13 :        $\hat{U}_i^1 = \hat{U}_i^1 + u_i(a_i, a_{-i})$ , if  $a_i \in \mathcal{N}_{a_i^1}^\delta$ .
14 :        $\hat{U}_i^2 = \hat{U}_i^2 + u_i(a_i, a_{-i})$ , if  $a_i \in \mathcal{N}_{a_i^2}^\delta$ .
15 :     end if
16 :     if ( $k_i = |S_i|$ )
17 :        $\alpha = \epsilon^{-\hat{U}_i^1}$ ,  $\beta = \epsilon^{-\hat{U}_i^2}$ .
18 :        $S_i = \begin{cases} \{a_i^1\} & \text{w.p. } \frac{\alpha}{\alpha+\beta}, \\ \{a_i^2\} & \text{otherwise.} \end{cases}$ 
19 :        $k_i = 1$ ,  $\hat{U}_i^1 = \hat{U}_i^2 = 0$ .
20 :     else
21 :        $k_i = k_i + 1$ .
22 :     end if
23 :   end if
24 : end repeat

```

The adaptive implementation differs from the CFCM algorithm by having each r_i as a state variable rather than a

constant. In analyzing the resulting Markov chain, we will use the same notation as before and the distinction should be clear from the context. Accordingly, the state of each agent i is denoted as

$$x_i = [S_i \quad k_i \quad \hat{U}_i^1 \quad \hat{U}_i^2 \quad r_i], \quad (57)$$

whereas the global state is the concatenation of the individual states, $x = [x_1, x_2, \dots, x_m]$, and \mathcal{X} denotes the state space. Since each r_i is monotonically increasing (see line 8 in Algorithm III), the Markov chain induced by the adaptive implementation is not irreducible unless $\nu(\mathcal{G}) = 0$. However, we will show that it has a unique limiting distribution that behaves the same as the limiting distribution under CFCM as ϵ goes down to zero.

Theorem 5.9. *Let $\mathcal{G} = (V, E)$ be a connected graph, and let all agents follow the adaptive CFCM algorithm. Then, the resulting Markov chain has a unique limiting distribution μ_ϵ^* such that*

$$\lim_{\epsilon \rightarrow 0^+} \mu_\epsilon^*(x) > 0 \Rightarrow x \in \mathcal{X}^*, \quad (58)$$

where

$$\mathcal{X}^* = \{x \in \mathcal{X} \mid r_i = \beta + \nu(\mathcal{G}), \forall i \in I, I_s(x) = I, |V_c(x)| \geq |V_c(x')|, \forall x' \in \mathcal{X}\}. \quad (59)$$

Proof. Let $\mathcal{G} = (V, E)$ be a connected graph, and let \mathcal{X}_c be the set of states such that each agent i has $k_i = \beta + \nu(\mathcal{G})$, i.e.,

$$\mathcal{X}_c = \{x \in \mathcal{X} \mid r_i = \beta + \nu(\mathcal{G}), \forall i \in I\}. \quad (60)$$

First, we will show that \mathcal{X}_c is the unique closed communicating class of the Markov chain, i.e., any $x \notin \mathcal{X}_c$ is transient, any state in \mathcal{X}_c is reachable from any other state in \mathcal{X}_c , and the probability of leaving \mathcal{X}_c is zero. Let $(v, v') \in E$ be such that $|\mathcal{N}_v^\delta \setminus \mathcal{N}_{v'}^\delta| = \nu(\mathcal{G})$. Note that, for any $\epsilon, r_i > 0$, any stationary agent i has a non-zero probability to experiment and pick any of the neighboring nodes as its next stationary position. Hence, for a connected graph \mathcal{G} , any agent i such that $r_i \neq \beta + \nu(\mathcal{G})$ can eventually be stationary at $a_i^1 = v$ and start an experiment by picking $a_i^2 = v'$. Note that each r_i is monotonically increasing under the adaptive CFCM (line 8), and such a feasible trajectory would result in $r_i = \beta + \nu(\mathcal{G})$. Furthermore, $\beta + \nu(\mathcal{G})$ is the maximum value that any r_i can attain under the adaptive CFCM due to the definition of $\nu(\mathcal{G})$ in (38). Hence, \mathcal{X}_c is reachable from any $x \notin \mathcal{X}_c$, and once inside \mathcal{X}_c , the probability of leaving is zero. Accordingly, any $x \notin \mathcal{X}_c$ is a transient state. Note that the behavior of the Markov chain inside \mathcal{X}_c is identical to the chain induced by CFCM with the same value of ϵ and $r = \beta + \nu(\mathcal{G})$. Hence, any state in \mathcal{X}_c is reachable from any other state in \mathcal{X}_c . Consequently, \mathcal{X}_c is the unique closed communicating class.

