

# Multiagent UAV Routing: A Game Theory Analysis With Tight Price of Anarchy Bounds

Omkar Thakoor, Jugal Garg<sup>ID</sup>, and Rakesh Nagi

**Abstract**—We study the multiagent unmanned aerial vehicle (UAV) routing problem where a set of UAVs needs to collect information via surveillance of an area of operation. Each UAV is autonomous and does not rely on a reliable communication medium to coordinate with other UAVs. We formulate the problem as a game where UAVs are players and their strategies are the different routes they can take. Our model also incorporates the useful concept of information fusion. This results in a new variant of weighted congestion-type games. We show that the price of anarchy (PoA) of the game is at most 2, irrespective of the number of UAVs and their sensor capabilities. This also validates the empirical results of earlier works. Furthermore, we identify classes of games for the existence of a pure Nash equilibrium. To the best of our knowledge, these are the first such theoretical results in the related literature. Finally, we conduct experimental studies using randomly generated instances with several multiagent UAV routing policies. Our insights are that PoA increases with the congestion level when the same number of UAVs search a smaller area or more UAVs search the same area, and on an average, our proposed policies are less than 10% worse than the centralized optimal for the problem scenarios attempted.

**Note to Practitioners**—UAVs are becoming increasingly popular for information collection tasks in defense and civilian applications alike. When the collection area is large, it is not unusual that a fleet of UAVs is deployed. Routing of a fleet can be performed in a centralized or decentralized manner. Decentralized routing might be the only possibility when centralized situational awareness is not possible due to bandwidth limitations and centralized optimal routes for each UAV in the fleet are too complex to compute. Autonomous solutions have several other advantages, let alone simplicity. For managers of UAV systems, our work provides the first theoretical characterization of how bad could decentralized routing be. Under various scenarios of information fusion, specifically weak and strong, and the attribution of information collected to each UAV of a team, we prove that the

fleet will collect at least 50% of the best-centralized solution. Empirically, we show that, in fact, the performance of the fleet is much better and generally not worse than 10% of the best-centralized solution. Hopefully, our routing strategies provide valuable guidance to the practicing engineer or manager of a UAV fleet.

**Index Terms**—Game theory, multiagent systems, price of anarchy (PoA), surveillance and reconnaissance, unmanned aerial vehicles (UAVs).

## I. INTRODUCTION

Unmanned aerial vehicles (UAVs) are increasingly being used for intelligence, defense, and civilian information gathering and monitoring. This is particularly due to their utility in high endurance and perilous environments that are—as characterized by [1]—dull, dirty, and dangerous. A popular application of UAVs is information collection via surveillance of an area of operations. Often, a fleet of UAVs is dispatched for large geographical coverage and multiple intelligence, surveillance, and reconnaissance (ISR) missions where the goal is to maximize the amount of information collected, and this results in the problem of routing and coordination.

This problem can be formulated into two different ways depending on the application and environment. In one formulation, all UAVs coordinate and cooperate either directly or through an omniscient planner or both. It relies on the existence of a reliable communication medium between all agents [UAVs, planner (if present), etc.]. In this case, a centralized and cooperative solution is desired which maximizes the system throughput, and it essentially becomes an optimization problem which has been well studied but remains computationally challenging (see Section I-B).

In another formulation, agents are autonomous, i.e., they need to decide their own route themselves. However, if multiple agents obtain the same information then they need to share the payoff received in an appropriate manner. Full autonomy is the general trend in the UAV research due to its nondependence on a reliable communication medium, which often can have several issues as discussed in [2], also for missions that require radio silence, jamming, limited communication radius, or due to different competitive units participating with different subobjectives, or due to competitive self-interested agents who want to maximize their own payoff (i.e., the amount of information collected), among others.

While the former problem has the optimal outcome, it is unclear as to how much efficiency is lost due to duplication of effort in the latter. This paper provides an answer to this

Manuscript received April 3, 2018; revised August 21, 2018, October 19, 2018, and January 1, 2019; accepted February 11, 2019. Date of publication March 20, 2019; date of current version January 9, 2020. This paper was recommended for publication by Associate Editor D. V. Dimarogonas and Editor S. Reveliotis upon evaluation of the reviewers' comments. The work of J. Garg was supported by NSF CRII under Award 1755619. The work of R. Nagi was supported in part by ONR through the Program Management of Dr. D. Wagner under Award N000014-16-1-2245. (Corresponding author: Jugal Garg.)

O. Thakoor is with the Department of Computer Science, University of Southern California, Los Angeles, CA 90089 USA (e-mail: othakoor@usc.edu).

J. Garg and R. Nagi are with the Department of Industrial and Enterprise Systems Engineering, University of Illinois at Urbana-Champaign, Urbana, IL 61801 USA (e-mail: jugal@illinois.edu; nag@illinois.edu).

Color versions of one or more of the figures in this article are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TASE.2019.2902360

TABLE I  
SUMMARY OF CONTRIBUTIONS

Game Type	$p$ -player ( $p \geq 2$ )
Single Step	<ul style="list-style-type: none"> <li>• Existence of PNE (Theorem 1)</li> <li>• PoA <math>\leq 2</math> (Theorem 5)</li> </ul>
Homogeneous fleet	• PoA, PoS $\leq 2 - \frac{1}{p}$ (Theorem 3)
(Mild/Strong fusion)	• Existence of PNE (Theorem 1)
Multi Step	• Non-existence of PNE
With temporal aspect	• PoA $\leq 2$ (Theorem 5)
Without temporal aspect	• PoA $\leq 2$ (Corollary 3)

question. We formulate the problem by dividing the region of surveillance into discrete cells each having an associated information value, which we treat as an abstract entity allowing flexibility for what it represents; possibly a prior probability of a target being present, which could be obtained using a prior surveillance activity (say using a satellite). Such discretization is common in models for UAV path planning such as [3]. Our basic information model has some similarities with the earlier studied models, e.g., [4], [5], with some interpretational differences with various parameters (see Section I-B for details). In addition, our model incorporates a useful concept of *information fusion*, which can be defined as “the combination of multiple sources to obtain improved information (cheaper, greater quality, or greater relevance)” [6] and has been used in robotics and military applications [7]. This makes the information collection problem versatile in terms of applicability, especially when the fleet is heterogeneous in sensing capabilities.

Despite the problem being a cooperative scenario on a broad level, the autonomy of the agents and the lack of communication with each other (due to various aforementioned reasons) make it plausible to assume that they can obtain the information about the current positions of other players but cannot coordinate with each other to come up with a cooperative solution. Hence, our model studies what happens if they simply follow a selfish behavior maximizing their own payoffs. A natural framework to model such a situation is a noncooperative game where UAVs are the players and their strategies are the different routes that they can take. The stable outcomes of such a game are given by the Nash equilibria (NE), where no player gains by a unilateral deviation. Furthermore, their inefficiency is measured by the notions of price of anarchy (PoA) and price of stability (PoS), where PoA (PoS) is defined as the ratio of the value of optimal output to the value of the worst (best) NE outcome. We provide results on the existence of NE and tight bounds on the PoA and PoS, summarized in Table I. To the best of our knowledge, these are the first such theoretical results in the related literature. We note that all our results are applicable to a general setting, where the search space is given by a directed graph whose nodes represent the regions and the edges represent the connectivity among regions.

#### A. Our Contributions

In the basic game, each UAV  $i$  has a sensor effectiveness of  $\rho_i \in (0, 1]$ . If  $i$  visits a region, having information value  $v$ ,

then it collects  $\rho_i v$  amount and the remaining  $(1 - \rho_i)v$  is left at this region. If two UAVs,  $i$  and  $j$ , visit a region together, then they collect a total of  $\rho_i v + \rho_j(1 - \rho_i)v = (1 - (1 - \rho_i)(1 - \rho_j))v$  and split it according to their  $\rho$  values, i.e.,  $i$ 's share is a  $\rho_i/(\rho_i + \rho_j)$  fraction of the total. In case of more than two UAVs visiting a region together, the total *value function*, which captures the total amount of information collected, and the individual shares are defined similarly. Note that it is independent of the order in which the UAVs are considered.

The UAV game turns out to be a novel variant of the class of well-known *weighted congestion-type* games, where  $\rho_i$  can be treated as the weight of player  $i$ . The critical difference is that the total value function is not a function of the total weight of the players. To the best of our knowledge, no result is known for the weighted congestion-type games with the total value function we consider in this paper. Hence, in addition to the UAV application, our results may be of independent interest.

We show that a single-step  $p$ -player ( $p \geq 2$ ) game always has a pure NE (PNE) by showing that it admits the *finite improvement property*. The existence of a pure equilibrium is particularly important for the UAV game, for its practical use—Nash's theorem [8] only implies the existence of a mixed equilibrium where players randomize their pure strategies for a stable outcome. We show that the PoA of the UAV game is at most 2 using the framework of  $(\lambda, \mu)$ -smoothness [9]. Furthermore, in case of a homogeneous fleet, i.e., when each UAV has the same  $\rho$ , we show that PoA and PoS are at most  $2 - 1/p$  and they are tight. These bounds validate previous empirical results on the basic model, e.g., [4], in an even more general setting.

We also study the role of information fusion in the UAV game. In many situations, UAVs together can obtain a more refined information of a region by visiting it simultaneously (via fusion gain) than by visiting singly, in which case, they are awarded a greater payoff for the improved quality. We incorporate this using the fusion parameters  $\gamma_j$ ,  $1 \leq j \leq p$  as follows: when  $j$  UAVs visit a region simultaneously, then their resulting payoffs are multiplied by  $\gamma_j$ . By definition, we have  $1 = \gamma_1 \leq \dots \leq \gamma_p$ . We consider two special cases of *mild* and *strong* fusions where the value function is monotone nonincreasing and monotone nondecreasing, respectively (observe that it is nonincreasing in case of no fusion). For these special cases, we show that a PNE always exists.