Any finite Markov chain with a unique closed communicating class is known to have unique stationary distribution that is determined by the transition probabilities within the closed communicating class. Hence, the chain induced by the adaptive CFCM has a unique limiting distribution μ_ϵ^* such that $\mu_\epsilon^*(x) = 0$ for any $x \notin \mathcal{X}_c$. Furthermore, for any $x \in \mathcal{X}_c$, it follows from Theorem 5.8 that if $\lim_{\epsilon \rightarrow 0^+} \mu_\epsilon^*(x) > 0$, then

$I_s(x) = I$ and $|V_c(x)| \geq |V_c(x')|$ for all $x' \in \mathcal{X}_c$. Since $V_c(x)$ is determined by only the agent positions, $|V_c(x)| \geq |V_c(x')$ for all $x' \in \mathcal{X}_c$ implies $|V_c(x)| \geq |V_c(x')|$ for all $x' \in \mathcal{X}$. Consequently, if $\lim_{\epsilon \rightarrow 0^+} \mu_\epsilon^*(x) > 0$, then $x \in \mathcal{X}^*$. \square

In light of Theorem 5.9, if all the agents follow the adaptive CFCM algorithm on any connected graph \mathcal{G} , then they asymptotically maintain maximum coverage with an arbitrarily high probability for arbitrarily small values of the noise parameter ϵ . The necessity of choosing a sufficiently large r is removed for the adaptive implementation since the agents learn a suitable value as they move on the graph.

VI. SIMULATION RESULTS

In this section, some simulation results are presented to demonstrate the performance of the proposed method. In the simulations, a group of 13 agents are initially placed at an arbitrary node of a connected random geometric graph. Each agent has a sensing range $\delta = 1$. The graph consists of 45 nodes and 62 edges, and it has $\nu(\mathcal{G}) = 2$.

In the first simulation, the agents follow the adaptive CFCM algorithm with $\epsilon = 0.03$ and $\beta = 0.01$. The number of covered nodes throughout a period of 150000 time steps is shown in Fig. 5, whereas the configuration of the agents on the graph at some instants are provided in Fig. 6. As depicted in Fig. 5, after a sufficient amount of time, the agents maintain complete coverage with a very high probability. For $90000 \leq t \leq 150000$, the average number of covered nodes at each time step is computed as 44.77.

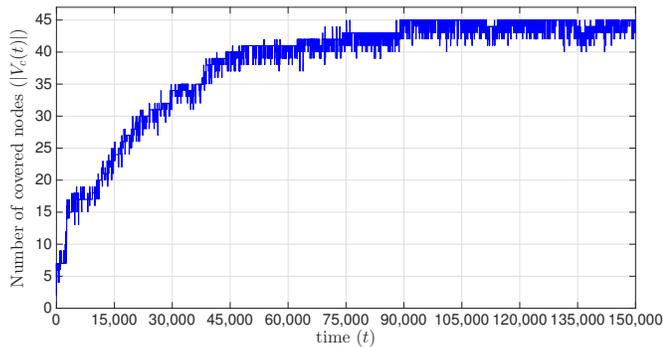


Fig. 5. The number of covered nodes as a function of time (adaptive CFCM).