In a multiple-step multiplayer game, each UAV needs to decide a walk to visit multiple regions, and its payoff is the sum of partial payoffs it gets by visiting each cell on its walk. We consider two different games, *temporal* and *nontemporal*, which differ in the payoff formulations, each being useful in different situations. In the former, the payoff of visiting a region is immediately awarded, and in the latter, it is awarded at the end with no regard to when the region was visited. For both formulations, we show that PNE may not exist. Furthermore, we show that PoA is at most 2, irrespective of the number of UAVs, their sensor capabilities, and the length of the walk. Furthermore, these are strong PoA bounds as

they apply to a more general solution concept called (*coarse*) *correlated equilibrium* of the game.

### B. Related Work

The problem of information collection has been widely studied in the search theory literature. A classical problem here is to maximize the probability of detecting a hidden target, e.g., as in [10]–[16]. On the other hand, Ortiz-Pena *et al.* [5] associate a potential information gain with each subregion based on an entropy-based function and aims to maximize the total gain. Our basic game model is similar to this and that of [4] in terms of discretization of the search space, time steps, and payoff formulations. Also, our representation of the surveillance region as a graph is similar to that of, say, [17] and [18]; however, we use the reasonable simplification of unweighted edges and the useful generalization of directed edges.

A large body of the previous work on the routing problem, e.g., [19]–[21], assumes a centralized control and full communication among the UAVs and the central controller. In situations where communication is feasible, challenges arise concerning key communication facets of energy consumption, power transmission, and transmission data rate parameters. Mozaffari *et al.* [22] investigate optimal UAV deployment with respect to wireless coverage, whereas Sikeridis *et al.* [23] present a holistic view of challenges in efficient resource management and optimal communication establishment in the deployment of UAVs. Several other works focusing on these issues, those in the context of UAVs and others, have also used game-theoretic models, for the corresponding resource allocation problems in minimizing energy or power consumption [24]–[27]. However, as mentioned earlier, in many situations such as when conducting UAV surveillance in sensitive regions, e.g., [28], the infeasibility of communication is a critical constraint, consequently, full autonomy is desirable. The benefits of autonomous UAVs for various domains have been highlighted, such as in [29] for persistent ISR missions.

A game-theoretic analysis is a perfect fit to tackle this problem under these circumstances. Game-theoretic models have been deployed in numerous other routing problems in transportation and networking applications such as [30]–[33]. The class of games, we formulate in this paper, is closely related to the class of congestion games or resource-selection games, and numerous other variants, defined, and studied in—most notably [34]–[40]. Our class of games has some critical differences with these well-studied classes of games, in terms of cost-sharing protocols and player weights, and hence, the results established for them do not directly apply in our case, making our theoretical results on the existence of pure equilibria and bounds on PoA and PoS, interesting and nontrivial.

### C. Organization of this Paper

In Section II, we present the UAV game model and establish its relationship to congestion-type games. Section III contains our results for the multiplayer single-step games. The existence

TABLE II  
NOTATION

$\rho_i$	Sensor effectiveness parameter for player $i$
$\mathcal{P}, p$	Set of players and number of players, i.e. $ \mathcal{P} $ , resp.
$\mathcal{C}$	Set of cells
$\gamma_i$	Fusion gain factor with $i$ UAVs
$l$	Number of time-steps
$v(c)$	Information in cell $c$
$\pi_i(s, c)$	$i$ 's payoff for visiting cell $c$ , when the outcome is $s$
$\pi_i(s)$	$i$ 's total payoff when the outcome is $s$ , i.e., $\sum_{c \in \mathcal{C}} \pi_i(s, c)$
$m_A(i)$	Multiplicity of $i$ in $A$
$A _i, A -i$	The multiset $m_A(i) \otimes \{i\}$ , and $A \setminus A _i$ resp.
$\pi_B^A(c)$	The combined payoff of visitors in multiset $B$ from visiting cell $c$ when preceded by exactly the visitors in multiset $A$ (Game with temporal aspect)
$\theta_B^A(c)$	The combined payoff of visitors in multiset $B$ from visiting cell $c$ when the complete set of visitors for cell $c$ is multiset $A$ (Game without temporal aspect)

of a PNE is shown in Section III-A. Section III-C presents the PoA and PoS results for the case of a homogeneous fleet. Section IV presents our results for the multiplayer multistep games. The results for multistep games with temporal aspects and without temporal aspects are presented in Sections IV-C and IV-D, respectively. In Section V, we perform numerical studies and compare the results for several routing strategies by simulating plausible scenarios with randomly generated game parameters.

## II. UAV GAME MODEL

### A. Preliminaries

Consider a game with  $n$  players. Let  $S_i$  be the strategy set for player  $i$ , and  $S = \times_i S_i$  be the set of joint strategy profiles. It is known that the NE always exists [8], however, a PNE may not exist, defined as follows.

**Definition 1:** A joint strategy profile  $s^* \in S$  is a PNE if no unilateral deviation in strategy by any single player is profitable for that player, that is,

$$\forall i, s_i \in S_i : \pi_i(s_i^*, s_{-i}^*) \geq \pi_i(s_i, s_{-i}^*)$$

where  $\pi_i(s)$  is the payoff function of player  $i$ , and  $s_{-i}$  denotes the strategies of all players except  $i$ .

The social welfare of an outcome is defined as the sum of the individual payoffs of the players. PoS and the PoA are the two well-known metrics used in economics and game theory, to quantify the *inefficiency* of the equilibria—that is, how bad the social welfare at an equilibrium is as compared to the optimum social welfare achievable. In the following definitions, the *best* and the *worst* equilibria refer to those which give the maximum and the minimum social welfare among all equilibria.

**Definition 2:** The PoS is defined as the ratio of the optimal social welfare to the welfare for the best NE. The PoA is defined as the ratio of the optimal social welfare to the welfare for the worst NE.

The smoothness of these games is defined as follows.

**Definition 3** ( $\lambda, \mu$ )-smoothness [41]: A payoff-maximization game—one where each player strives to



maximize his payoff  $\pi(s)$ —is called  $(\lambda, \mu)$ -smooth if

$$\forall s, s^* \in \mathcal{S}, \sum_{i \in \mathcal{P}} \pi_i(s_i^*, s_{-i}) \geq \lambda \sum_{i \in \mathcal{P}} \pi_i(s^*) - \mu \sum_{i \in \mathcal{P}} \pi_i(s). \quad (1)$$

The UAV game we model is closely related to the congestion-type games defined as follows.

**Definition 4:** An arbitrary congestion-type payoff-maximization game consists of a resource set  $\mathcal{C}$  and a player set  $\mathcal{P} = \{1, \dots, p\}$ , where each resource  $c \in \mathcal{C}$  has a joint value function  $M_c : 2^{\mathcal{P}} \rightarrow \mathbb{R}^+$  defined on the subsets of  $\mathcal{P}$ , which describes the worth of the resource as a function of the set of players sharing it. Each  $i \in \mathcal{P}$  has a strategy set  $\mathcal{S}_i \subseteq 2^{\mathcal{C}}$  and a per-resource payoff function  $U_i : \mathbb{R}^+ \rightarrow \mathbb{R}^+$ —if  $\mathcal{Q} \subseteq \mathcal{P}$  is the set of players using  $c \in \mathcal{C}$  (where  $i \in \mathcal{Q}$ ), the value of  $c$  is  $M_c(\mathcal{Q})$  and consequently,  $i$ 's payoff for using  $c$  is  $U_i(M_c(\mathcal{Q}))$ . The net payoff for  $i \in \mathcal{P}$  on playing strategy  $s_i \in \mathcal{S}_i$  is the sum of its payoffs for using resources  $c \in s_i$ .

Thus, these games can be defined as a tuple  $(\mathcal{P}, \mathcal{C}, (\mathcal{S}_i)_{i \in \mathcal{P}}, (M_c)_{c \in \mathcal{C}}, (U_i)_{i \in \mathcal{P}})$  with each entity as defined earlier. Arbitrary congestion-type cost-minimization games are similarly defined as well. However, this paper only deals with the payoff-maximization games, and for convenience, we simply call these as the arbitrary congestion-type games. When each strategy has a single resource, i.e.,  $|s_i| = 1, \forall s_i \in \mathcal{S}_i, \forall i \in \mathcal{P}$ , then we call them *singleton congestion-type games*.

Our game model extensively uses the *multisets*, a generalization of the concept of sets, with the distinction that a multiset can have multiple instances of any of its elements.

**Definition 5:** A multiset can be represented by a two-tuple  $(X, m)$ , where  $X$  is the set of distinct elements in the multiset, and  $m$  is the multiplicity function, such that, for each  $x \in X$ ,  $m(x)$  is the number of instances of  $x$  in the multiset. By convention, we have  $x \notin X \Leftrightarrow m(x) = 0$ .

Thus,  $X$  can be viewed as the support set of the function  $m$ . Hence, we often do not explicitly give the ground set  $X$ , as just the  $m$  can be used to precisely capture the containment of an element in the multiset. For a multiset  $A$ , we denote its multiplicity function as  $m_A$ . We say, “ $A$  is a multiset with support in  $X$ ,” to mean that the support of  $m_A$ , is a subset of  $X$ . Finally, some standard operations on multisets are defined as follows: containment as  $x \in A \Leftrightarrow m_A(x) > 0$ , intersection as  $m_{A \cap B}(x) = \min(m_A(x), m_B(x))$ , union as  $m_{A \cup B}(x) = \max(m_A(x), m_B(x))$ , sum as  $m_{A \uplus B}(x) = m_A(x) + m_B(x)$ , and subtraction as  $m_{A \setminus B}(x) = \max(m_A(x) - m_B(x), 0)$ .