In the second simulation, we consider the same scenario for the agents following the CFCM algorithm. In CFCM, the parameter r sets the likelihood of each stationary agent to update its action. Accordingly, increasing r also increases the mixing time of the resulting Markov chain. Hence, in practice, it is desirable to choose r as small as possible while ensuring the stochastic stability of the optimal configurations. Note that a suitable choice of r depends on the graph structure, and we provided $r > \nu(\mathcal{G})$ as a sufficient condition. However, it may not be necessary in many cases since simultaneously updating agents do not necessarily influence the utility estimations of each other, especially when they are sufficiently far from each other. In this simulation, we demonstrate such a case. The

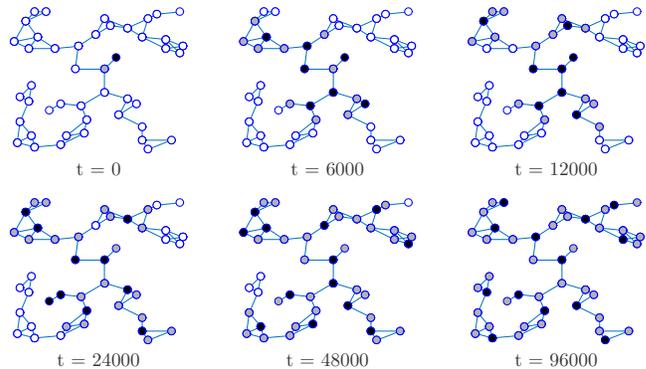


Fig. 6. The configuration of 13 agents on the graph at some instants of the simulation (adaptive CFCM). The nodes occupied by at least one agent are black, the nodes covered by at least one agent are gray, and the nodes that are not covered are white.

agents start at the same initial condition as the previous simulation, and they follow the CFCM algorithm with $\epsilon = 0.03$ and $r = 1.6$. The number of covered nodes throughout a period of 75000 time steps is shown in Fig. 7, whereas the configuration of the agents on the graph at some instants are provided in Fig. 8. For $45000 \leq t \leq 75000$, the average number of covered nodes at each time step is computed as 44.48.

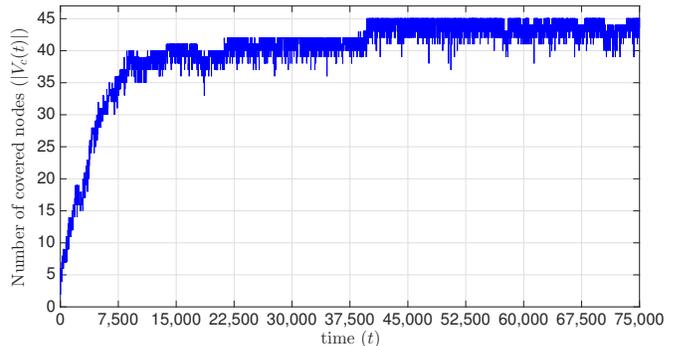


Fig. 7. The number of covered nodes as a function of time (CFCM).

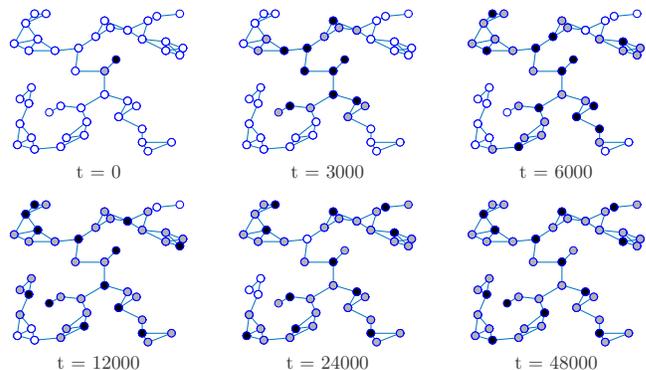


Fig. 8. The configuration of 13 agents on the graph at some instants of the simulation (CFCM). The nodes occupied by at least one agent are black, the nodes covered by at least one agent are gray, and the nodes that are not covered are white.

Finally, in order to compare the performance with a setting

that allows for communications, we also present a simulation of the same scenario with BLLL. The agents start at the same initial condition as the previous simulations, and BLLL is executed with the same noise parameter $\epsilon = 0.03$. The number of covered nodes throughout a period of 5000 time steps is shown in Fig. 9, whereas the configuration of the agents on the graph at some instants are provided in Fig. 10. As illustrated in Fig. 9, after a sufficient amount of time, the agents maintain complete coverage with a very high probability. For $3000 \leq t \leq 5000$, the average number of covered nodes at each time step is computed as 44.64.

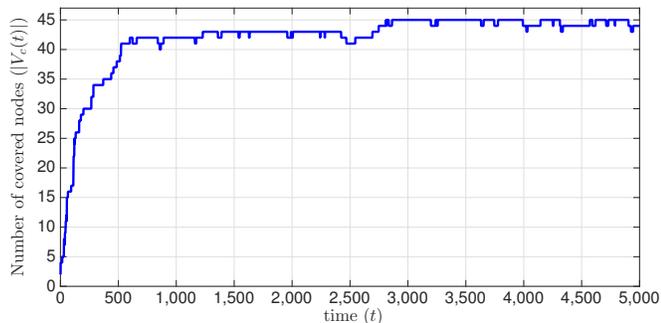


Fig. 9. The number of covered nodes as a function of time (BLLL).