### B. Model Description

We model the UAV routing problem as a game between the UAVs. Let  $\mathcal{P} = \{1, \dots, p\}$  denote a finite set of  $p$  players, each corresponding to a UAV. The geographical region of surveillance is discretized into a finite set  $\mathcal{C}$  of smaller subregions referred to as the cells. The information collection environment is then represented as a directed graph, where the cells are the vertices, and the directed edges of the graph capture the connectivity between these cells. The time is also assumed to be discretized into *time steps*. The division of cells

and their connectivity is assumed to be such that moving along any edge and surveilling the subsequent cell, altogether takes an equal time—of 1 time step. The number of time steps for which the game lasts is denoted by  $l$ . Consequently, the goal for each player is to move in this network for 1 time step, while capturing the information from the cells visited along the route, with the goal of maximizing this information captured. Thus, the set of strategies for player  $i$ , denoted by  $\mathcal{S}_i$  is a set of walks of length  $l$  starting from player  $i$ 's initial cell. The set of “joint strategy profiles,” or simply “outcomes,” is denoted by  $\mathcal{S} = \times_{i \in \mathcal{P}} \mathcal{S}_i$ . Each cell has an associated information value denoted by a function  $v : \mathcal{C} \rightarrow \mathbb{R}^+$ . Thus,  $v(c)$  denotes the information initially available in cell  $c$ .

The UAVs have sensors through which they collect the available information—better the quality of sensors, greater is the fraction of information they can collect from what is available. Consequently, each player  $i \in \mathcal{P}$  is assigned a sensor effectiveness parameter denoted by  $\rho_i \in (0, 1]$ , which determines how much information the player can collect from what is available in the cell it visits. Finally, the payoff of  $i$ , which is simply the collected information, depends on the outcome, and is denoted by  $\pi_i : \mathcal{S} \rightarrow \mathbb{R}^+$ .

### III. SINGLE-STEP GAMES

In this section, we study *single-step* games where the number of time steps is  $l = 1$ . These games have some nice properties that the general multistep games do not. They have a simple structure which allows us to incorporate an important facet—that of *information fusion*. The strategies of a player are simply the cells adjacent to its initial position. We define the payoffs as follows: as aforementioned, the parameter  $\rho_i$  denotes the fraction of information a player can collect from what is available. Hence, player  $i$  on visiting cell  $c$  alone, gets  $\rho_i v(c)$ , leaving  $(1 - \rho_i)v(c)$  amount of information in the cell. In other words, the information value of a cell depletes by a factor of  $(1 - \rho_i)$  after  $i$ 's visit. Consequently, if a set of players  $\mathcal{Q} \subseteq \mathcal{P}$  simultaneously visits  $c \in \mathcal{C}$ , the information available depletes by a factor of  $\prod_{i \in \mathcal{Q}} (1 - \rho_i)$ —accordingly, the total information collected from cell  $c$  by players in  $\mathcal{Q}$  is  $(1 - \prod_{i \in \mathcal{Q}} (1 - \rho_i))v(c)$  which we define the aggregate payoff of the players to be. Furthermore, each player  $i \in \mathcal{Q}$  is said to get as an individual payoff, a share of the aggregate payoff that is proportional to its  $\rho_i$ . Thus, this equals

$$\frac{\rho_i}{\sum_{i \in \mathcal{Q}} \rho_i} \left( 1 - \prod_{i \in \mathcal{Q}} (1 - \rho_i) \right) v(c). \quad (2)$$

Next, we extend the model to incorporate information fusion.

**Information Fusion:** In case of fusion, a combination of multiple sources of information can be utilized for a greater quality of information. Consequently, UAVs can obtain a more refined information from a cell by visiting it simultaneously (via fusion gain) than by visiting singly, in which case, they are awarded a greater payoff for the improved quality. We model this by introducing fusion parameters  $\gamma_1, \dots, \gamma_p$ , where  $\gamma_j$  captures the enhancement factor in the information quality as well as in the resultant payoffs when any  $j$  players visit a cell simultaneously. We assume that these parameters are

specified *a priori* as input to the problem depending on the complementarity and similarity of the UAVs functionalities. By definition, we have  $\gamma_1 = 1$  since there is no information fusion with just one UAV, and  $\gamma_1 \leq \dots \leq \gamma_p$ , since more UAVs participating in the information fusion should not decrease the fusion gain factor. The case of no information fusion is simply when  $\gamma_1 = \dots = \gamma_p = 1$ , and the payoffs are given by the expression obtained in (2). As the payoffs get improved by the appropriate fusion gain factors, the payoff of  $i \in \mathcal{P}$  on visiting  $c \in \mathcal{C}$  simultaneously with a set of players  $\mathcal{Q} \subseteq \mathcal{P}$ , equals

$$\gamma_{i|\mathcal{Q}} \frac{\rho_i}{\sum_{i \in \mathcal{Q}} \rho_i} \left( 1 - \prod_{i \in \mathcal{Q}} (1 - \rho_i) \right) v(c). \quad (3)$$

Note that the payoff function in (3) can be written in a general form of  $U_i(M_c(\mathcal{Q}))$ , where  $M_c : 2^{\mathcal{P}} \rightarrow \mathbb{R}^+$  for each  $c \in \mathcal{C}$  is given by  $M_c(\mathcal{Q}) = \gamma_{i|\mathcal{Q}} (1 - \prod_{i \in \mathcal{Q}} (1 - \rho_i)) / (\sum_{i \in \mathcal{Q}} \rho_i v(c))$ , and  $U_i : \mathbb{R}^+ \rightarrow \mathbb{R}^+$  by  $U_i(x) = \rho_i x$ . Consequently, the single-step UAV game can be represented as a tuple  $(\mathcal{P}, \mathcal{C}, (S_i)_{i \in \mathcal{P}}, (M_c)_{c \in \mathcal{C}}, (U_i)_{i \in \mathcal{P}})$  with all the entities defined as earlier and, thus, is a singleton congestion-type game.

We now characterize a class of singleton congestion-type games which this game lies in and show an important result regarding the existence of PNE for this class of games.

**Definition 6:** A singleton congestion-type game  $(\mathcal{P}, \mathcal{C}, (S_i)_{i \in \mathcal{P}}, (M_c)_{c \in \mathcal{C}}, (U_i)_{i \in \mathcal{P}})$  is called well-behaved if:

- 1)  $\forall i$ ,  $U_i$  is monotonically nondecreasing;
- 2)  $(M_c)_{c \in \mathcal{C}}$  are either all monotonically nondecreasing, or all monotonically nonincreasing.<sup>1</sup>

The first condition simply says that if the value of a resource increases, a player's utility for using it should not decrease. The second condition says that the effect of congestion in a resource on its value is always in the same direction, for all the resources—either adding more players to any resource never increases its value, or it never decreases. We need the following lemma for the containment of single-step UAV game in the class of well-behaved singleton congestion-type games.

**Lemma 1:** Let  $0 < y \leq 1$ , and let  $X$  be a set of  $n(\geq 0)$  numbers such that  $\forall x \in X, 0 < x \leq 1$ . Then

$$\prod_{x \in X} (1 - x) \leq \frac{1 - (1 - y) \prod_{x \in X} (1 - x)}{y + \sum_{x \in X} x} \leq \frac{1 - \prod_{x \in X} (1 - x)}{\sum_{x \in X} x}.$$

*Proof:* Let  $S = \sum_{x \in X} x$ , and  $P = \prod_{x \in X} (1 - x)$ . We want to show  $P \leq 1 - (1 - y)P/y + S \leq 1 - P/S$ . Consider the first inequality

$$\begin{aligned} P \leq \frac{1 - (1 - y)P}{y + S} &\Leftrightarrow (y + S)P \leq 1 - (1 - y)P \\ &\Leftrightarrow (1 + S)P \leq 1. \end{aligned}$$

Next, consider the second inequality

$$\begin{aligned} \frac{1 - (1 - y)P}{y + S} &\leq \frac{1 - P}{S} \Leftrightarrow (1 - P + Py)S \leq y + S(1 - P) \\ &\Leftrightarrow yPS \leq y(1 - P) \Leftrightarrow (1 + S)P \leq 1. \end{aligned} \quad (4)$$

<sup>1</sup> $M_c$  is monotonically nondecreasing, if  $\forall \mathcal{Q}, \mathcal{Q}' \subseteq \mathcal{P}, \mathcal{Q} \subseteq \mathcal{Q}' \Rightarrow M_c(\mathcal{Q}) \leq M_c(\mathcal{Q}')$ . (Vice-versa for nonincreasing.)

Thus, proving (4) proves both the inequalities. Now, by the inequality of arithmetic and geometric means, we have  $(n - S)/n \geq P^{1/n}$  and hence

$$(1 - S/n)^n \geq P. \quad (5)$$

Also, using the binomial theorem

$$(1 + S/n)^n = \sum_{i=0}^n \binom{n}{i} (S/n)^i \geq 1 + S. \quad (6)$$

Hence, combining (5) and (6), we get  $(1 + S)P \leq (1 + S/n)^n (1 - S/n)^n = (1 - S^2/n^2)^n \leq 1$ . Thus, this proves (4) as required.  $\square$

**Lemma 2:** Single-step games with no fusion are well behaved.

*Proof:* Since  $U_i(x) = \rho_i x$  for each player  $i$  and, thus, is monotonically increasing, it satisfies the first condition for being well behaved. Now, when there is no fusion,  $M_c(\mathcal{Q}) = (1 - \prod_{i \in \mathcal{Q}} (1 - \rho_i)) / \sum_{i \in \mathcal{Q}} \rho_i v(c)$  for each cell  $c$ . As  $v(c)$  is a constant, next, we show that  $M_c(\mathcal{Q})/v(c)$  is monotonically nonincreasing.

Let  $\mathcal{Q} \subset \mathcal{Q}' \subseteq \mathcal{P}$ . Suppose  $\mathcal{Q}' \setminus \mathcal{Q}$  has  $k$  players, without loss of generality, say players  $1, \dots, k$ . Next, define  $\mathcal{Q}_0, \mathcal{Q}_1, \dots, \mathcal{Q}_k$ , with  $\mathcal{Q}_0 = \mathcal{Q}$ , and  $\mathcal{Q}_i = \mathcal{Q}_{i-1} \cup \{i\}, \forall i \in [k]$ , so that,  $\mathcal{Q}_k = \mathcal{Q}'$ . Then, using the second inequality from Lemma 1, we have

$$\begin{aligned} \forall i \in [k], \quad \frac{1 - \prod_{j \in \mathcal{Q}_i} (1 - \rho_j)}{\sum_{j \in \mathcal{Q}_i} \rho_j} &= \frac{1 - (1 - \rho_i) \prod_{j \in \mathcal{Q}_{i-1}} (1 - \rho_j)}{\rho_i + \sum_{j \in \mathcal{Q}_{i-1}} \rho_j} \\ &\leq \frac{1 - \prod_{j \in \mathcal{Q}_{i-1}} (1 - \rho_j)}{\sum_{j \in \mathcal{Q}_{i-1}} \rho_j}. \end{aligned}$$

Hence,  $M(\mathcal{Q}_0) \geq \dots \geq M(\mathcal{Q}_k)$ . Thus,  $M(\mathcal{Q}) \geq M(\mathcal{Q}')$  whenever  $\mathcal{Q} \subset \mathcal{Q}'$ , which proves the claim.  $\square$

As shown in Lemma 2, when there is no fusion, the value function for each cell is monotonically nonincreasing. The fusion parameters  $\gamma_1, \gamma_2, \dots, \gamma_p$  are nondecreasing and, thus, need not preserve the monotonicity of the value functions. We next consider two interesting special cases.

**Definition 7:** Given the players and the sensor effectiveness parameters, we say that information fusion is mild if the fusion parameters are gradually increasing so that all the value functions are monotonically nonincreasing. On the other hand, we say that it is strong if the fusion parameters are so rapidly increasing that all the value functions become monotonically nondecreasing.

By definition, single-step UAV games in both the above-mentioned cases are well-behaved singleton congestion-type games.

#### A. Existence of a Pure Nash Equilibrium

We prove the existence of a PNE using the *finite improvement property* [42] defined as follows.

**Definition 8:** Finite Improvement Property (FIP): A sequence of strategy-tuples in which each tuple differs from the preceding one in one coordinate (such a sequence is called a

path), and the unique deviator in each step strictly increases its payoff (an improvement path), is finite. Clearly, any maximal improvement path is terminated by an equilibrium.

**Theorem 1:** Every well-behaved singleton congestion-type game admits the FIP and, consequently, has a PNE.

*Proof:* We extend the argument of [35], which is used there for proving the FIP for symmetric congestion games. Suppose, for a contradiction, that there is an infinite improvement path. Since there are only finitely many joint strategies, there is an improvement cycle, say, of size  $k$ , given by  $\sigma_1, \sigma_2, \dots, \sigma_k, \sigma_1$ , where each  $\sigma_j \in \mathcal{S}$  is the outcome in the  $j$ th step. Furthermore, let  $\mathcal{Q}_j(c)$  denote the set of players going to cell  $c$  in the  $j$ th step of the improvement cycle. Let  $\mathcal{C}^\# \subseteq \mathcal{C}$  denote the set  $\{c \mid \exists i, j, \mathcal{Q}_i(c) \neq \mathcal{Q}_j(c)\}$ , i.e., those cells which are not occupied by the same set of players throughout the whole improvement cycle. For a well-behaved game, we have that  $U_i$  is monotonically nonincreasing for each  $i \in \mathcal{P}$ , and, we have two possibilities for the value functions  $(M_c)_{c \in \mathcal{C}}$ . First, we prove the result for the case where  $(M_c)_{c \in \mathcal{C}}$  are all monotonically nonincreasing.

Without loss of generality, suppose the improvement cycle and the cells are enumerated such that  $\min_{1 \leq j \leq k, c \in \mathcal{C}^\#} M_c(\mathcal{Q}_j(c)) = M_{c_1}(\mathcal{Q}_k(c_1))$ . Now, since  $c_1 \in \mathcal{C}^\#$ , there exists  $j$  such that  $\mathcal{Q}_j(c_1) \neq \mathcal{Q}_k(c_1)$ . Consider the largest such  $j$ , i.e.,  $\mathcal{Q}_{j+1}(c_1) = \mathcal{Q}_k(c_1)$ . Since each  $M_c$  is monotonically nonincreasing, and since the minimum value of  $M_{c_1}(\cdot)$  is attained for  $\mathcal{Q}_k(c_1)$ , thus, also for  $\mathcal{Q}_{j+1}(c_1)$ , it follows that  $\mathcal{Q}_j(c_1) \subset \mathcal{Q}_{j+1}(c_1)$ . Thus, the unique deviator between  $\sigma_j$  and  $\sigma_{j+1}$ , without loss of generality, say player 1, must be changing its strategy to  $c_1$  from some other cell  $c_i$ , say. Thus,  $c_i \in \mathcal{C}^\#$ , and further, for this deviation to be an improvement for player 1, it must be that  $U_1(M_{c_i}(\mathcal{Q}_j(c_i))) < U_1(M_{c_1}(\mathcal{Q}_{j+1}(c_1)))$ . Since each  $U_i$  is monotonically nondecreasing, hence,  $M_{c_i}(\mathcal{Q}_j(c_i)) < M_{c_1}(\mathcal{Q}_{j+1}(c_1)) = M_{c_1}(\mathcal{Q}_k(c_1))$ . This contradicts the assumption that  $\min_{1 \leq j \leq k, c \in \mathcal{C}^\#} M_c(\mathcal{Q}_j(c)) = M_{c_1}(\mathcal{Q}_k(c_1))$ , and hence, there cannot exist an improvement cycle.

The case where  $(M_c)_{c \in \mathcal{C}}$  are all monotonically nondecreasing can be shown similarly.  $\square$

Since the single-step game in the cases of mild or strong fusion is well behaved, it follows the following corollary.

**Corollary 1:** The single-step game in the cases of mild or strong information fusion has a PNE.

Next, we discuss the time complexity of NE computation.

### B. Equilibrium Computational Complexity

Computing a NE in a two-player game is, in general, polynomial parity arguments on directed graphs (PPAD) complete [43]. For the single-step UAV game, however, the existence of FIP trivially implies that the problem of computing a PNE is in the class polynomial local search (PLS) [44]—as the problem can be reduced to finding a sink in a directed acyclic graph formed over the outcomes (as vertices) with the directed edges capturing the unilateral improvement deviations. This also gives a finite time algorithm for computing a PNE in the single-step UAV game. However, for the case of strong information fusion, we give an efficient algorithm based on a greedy strategy, whereas it remains to

be seen whether the problem can be computed efficiently for the mild information fusion case, or, is PLS-complete—like it is for the closely related class of congestion games.

**Efficient Algorithm for Strong Information Fusion:** In this case, the value function for each cell is monotonically nondecreasing, i.e., more the number of players in a cell, larger the value of the cell and, in turn, the individual payoff for each player there as well. Let  $\mathcal{Q}(c) \subseteq \mathcal{P}$  denote the set of players which can visit a cell  $c$  in one step as a possible strategy, i.e.,  $\mathcal{Q}(c) = \{i \in \mathcal{P} \mid c \in \mathcal{S}_i\}$ .

The algorithm consists of a number of iterations. Starting with the set of all players and the set of all cells, in each iteration, some players are assigned a particular cell as their strategy to play and the set of remaining players and remaining cells are carried forward to the next iteration.

---

#### Algorithm 1 Algorithm to Compute a PNE

---

```

1:  $\mathcal{P}^1 \leftarrow \mathcal{P}, \quad \mathcal{C}^1 \leftarrow \mathcal{C}, \quad i \leftarrow 1$ 
2: while  $\mathcal{P}^i \neq \emptyset$  do ▷ Terminate if no players remaining
3:    $c_i^{\max} \leftarrow \text{Null}, \quad \text{maxScore} \leftarrow -\infty$ 
4:   for  $c \in \mathcal{C}^i$  do
5:      $\mathcal{Q}^i(c) = \mathcal{Q}(c) \cap \mathcal{P}^i, \quad \text{score} \leftarrow M_c(\mathcal{Q}^i(c))$ 
6:     if  $\text{score} > \text{maxScore}$  then
7:        $\text{maxScore} \leftarrow \text{score}, \quad c_i^{\max} \leftarrow c$ 
8:     end if
9:   end for ▷  $c_i^{\max}$  is computed.
10:  for  $p \in \mathcal{Q}^i(c_i^{\max})$  do
11:    Fix cell  $c_i^{\max}$  as strategy for  $p$ 
12:  end for
13:   $\mathcal{P}^{i+1} = \mathcal{P}^i \setminus \mathcal{Q}^i(c_i^{\max}), \quad \mathcal{C}^{i+1} = \mathcal{C}^i \setminus \{c_i^{\max}\}, \quad i \leftarrow i + 1$ 
14: end while
```

---

In each iteration  $i$ , for each cell  $c$  in  $\mathcal{C}^i$ , we compute  $M_c(\mathcal{Q}(c) \cap \mathcal{P}^i)$ , where  $\mathcal{P}^i$  is the set of remaining players in that iteration, and  $\mathcal{C}^i$  is the set of remaining cells. Then, we choose the cell  $c_i^{\max}$  for which the value thus computed is maximum and assign this cell as the strategy for all the players in  $\mathcal{Q}(c_i^{\max}) \cap \mathcal{P}^i$ . Subsequently, we update the set of remaining players  $\mathcal{P}^{i+1} = \mathcal{P}^i \setminus (\mathcal{Q}(c_i^{\max}) \cap \mathcal{P}^i)$  and  $\mathcal{C}^{i+1} = \mathcal{C}^i \setminus \{c_i^{\max}\}$  and move on to the next iteration.