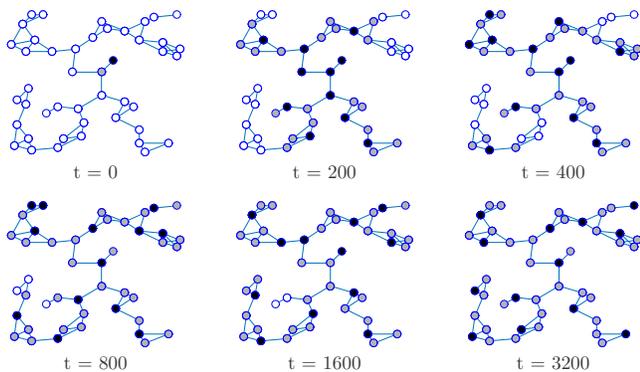


Fig. 10. The configuration of 13 agents on the graph at some instants of the simulation (BLLL). The nodes occupied by at least one agent are black, the nodes covered by at least one agent are gray, and the nodes that are not covered are white.

The simulation results show that the agents achieve optimal coverage in all three cases. When the agents are allowed to communicate, they can maximize the coverage much faster as expected. Despite the slower convergence to the limiting distribution, the main advantage of the CFCM algorithm is that the agents do not require any explicit communications, and optimal coverage is achieved if the parameter r is selected sufficiently large. Furthermore, the adaptive implementation of CFCM achieves optimal coverage without relying on the a-priori selection of a suitable r . Accordingly, CFCM or its adaptive implementation can be employed to optimally distribute some mobile resources on networks, even in scenarios that do not allow for explicit communications.

VII. CONCLUSION

In this paper, a game theoretic approach was proposed for distributed coverage of networked systems by mobile agents with local capabilities. We considered a distributed graph coverage (DGC) problem, where the network is modeled as an undirected graph, and the agents are located on some nodes of the graph. Each agent can sense the graph structure and the presence of the other agents within its δ -neighborhood, where δ is the sensing range. Any node of the graph is covered if it is within the sensing range of at least one agent. The agents move locally on the graph, and they aim to maximize the number of covered nodes. We studied this problem particularly for agents with no explicit communications among themselves.

A game theoretic formulation of the DGC problem was obtained by designing a potential game, Γ_{DGC} . In Γ_{DGC} , the action of each agent is defined as its position on the graph, and the utility of each agent is equal to the number of nodes covered only by itself. It was shown that Γ_{DGC} can be paired with a learning algorithm such as BLLL to maximize the coverage. However, such learning algorithms require the agents to measure their current utilities. In Γ_{DGC} , the actual utilities can not be computed without explicit communications since the agents with overlapping coverage are not necessarily within the sensing range of each other. In order to address this issue, we presented a communication-free learning algorithm, namely the CFCM. In CFCM, the agents follow a noisy best-response policy based on the estimated utilities gathered by moving around their current positions. The algorithm has a noise parameter, $\epsilon \in \mathbb{R}^+$, and a second parameter, $r \in \mathbb{R}^+$, that sets the likelihood of remaining stationary. We showed that the CFCM algorithm induces a regular perturbed Markov chain and the stochastically stable states are the coverage maximizers for sufficiently large values of r . A sufficient value of r was derived from the topology of the graph. Furthermore, we presented an adaptive implementation of CFCM such that the agents learn a suitable value of r based on their local observations as they move on the graph. As such, they achieve maximum coverage on any connected graph, without any prior information. Some simulation results were also presented to demonstrate the performance of the proposed method.

REFERENCES

- [1] W. Goddard, S. M. Hedetniemi, and S. T. Hedetniemi, "Eternal security in graphs," *J. Combin. Math. Combin. Comput.*, vol. 52, pp. 169–180, 2005.
- [2] T. C. Du, E. Y. Li, and A.-P. Chang, "Mobile agents in distributed network management," *Communications of the ACM*, vol. 46, no. 7, pp. 127–132, 2003.