**Theorem 2:** Algorithm 1 computes a PNE of the single-step game with strong information fusion in  $O(|\mathcal{C}| \cdot |\mathcal{P}|)$  time.

*Proof:* Without loss of generality, let the cells be enumerated such that  $\forall i, c_i^{\max} = c_i$ . We first note that  $i < j \Leftrightarrow \mathcal{P}^j \subset \mathcal{P}^i \Rightarrow \mathcal{Q}^j(c_j) \subseteq \mathcal{Q}^i(c_j)$ . Suppose player 1 is assigned the cell  $c_i$  by the algorithm. Hence,  $1 \notin \mathcal{Q}(c_j)$ ,  $\forall j < i$ , since otherwise, it would have been assigned a cell before the  $i$ th iteration. Thus, the only cells player 1 could possibly deviate to are  $\{c_j\}_{j>i}$ . The payoff of player 1 before deviation, by playing  $c_i$ , is  $\rho_1 M_{c_i}(\mathcal{Q}^i(c_i))$ , which is at least  $\rho_1 M_{c_j}(\mathcal{Q}^i(c_j)) \forall j > i$ , since  $c_i^{\max} = c_i$ . Furthermore,  $\forall j > i, 1 \in \mathcal{Q}(c_j) \Rightarrow 1 \in \mathcal{Q}^i(c_j)$ . This further implies that since  $\mathcal{Q}^j(c_j) \subseteq \mathcal{Q}^i(c_j)$ , it must also be that,  $(\mathcal{Q}^j(c_j) \cup \{1\}) \subseteq \mathcal{Q}^i(c_j)$ , whenever player 1 can access  $c_j$ . Hence, by the monotonically nondecreasing behavior of each  $M_c$ , player 1's payoff after deviation to  $c_j$ , i.e.,  $U_1(M_{c_j}(\mathcal{Q}^j(c_j) \cup \{1\}))$ , is at most  $U_1(M_{c_j}(\mathcal{Q}^i(c_j)))$ , which is at most its payoff before



deviation. Hence, it has no incentive to switch, and the same argument applies to each player. Hence, the algorithm does produce a NE. It is easy to check that the algorithm takes time  $O(|C| \cdot |\mathcal{P}|)$ .  $\square$

### C. PoS and PoA Bounds

Since single-step games are special cases of multistep games which we analyze in Section IV, the upper bound of 2 on the PoA and PoS proved there applies here as well and can be shown to be tight. In this section, we establish a stronger bound for the special case of *homogeneous fleet*, i.e., all players are identical, and there is either no fusion or mild fusion.

When all UAVs have same  $\rho$ , the payoff of a player simply depends on the number of players it shares a cell with and not the actual subset. This makes it an unweighted singleton congestion game. While results on the PoA bound for these games have been shown for various classes of payoff (or cost) functions, such as affine, polynomial, etc., to the best of our knowledge, the payoff function of the UAV game herein does not fall in any of the previously studied classes. Kleinberg and Oren [40] study a class of games called the *Project Game*. The scenario therein with identical players very closely resembles the setting studied here and for no information fusion, the following result can be derived from the corresponding results for the Project Game. However, our tight bounds shown below are for the more general case of mild information fusion which does not follow from [40]. The main result is in the following.

**Theorem 3:** *The PoS and PoA, in unweighted singleton congestion games with either no fusion or mild fusion, are at most  $2 - 1/p$ . Furthermore, these bounds are tight.*

**Proof:** We denote the individual payoff of a player when  $n$  players share a cell  $c$ , by  $v_n(c)$ , given by  $\gamma_n v(c) 1 - (1 - \rho)^n / n$ . Let  $\sigma^e$  be an equilibrium and let  $\sigma^m$  be a joint strategy which provides the maximum social welfare. Suppose, starting with  $\sigma^e$ ,  $\sigma^m$  is achieved by a series of deviations, where each deviation refers to a player switching from a cell  $c_i$  to a cell  $c_j$ . Since the players are identical, only the cells involved in a deviation matter, and not the player who deviates. We represent this as a *deviation graph*  $G$ , where each cell is a vertex and a deviation from one cell to another is represented as a directed edge. Note that since only the number of players in a cell matters in computing any payoffs, any path in the graph of length more than 1, say between nodes  $u$  and  $v$ , can be replaced by a single edge  $(u, v)$ , since both equivalently result in the number of players at  $u$  decreasing by 1, the number of players at  $v$  increasing by 1, and other cells on the path being unaffected. Thus,  $G$  can be reduced to, say  $G^*$ , that does not have a cycle, nor a path of length more than 1. Thus,  $G^*$  only has sources, sinks, and isolated vertices. Fig. 1 illustrates this with an example. Vertices such as  $c_3$  with a larger in-degree than out-degree in  $G$  become sinks in  $G^*$ . Similarly, vertices like  $c_1$  having a larger out-degree in  $G$  become sources in  $G^*$ , and the remaining ones like  $c_7$  where the in- and out-degrees are equal in  $G$ , become isolated in  $G^*$ .

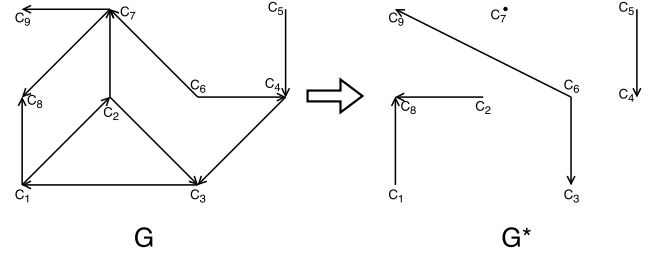


Fig. 1. Reduction of a deviation graph.

Now, consider the group of players who are in a cell  $c$  at equilibrium. We consider the following three cases.

- 1)  $c$  is an isolated vertex in  $G^*$ : The payoff of every player here remains the same.
- 2)  $c$  is a sink in  $G^*$ : There are at least as many players in  $c$  at  $\sigma^m$ , as there were at the equilibrium  $\sigma^e$ . Hence, the payoff of these players is bounded above by their payoff at the equilibrium.
- 3)  $c$  is a source in  $G^*$ : Consider a player  $i$  who is in cell  $c$  at  $\sigma^e$  and is in  $c'$  at  $\sigma^m$ . By the nature of  $G^*$ ,  $c'$  must be a sink in  $G^*$ . Hence, there are strictly more players in  $c'$  at  $\sigma^m$ , than there were at  $\sigma^e$ . Thus, since we have mild fusion or no fusion, if  $i$  were to be the unique deviator at equilibrium from  $c$  to  $c'$ , it would have obtained at least as much a payoff as it would have by playing  $c'$  at  $\sigma^m$ . Furthermore, since  $\sigma^e$  is an equilibrium, the payoff of player  $i$  at  $\sigma^e$  is at least as much as it would get by deviating to any other cell, in particular,  $c'$ , and in turn, greater than its payoff at  $\sigma^m$ . Now, suppose there were  $x$  players in cell  $c$  at  $\sigma^e$ , of which  $y$  are not in  $c$  at  $\sigma^m$ , while  $x - y$  players continue to be in  $c$  at  $\sigma^m$ . Then, as per the argument earlier, the payoff at  $\sigma^m$ , of each of the  $y$  deviating players, is at most as much as their payoff at  $\sigma^e$ , i.e.,  $v_x(c)$ . The payoff of any of the remaining  $x - y$  players improves from  $v_x(c)$  to  $v_{x-y}(c)$ . Hence, if all the players deviate from  $c$ , i.e.,  $y = x$ , the total welfare of this group of  $x$  players cannot increase. On the other hand, if  $y < x$ , the total social welfare for this group of players can increase by a factor of at most  $(x - y)v_{x-y}(c) + yv_x(c) / xv_x(c) = \gamma_{x-y}(1 - (1 - \rho)^{x-y}) / \gamma_x(1 - (1 - \rho)^x) + y/x$ . It can be shown that this expression is a monotonically increasing function of  $\rho$  for  $\rho \in (0, 1]$  (by showing its derivative with respect to  $\rho$  to be strictly positive for  $\rho \in (0, 1]$  when  $y < x$ ). Hence, its maximum value is when  $\rho = 1$ , which comes out to  $\gamma_{x-y}/\gamma_x + y/x$ . Using  $y < x \leq p$ , we have  $\gamma_{x-y}/\gamma_x \leq 1$ , and  $y/x \leq p - 1/p$ , giving a bound of  $2 - 1/p$  on the welfare gain ratio (which can occur at  $y = p - 1$ ,  $x = p$ , that is, when all players visit the same cell at equilibrium, and only one of them visits the cell to obtain the maximum social welfare).

Thus, for a group of players which are in a particular cell at equilibrium, the sum of their payoffs either remains the same, decreases, or increases by a factor of at most  $2 - 1/p$  as analyzed for the three cases mentioned above. Hence, the total social welfare of all the players, which is the sum of welfare of all such groups, can increase by a factor of at most as

much as any of the individual groups, which is nothing but  $2 - 1/p$ , giving us the required bound. (Note that the analysis mentioned above holds for any equilibrium, and thus, the worst equilibrium, in particular, giving the bound on PoA.)

Furthermore, this bound can be shown to be tight not only for PoA but also PoS, with the following example. Let there be  $p$  players with  $\rho = 1$  for each player, and no information fusion. Let there be  $p$  cells with every cell being a valid strategy for every player. Let the information available in various cells be as follows:  $v(c_1) = p, v(c) = 1 - \epsilon \forall c \neq c_1; \epsilon > 0$ . Clearly,  $c_1$  is a dominant strategy for every player, giving a unique equilibrium  $(c_1, \dots, c_1)$ . The total welfare here is  $p$ . However, it is maximum for the joint strategy  $(c_1, c_2, \dots, c_p)$  which equals  $p + (p - 1)(1 - \epsilon)$ . Hence, the PoS is  $1 + (1 - 1/p)(1 - \epsilon)$  which approaches  $2 - 1/p$  as  $\epsilon \rightarrow 0$ .  $\square$

#### IV. MULTISTEP GAMES

In this section, we extend our analysis from single step to multiple steps. Recall the notation from Table II. The payoff of  $i$  depends on the outcome and is denoted by  $\pi_i : \mathcal{S} \rightarrow \mathbb{R}^+$ . It is the sum of the partial payoffs  $i$  gets by visiting each cell on its walk. With a slight abuse of notation, we let  $\pi_i(s, c)$  denote  $i$ 's payoff for visiting  $c \in \mathcal{C}$ , when the outcome is  $s \in \mathcal{S}$ , so that  $\pi_i(s) = \sum_{c \in \mathcal{C}} \pi_i(s, c)$ . Naturally,  $\pi_i(s, c)$  is zero if  $i$  does not visit  $c$  at all when playing  $s_i$ . However, when it does visit the cell (possibly more than once), the value  $\pi_i(s, c)$  can be defined in two different ways depending on the logistics of the real-world scenario, giving rise to two different games—*temporal* and *nontemporal*.

##### A. With Temporal Aspect

In this case, a player gets an instant payoff after visiting a cell (in a manner described below), and these payoffs get accumulated constituting its net payoff. Any player  $i$ , on visiting a cell  $c$ , gets a payoff that is  $\rho_i$  fraction of the value left in  $c$  at the time of its visit, leaving behind  $(1 - \rho_i)$  fraction of that value. Thus, if a sequence of  $k$  players say  $(x_1, x_2, \dots, x_k)$  visit  $c$  one after the other, then the  $i$ th visitor  $x_i$  gets a payoff of  $\rho_{x_i} \left( \prod_{j < i} (1 - \rho_{x_j}) \right) v(c)$  corresponding to that visit; if the same player is also the  $j$ th visitor for some  $j \neq i$ , it will get a payoff for each such visit defined similarly. The combined payoff of all these players, from visiting  $c$  is

$$\left( 1 - \prod_{j \leq k} (1 - \rho_{x_j}) \right) v(c). \quad (7)$$

Note that this combined payoff is independent of the order of the players. Thereby, if these players visit  $c$ , in the same time step, then we define their aggregate payoff as above, and further, the payoff of  $x_i$  as the share of this combined payoff is proportional to  $\rho_{x_i}$ , that is,

$$\frac{\rho_{x_i}}{\sum_{j \leq k} \rho_{x_j}} \left( 1 - \prod_{j \leq k} (1 - \rho_{x_j}) \right) v(c). \quad (8)$$

Thus, a player's payoff from visiting a cell depends on which players visit before it and which players visit simultaneously.

##### B. Without Temporal Aspect

In this case, the payoff from visiting a cell is determined at the end of the game, regardless of the order in which the players visit the cell. Since the order is immaterial, we can represent the visitors of a cell  $c$  as a multiset, say  $\mathcal{P}'$ , having support in  $\mathcal{P}$  and an associated multiplicity function denoted by  $m_{\mathcal{P}'}(\cdot)$ . In case of no ambiguity, we drop the subscript and denote it as simply  $m(\cdot)$ . The payoff of a visitor from a single visit is precisely as in (8), and thus, with possibly multiple visits, the payoff of  $i \in \mathcal{P}'$  is given by

$$\frac{\rho_i m(i)}{\sum_{j \in \mathcal{P}'} \rho_j m(j)} \left( 1 - \prod_{j \in \mathcal{P}'} (1 - \rho_j)^{m(j)} \right) v(c). \quad (9)$$

Next, we establish results for both these games on existence of pure equilibria, PoA bounds, etc.

##### C. Multistep Games With Temporal Aspect

In this section, we analyze the game with a temporal aspect. As discussed earlier, the payoff of a player from visiting a cell is not only merely dependent on which players visit the cell but also on the order in which they visit the cell. The combined payoff, however, when a sequence of players visit a cell  $c$  (some of them possibly simultaneously), does not depend on their order, and can be easily computed as in (8). Let  $A, B$  be multisets with support in  $\mathcal{P}$ . For a cell  $c$ , let  $\pi_B^A(c)$  denote the combined payoff which the visitors in  $B$  would obtain by visiting cell  $c$  (as many times as the respective multiplicities in  $B$ ) when preceded by all (and only) the visitors as represented by  $A$ . Note that the multiset representation is sufficient for this to be well defined, since the order of visitors in  $A$  among themselves, and similarly of those in  $B$  among themselves, does not matter when computing the said *combined* payoff. Indeed, the exact expression can be easily obtained to be

$$\pi_B^A(c) = \left( 1 - \prod_{j \in B} (1 - \rho_j)^{m_B(j)} \right) \prod_{j \in A} (1 - \rho_j)^{m_A(j)} v(c). \quad (10)$$

Here, the entity  $\prod_{j \in A} (1 - \rho_j)^{m_A(j)}$  denotes the fraction of  $v(c)$  left in  $c$  after visitors in  $A$  have visited, and the fraction of it collected by  $B$  is computed similarly.

*Lemma 3:* Let  $A, A', B, B'$ , and  $D$  be multisets with support in  $\mathcal{P}$  such that  $A' \subseteq A$  and  $B' \subseteq B$ . Then, (1)  $\pi_B^A(c) + \pi_D^{A \uplus B}(c) = \pi_{B \uplus D}^A(c)$ , (2)  $\pi_{B'}^A(c) \leq \pi_B^A(c)$ , and (3)  $\pi_B^A(c) \leq \pi_{B'}^{A'}(c)$ .

*Proof:* For the first part, note that  $B \uplus D$  is the multiset sum of  $B$  and  $D$  and thus represents the combined visitors in  $B$  as well as  $D$ . Thus, the result follows from the definition of  $\pi_B^A(c)$ , as both sides equal the combined payoff of visitors in  $B$  and  $D$  when they are preceded by visitors in  $A$ . For the second part, we apply the first part on  $A, B', B \setminus B'$ , we get  $\pi_B^A(c) - \pi_{B'}^A(c) = \pi_{B \setminus B'}^{A \uplus B'}(c) \geq 0$ . Rearranging gives the required result. For the third part, using (10), it is easy to see that  $\pi_B^A(c) / \pi_{B'}^{A'}(c) = \prod_{j \in A \setminus A'} (1 - \rho_j)^{m_A(j) - m_{A'}(j)} \leq 1$ . Rearranging gives the required result.  $\square$



Next, suppose  $c$  is a cell, and  $A$  is a multiset with support in  $\mathcal{P}$ . For each player  $j$ , let  $S'_j \subseteq S_j$  denote the subset of strategies in which  $j$  visits  $c$  exactly  $m_A(j)$  times. Then,  $S' = \times_j S'_j$  is the set of outcomes for which the multiset  $A$  precisely captures which players visit cell  $c$  and how often. In case of such an outcome, we refer to  $A$  as the *visitor set* for  $c$ . Also, for any multiset  $A$  and a player  $i \in A$ , let  $A|_i$  denote the multiset  $m_A(i) \otimes \{i\}$ , i.e.,  $A|_i$  only contains  $i$ —with the same multiplicity as  $A$ —and let  $A|_{-i}$  denote the multiset  $A \setminus A|_i$ . The following lemma shows an important result.