- [3] G. Berbeglia, J.-F. Cordeau, and G. Laporte, "Dynamic pickup and delivery problems," *European Journal of Operational Research*, vol. 202, no. 1, pp. 8–15, 2010.
- [4] J. Reese, "Solution methods for the p-median problem: An annotated bibliography," *Networks*, vol. 48, no. 3, pp. 125–142, 2006.
- [5] N. Megiddo, E. Zemel, and S. L. Hakimi, "The maximum coverage location problem," *SIAM Journal on Algebraic Discrete Methods*, vol. 4, no. 2, pp. 253–261, 1983.
- [6] S. Khuller, A. Moss, and J. S. Naor, "The budgeted maximum coverage problem," *Information Processing Letters*, vol. 70, no. 1, pp. 39–45, 1999.
- [7] S. H. Owen and M. S. Daskin, "Strategic facility location: A review," *European Journal of Operational Research*, vol. 111, no. 3, pp. 423–447, 1998.

[8] A. Caprara, P. Toth, and M. Fischetti, "Algorithms for the set covering problem," *Annals of Operations Research*, vol. 98, no. 1-4, pp. 353–371, 2000.

[9] A. Howard, M. J. Mataric, and G. S. Sukhatme, "Mobile sensor network deployment using potential fields: A distributed, scalable solution to the area coverage problem," in *Distributed Autonomous Robotic Systems 5*, pp. 299–308, Springer, 2002.

[10] M. Schwager, D. Rus, and J.-J. Slotine, "Decentralized, adaptive coverage control for networked robots," *International Journal of Robotics Research*, vol. 28, no. 3, pp. 357–375, 2009.

[11] S. Poduri and G. S. Sukhatme, "Constrained coverage for mobile sensor networks," in *IEEE International Conference on Robotics and Automation*, pp. 165–171, 2004.

[12] J. Cortés, S. Martínez, T. Karatas, and F. Bullo, "Coverage control for mobile sensing networks," *IEEE Transactions on Robotics and Automation*, vol. 20, no. 2, pp. 243–255, 2004.

[13] S. Lloyd, "Least squares quantization in pcm," *IEEE Transactions on Information Theory*, vol. 28, no. 2, pp. 129–137, 1982.

[14] J. Cortes, S. Martinez, and F. Bullo, "Spatially-distributed coverage optimization and control with limited-range interactions," *ESAIM: Control, Optimisation and Calculus of Variations*, vol. 11, no. 4, pp. 691–719, 2005.

[15] A. Kwok and S. Martínez, "Deployment algorithms for a power-constrained mobile sensor network," *International Journal of Robust and Nonlinear Control*, vol. 20, no. 7, pp. 745–763, 2010.

[16] L. Pimenta, V. Kumar, R. C. Mesquita, and G. Pereira, "Sensing and coverage for a network of heterogeneous robots," in *IEEE Conference on Decision and Control*, pp. 3947–3952, 2008.

[17] J. W. Durham, R. Carli, P. Frasca, and F. Bullo, "Discrete partitioning and coverage control with gossip communication," in *ASME Dynamic Systems and Control Conference*, pp. 225–232, 2009.

[18] S. Yun and D. Rus, "Distributed coverage with mobile robots on a graph: Locational optimization," in *IEEE International Conference on Robotics and Automation*, pp. 634–641, 2012.

[19] M. Zhu and S. Martínez, "Distributed coverage games for energy-aware mobile sensor networks," *SIAM Journal on Control and Optimization*, vol. 51, no. 1, pp. 1–27, 2013.

[20] A. Y. Yazıcıoğlu, M. Egerstedt, and J. S. Shamma, "A game theoretic approach to distributed coverage of graphs by heterogeneous mobile agents," in *IFAC Workshop on Distributed Estimation and Control in Networked Systems*, pp. 309–315, 2013.

[21] G. Arslan, J. Marden, and J. S. Shamma, "Autonomous vehicle-target assignment: a game theoretical formulation," *ASME Journal of Dynamic Systems, Measurement, and Control*, pp. 584–596, 2007.

[22] A. Arsie, K. Savla, and E. Frazzoli, "Efficient routing algorithms for multiple vehicles with no explicit communications," *IEEE Transactions on Automatic Control*, vol. 54, no. 10, pp. 2302–2317, 2009.