**Lemma 4:** *Let  $A$  be any multiset with support in  $\mathcal{P}$ ,  $i \in A$ ,  $c$  be any cell, and  $S' \subseteq S$  be the set of outcomes for which  $A$  is the visitor set for  $c$ . Then,  $\forall s \in S'$ ,  $\pi_i(s, c) \geq \pi_{A|_i}^{A|_{-i}}(c)$ .*

*Proof:* Let  $s$  be an outcome with the visitor set for  $c$  being  $A$ . Suppose player  $i$  visits  $c$   $m$  times when the visitor set is  $A$ . The outcome  $s$  can be naturally associated with two well-defined sequences  $X_1, \dots, X_m$  and  $Y_1, \dots, Y_m$  as follows. For each  $j$ ,  $X_j$  denotes the set of players visiting  $c$  in the same time step as the  $j$ th visit of  $i$ , and  $Y_j$  denotes the multiset of all the visitors visiting strictly before. Naturally, each  $X_j$  must contain  $i$ , and each  $Y_j$  must contain  $i$  with a multiplicity of  $j - 1$ . Also, we have  $\emptyset \subseteq Y_1 \subset \dots \subset Y_m \subseteq A \setminus \{i\}$  by definition. Let  $v$  be the information initially available in  $c$ . The combined payoff of the visitors in  $X_j$ , as per our notation, is  $\pi_{X_j}^{Y_j}(c)$ . Player  $i$  gets a share of it proportional to  $\rho_i$ , which summed over all visits gives its total payoff from visiting  $c$

$$\pi_i(s, c) = \sum_{j=1}^m \frac{\rho_i}{\sum_{k \in X_j} \rho_k} \pi_{X_j}^{Y_j}(c). \quad (11)$$

For each  $j$ , let  $v_j$  denote the information available in  $c$  just before the visitors in  $X_j$  visit, which evaluates to  $v_j = v \prod_{k \in Y_j} (1 - \rho_k)^{m_{Y_j}(k)}$ . Then, in terms of  $v_j$ , we get  $\pi_{X_j}^{Y_j}(c) = v_j (1 - \prod_{k \in X_j} (1 - \rho_k))$ . Hence, (11) can be written as

$$\pi_i(s, c) = \sum_{j=1}^m \rho_i v_j \frac{(1 - \prod_{k \in X_j} (1 - \rho_k))}{\sum_{k \in X_j} \rho_k}. \quad (12)$$

Now, let  $s'$  be another outcome with similarly defined sequences  $X'_1, \dots, X'_m$  and  $Y'_1, \dots, Y'_m$  such that the only difference from  $s$  is that for each  $j$ , the players visiting  $c$  in the same time step as the  $j$ th visit of player  $i$  in the outcome  $s$ , now visit strictly before it, in  $s'$ . Formally, for each  $j$ , we have  $Y'_j = Y_j \uplus (X_j \setminus \{i\})$  and  $X'_j = \{i\}$ . Thus, the visitor set remains  $A$  for  $s'$ . Now, by definition,  $v_j$  is the information available in  $c$  after the visitors in  $Y_j$  have visited it. Hence, the information left after players in  $X_j \setminus \{i\}$  subsequently visit it is  $v_j \prod_{k \in X_j \setminus \{i\}} (1 - \rho_k)$ . Hence, as player  $i$ 's  $j$ th visit to  $c$  follows, it gets a payoff that is  $\rho_i$  fraction of the value available, and this summed over all the visits gives the total payoff of player  $i$  from visiting  $c$ , for the outcome  $s'$

$$\pi_i(s', c) = \sum_{j=1}^m \left( \rho_i v_j \prod_{k \in X_j \setminus \{i\}} (1 - \rho_k) \right). \quad (13)$$

Now, applying the first inequality from Lemma 1 on  $y = \rho_i$ ,  $X = \{\rho_k | k \in X_j \setminus \{i\}\}$ , we get

$$\begin{aligned} \pi_i(s', c) &\leq \sum_{j=1}^m \rho_i v_j \frac{(1 - (1 - \rho_i) \prod_{k \in X_j \setminus \{i\}} (1 - \rho_k))}{\rho_i + \sum_{k \in X_j \setminus \{i\}} \rho_k} \\ &= \sum_{j=1}^m \rho_i v_j \frac{(1 - \prod_{k \in X_j} (1 - \rho_k))}{\sum_{k \in X_j} \rho_k} = \pi_i(s, c). \end{aligned} \quad (14)$$

Furthermore, let  $s''$  be another outcome with similarly defined sequences  $X''_1, \dots, X''_m$  and  $Y''_1, \dots, Y''_m$  and the visitor set for  $c$  being still  $A$ , such that all the visits of player  $i$  are strictly after all the visits of all the players. Formally, for each  $j$ , we have  $Y''_j = A|_{-i} \uplus ((j-1) \otimes \{i\})$  and  $X''_j = \{i\}$ . Furthermore,  $\pi_i(s'', c) = \pi_{A|_i}^{A|_{-i}}(c)$ . Now, since  $s'$  and  $s''$  are such that all the visits of player  $i$  are unaccompanied, we can write its payoff from the  $j$ th visit as simply  $\pi_{\{i\}}^{Y'_j}(c)$  and  $\pi_{\{i\}}^{Y''_j}(c)$ , respectively. Now, the multiplicity of  $i$  is the same ( $= j - 1$ ) in  $Y'_j$  and  $Y''_j$ , whereas all other players reside in  $Y''_j$  with the maximum multiplicity possible for the visitor set  $A$ . Thus,  $Y'_j \subseteq Y''_j$ , and using Lemma 3, we have  $\pi_{\{i\}}^{Y'_j}(c) \geq \pi_{\{i\}}^{Y''_j}(c)$ ,  $\forall j$ . Hence,

$$\begin{aligned} \sum_{j=1}^m \pi_{\{i\}}^{Y'_j}(c) &\geq \sum_{j=1}^m \pi_{\{i\}}^{Y''_j}(c) \\ \pi_i(s', c) &\geq \sum_{j=1}^m \pi_{\{i\}}^{A|_{-i} \uplus ((j-1) \otimes \{i\})}(c) \\ &\quad \text{( By defn of } Y''_j \text{)} \\ &\geq \sum_{j=1}^m (\pi_{A|_{-i} \uplus (j \otimes \{i\})}^\emptyset(c) - \pi_{A|_{-i} \uplus ((j-1) \otimes \{i\})}^\emptyset(c)) \\ &\quad \text{( Using Lemma 3)} \\ &\geq \pi_{A|_{-i} \uplus (m \otimes \{i\})}^\emptyset(c) - \pi_{A|_{-i}}^\emptyset(c) \\ &\geq \pi_{(m \otimes \{i\})}^{A|_{-i}}(c) \quad \text{( Using Lemma 3)} \\ &\geq \pi_{A|_i}^{A|_{-i}}(c). \end{aligned} \quad (15)$$

Thus, it follows from (14) and (15) that  $\pi_i(s, c) \geq \pi_{A|_i}^{A|_{-i}}(c)$ . Since this holds for any outcome  $s$  for which the visitor set for cell  $c$  is  $A$ , the lemma is proven.  $\square$

Using these, we now show that this game is (1,1)–smooth. Let  $s$  and  $s^*$  be any two outcomes. For every player  $i$ , let  $q^i$  denote the outcome  $(s^*_i, s_{-i})$ . For any cell  $c$ , let multisets  $A_c$  and  $A_c^*$  denote the visitor sets for cell  $c$  when the outcomes are  $s$  and  $s^*$ , respectively. Note that when the outcome is  $q^i$ , the visitor set of cell  $c$  can be written as  $A_c^*|_i \uplus A_c|_{-i}$ . With this notation, we can write

$$\sum_{i \in \mathcal{P}} \pi_i(s) = \sum_{i \in \mathcal{P}} \sum_{c \in \mathcal{C}} \pi_i(s, c) = \sum_{c \in \mathcal{C}} \sum_{i \in \mathcal{P}} \pi_i(s, c) = \sum_{c \in \mathcal{C}} \pi_{A_c}^\emptyset(c). \quad (16)$$

$$\sum_{i \in \mathcal{P}} \pi_i(s^*) = \sum_{c \in \mathcal{C}} \pi_{A_c^*}^\emptyset(c). \quad (17)$$

Finally, for the outcomes  $q^i$ , we can write  $\sum_{i \in \mathcal{P}} \pi_i(q^i) = \sum_{i \in \mathcal{P}} \sum_{c \in \mathcal{C}} \pi_i(q^i, c) = \sum_{c \in \mathcal{C}} \sum_{i \in \mathcal{P}} \pi_i(q^i, c)$ . Now, if player

$i$  does not visit cell  $c$  when playing strategy  $s_i^*$ , equivalently, if it is not contained in the visitor set  $A_c^*$ , its payoff from visiting  $c$  is simply zero. Hence, we get

$$\sum_{i \in \mathcal{P}} \pi_i(q^i) = \sum_{c \in \mathcal{C}} \sum_{i \in A_c^*} \pi_i(q^i, c) \geq \sum_{c \in \mathcal{C}} \sum_{i \in A_c^*} \pi_{A_c^*|i}^{A_c|-i}(c). \quad (18)$$

Here, (18) follows from Lemma 4. Finally, adding (16) and (18) and subtracting (17) gives

$$\begin{aligned} & \sum_{i \in \mathcal{P}} \pi_i(q^i) + \sum_{i \in \mathcal{P}} \pi_i(s) - \sum_{i \in \mathcal{P}} \pi_i(s^*) \\ & \geq \sum_{c \in \mathcal{C}} \left( \pi_{A_c}^\emptyset(c) + \sum_{i \in A_c^*} \pi_{A_c^*|i}^{A_c|-i}(c) - \pi_{A_c^*}^\emptyset(c) \right). \quad (19) \end{aligned}$$

Now, we show that each term of the summation on the right-hand side of (19) is always nonnegative.

**Lemma 5:** *Let  $A$  and  $A^*$  be the visitor sets of a cell  $c \in \mathcal{C}$  for outcomes  $s, s^* \in S$ , respectively. Then*

$$\pi_A^\emptyset(c) + \sum_{i \in A^*} \pi_{A^*|i}^{A|-i}(c) \geq \pi_{A^*}^\emptyset(c). \quad (20)$$

*Proof:* We will show that

$$\pi_A^\emptyset(c) + \sum_{i \in A^* \setminus A} \pi_{A^*|i}^{A|-i}(c) \geq \pi_{A \cup A^*}^\emptyset(c). \quad (21)$$

Recall the definitions related to multisets from Section IV. It is easy to see that the left-hand side of (20) is not less than that of (21) since the latter possibly excludes some terms in the summation, and each term is nonnegative by definition. It is also easy to see that the right-hand side of (20) is no greater than that of (21) by Lemma 3. Thus, it suffices to prove (21) to prove this lemma.