[23] J. Huang, Z. Han, M. Chiang, and H. V. Poor, "Auction-based resource allocation for cooperative communications," *IEEE Journal on Selected Areas in Communications*, vol. 26, no. 7, pp. 1226–1237, 2008.

[24] L. Jia, R. Rajaraman, and T. Suel, "An efficient distributed algorithm for constructing small dominating sets," *Distributed Computing*, vol. 15, no. 4, pp. 193–205, 2002.

[25] F. Kuhn and R. Wattenhofer, "Constant-time distributed dominating set approximation," *Distributed Computing*, vol. 17, no. 4, pp. 303–310, 2005.

[26] Z. Abrams, A. Goel, and S. Plotkin, "Set k-cover algorithms for energy efficient monitoring in wireless sensor networks," in *International Symposium on Information Processing in Sensor Networks*, pp. 424–432, 2004.

[27] J. R. Marden, G. Arslan, and J. S. Shamma, "Cooperative control and potential games," *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, vol. 39, no. 6, pp. 1393–1407, 2009.

[28] L. E. Blume, "The statistical mechanics of strategic interaction," *Games and Economic Behavior*, vol. 5, no. 3, pp. 387–424, 1993.

[29] J. R. Marden and J. S. Shamma, "Revisiting log-linear learning: Asynchrony, completeness and payoff-based implementation," *Games and Economic Behavior*, vol. 75, no. 2, pp. 788–808, 2012.

[30] H. P. Young, "The evolution of conventions," *Econometrica: Journal of the Econometric Society*, vol. 61, no. 1, pp. 57–84, 1993.

[31] M. Kandori, G. J. Mailath, and R. Rob, "Learning, mutation, and long run equilibria in games," *Econometrica: Journal of the Econometric Society*, pp. 29–56, 1993.

[32] S. Boyd, A. Ghosh, B. Prabhakar, and D. Shah, "Randomized gossip algorithms," *IEEE Transactions on Information Theory*, vol. 52, no. 6, pp. 2508–2530, 2006.

[33] R. Tarjan, "Depth-first search and linear graph algorithms," *SIAM Journal on Computing*, vol. 1, no. 2, pp. 146–160, 1972.



A. Yasin Yazıcıoğlu is a postdoctoral associate at the Laboratory for Information and Decision Systems, Massachusetts Institute of Technology. He received the B.S. and M.S. degrees in Mechatronics Engineering from Sabancı University, Turkey, in 2007 and 2009 respectively, and the Ph.D. degree in Electrical and Computer Engineering from the Georgia Institute of Technology in 2014. His recent research is primarily focused on control theory, multi-agent systems, game theory, and networks.



Magnus Egerstedt is the Schlumberger Professor in the School of Electrical and Computer Engineering at the Georgia Institute of Technology, where he serves as Associate Chair for Research. He received the M.S. degree in Engineering Physics and the Ph.D. degree in Applied Mathematics from the Royal Institute of Technology, Stockholm, Sweden, the B.A. degree in Philosophy from Stockholm University, and was a Postdoctoral Scholar at Harvard University. Dr. Egerstedt is the director of the Georgia Robotics and Intelligent Systems Laboratory (GRITS Lab), a Fellow of the IEEE, and a recipient of a number of research and teaching awards, including the Ragazzini Award from A2C2.



Jeff S. Shamma is the Julian T. Hightower Chair in Systems & Control (currently on leave) in the School of Electrical and Computer Engineering at the Georgia Institute of Technology (Georgia Tech) and a Professor of Electrical Engineering in the King Abdullah University of Science and Technology (KAUST). He received a BS in Mechanical Engineering from Georgia Tech in 1983 and a PhD in Systems Science and Engineering from the Massachusetts Institute of Technology in 1988. He held faculty positions at the University of Minnesota, University of Texas-Austin, and University of California-Los Angeles. Prof. Shamma is a recipient of the NSF Young Investigator Award (1992), the American Automatic Control Council Donald P. Eckman Award (1996), and the Mohammed Dahleh Award (2013), and he is a Fellow of the IEEE (2006). He is currently an Associate Editor for *Games* (2012–present) and a Senior Editor for the *IEEE Transactions on Control of Network Systems* (2013–present). Prof. Shamma's research is in the general area of feedback control and systems theory. His most recent research has been in decision and control for distributed multiagent systems and the related topics of game theory and network science, with applications to cyberphysical and societal network systems.