Now, we prove (21) by induction on the number of players in  $A^* \setminus A$ , denoted by, say,  $a$ . The base case  $a = 0$  is when  $A^* \subseteq A$ , i.e.,  $A^* \setminus A = \emptyset$ . This holds trivially, as both the sides of the inequality to be proven, become equal to  $\pi_A^\emptyset(c)$ . Assume, as inductive hypothesis, that (21) holds whenever  $a < a_0$ , for some  $a_0 \in \mathbb{Z}^+$ . Now, consider the case when  $a = a_0$ . Arbitrarily fix some  $x \in A^* \setminus A$ . Then,  $(A^* \setminus A)|_{-x}$ , or equivalently,  $A^*|_{-x} \setminus A$  has  $a_0 - 1$  distinct elements. We get

$$\begin{aligned} & \pi_A^\emptyset(c) + \sum_{i \in A^* \setminus A} \pi_{A^*|i}^{A|-i}(c) \\ & = \pi_A^\emptyset(c) + \sum_{\substack{i \in A^* \setminus A \\ i \neq x}} \pi_{A^*|i}^{A|-i}(c) + \pi_{A^*|x}^{A|-x}(c) \\ & = (\pi_A^\emptyset(c) + \sum_{i \in A^*|_{-x} \setminus A} \pi_{(A^*|_{-x})|i}^{A|-i}(c) + \pi_{A^*|x}^{A|-x}(c)) \\ & \geq \pi_{A \cup (A^*|_{-x})}^\emptyset(c) + \pi_{A^*|x}^{A|-x}(c) \quad (\text{using the Ind. Hyp.}) \\ & \geq \pi_{(A \cup A^*)|_{-x}}^\emptyset(c) + \pi_{A^*|x}^{A|-x}(c) \\ & \quad (\text{since } (A \cup A^*)|_{-x} \subseteq A \cup (A^*|_{-x})) \\ & \quad (\text{and using Lemma 3}) \\ & \geq \pi_{(A \cup A^*)|_{-x}}^\emptyset(c) + \pi_{A^*|x}^{(A \cup A^*)|_{-x}}(c) \\ & \quad (\text{since } (A \cup A^*)|_{-x} \supseteq A|_{-x}) \\ & \quad (\text{and using Lemma 3}) \end{aligned}$$

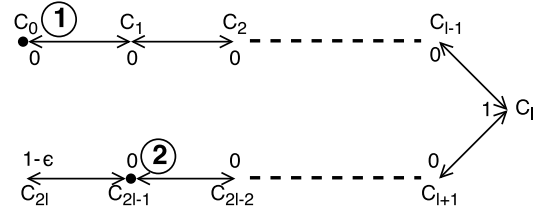


Fig. 2. Two-player game with temporal aspect. The only equilibrium has a welfare of 1, the maximum possible being  $2 - \epsilon$ , showing that the PoA can be arbitrarily close to 2.

$$\begin{aligned} & = \pi_{(A \cup A^*)|_{-x}}^\emptyset(c) + \pi_{(A \cup A^*)|_x}^{(A \cup A^*)|_{-x}}(c) \\ & \quad (\text{since } x \in A^* \setminus A \Rightarrow (A \cup A^*)|_x = A^*|_x) \\ & = \pi_{A \cup A^*}^\emptyset(c). \end{aligned}$$

Hence, this completes the inductive step and the proof by induction for (21), as required.  $\square$

Thus, it follows from (20) and (19) that  $\sum_{i \in \mathcal{P}} \pi_i(q^i) + \sum_{i \in \mathcal{P}} \pi_i(s) \geq \sum_{i \in \mathcal{P}} \pi_i(s^*)$ . Hence, using Definition 3, we have the desired result as follows.

**Theorem 4:** *The multistep game with temporal aspect is  $(1, 1)$ -smooth.*

As shown in [41], a  $(\lambda, \mu)$ -smooth payoff-maximization game has a PoA at most  $1 + \mu/\lambda$ , and this bound applies to the PoA with respect to all equilibrium concepts (mixed, correlated, and not just pure).

**Corollary 2:** *The multistep game with temporal aspect has PoA of at most 2.*

Thus, the game has a constant PoA bound independent of the number of players, number of time steps. We now show with an example, that this bound is tight.

**1) Tight Example:** Consider the game as shown in Fig. 2. Let the number of time steps be  $l(\geq 2)$ . The graph is a simple path as shown with cells  $c_i$  and  $c_{i+1}$  being neighbors of each other for each  $i$ . Initially, the information in all the cells is 0, except for  $c_l$  and  $c_{2l}$ —these two cells have an information of 1 and  $1 - \epsilon$  respectively, where  $\epsilon$  is a small positive constant. There are two players with sensor effectiveness  $\rho_1 = \rho_2 = 1$ , initially in cells  $c_0$  and  $c_{2l-1}$ , respectively. Observe that player 2 cannot grab information from both  $c_l$  and  $c_{2l}$  within one time step. Furthermore, it follows that the path  $c_{2l-1} \rightarrow c_{2l-2} \rightarrow \dots \rightarrow c_l \rightarrow c_{l-1}$  is a dominant strategy for player 2, in which case it gets a payoff of 1 and in response, player 1 gets 0 from any strategy. Clearly, this is a pure equilibrium, leading to a social welfare of 1. On the other hand, if player 2 captures information from (only)  $c_{2l}$  allowing player 1 to capture from  $c_l$  which it can reach in the  $l$ th time step, then the social welfare can reach its maximum value of  $2 - \epsilon$ . Since there is only one equilibrium, the PoS, as well as the PoA for this game is  $2 - \epsilon$ , i.e., it approaches 2 as  $\epsilon$  approaches 0. Thus, we have the following theorem.

**Theorem 5:** *The multistep game with temporal aspect has a PoS and PoA of 2.*

**2) Existence of Pure Equilibria:** Unlike the single-step game, the general multistep game may not always have a PNE, as demonstrated by the following example.

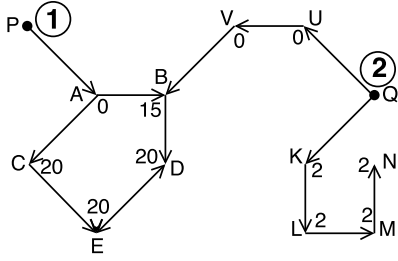


Fig. 3. Counterexample where PNE does not exist. Players 1 and 2, with  $\rho_1 = \rho_2 = 1$ , start in cells  $P$  and  $Q$ , respectively, and play the game (with temporal aspect) for 4 time steps.

Consider the game as shown in Fig. 3. The connectivity between cells is given by the directed edges and the information initially available in each cell is shown. The number of time steps is 4. Players 1 and 2, with sensor effectiveness  $\rho_1 = \rho_2 = 1$ , are initially in cells  $P$  and  $Q$ , respectively. Thus, player 1 has two strategies: paths  $P \rightarrow A \rightarrow C \rightarrow E \rightarrow D$  and  $P \rightarrow A \rightarrow B \rightarrow D \rightarrow E$ . Let these be called “Left” and “Right,” respectively. Player 2 similarly has two strategies, say “Up” and “Down” corresponding to paths  $Q \rightarrow U \rightarrow V \rightarrow B \rightarrow D$  and  $Q \rightarrow K \rightarrow L \rightarrow M \rightarrow N$ . Then, the payoff matrix is given by

$$\begin{array}{cc} & \begin{array}{cc} \text{Up} & \text{Down} \end{array} \\ \begin{array}{c} \text{Left} \\ \text{Right} \end{array} & \begin{bmatrix} (50, 25) & (60, 8) \\ (55, 0) & (55, 8) \end{bmatrix}. \end{array}$$

It is easy to check that there is no PNE in this case.

#### D. Multistep Games Without Temporal Aspect

In this section, we analyze the game without temporal aspect. Here, the payoff of a player from a visit to a cell is merely dependent on which players visit the cell over the complete course of the game and how many times, regardless of the order in which they visit the cell. Thereby, the combined payoff of players from their visits to a cell  $c$  (some of them possibly simultaneously), also does not depend on the order of visits, and can be easily computed using (9). Due to the different settings in this game than the one with temporal aspect, we opt for a slightly different notation. Let  $A$ ,  $B$  be multisets with support in  $\mathcal{P}$ , with  $m_A$ ,  $m_B$  the respective multiplicity functions, such that  $B \subseteq A$ . Then, for a cell  $c$ , let  $\theta_B^A(c)$  denote the combined payoff which the visitors in  $B$  would obtain by visiting cell  $c$  (as many times as the respective multiplicities in  $B$ ), when the complete set of visitors for  $c$  is given by  $A$ . Naturally, this is only meaningful when  $B \subseteq A$ .

**Lemma 6:** Let  $A$ ,  $B$ ,  $B'$ , and  $D$  be multisets with support in  $\mathcal{P}$  subject to  $B' \subseteq B \subseteq A$  and  $B \uplus D \subseteq A$ . Then, (1)  $\theta_B^A(c) + \theta_D^A(c) = \theta_{B \uplus D}^A(c)$ , (2)  $\theta_{B'}^A(c) \leq \theta_B^A(c)$ , and (3)  $\theta_{A|_i}^A(c) \geq \pi_{A|_i}^{A|-i}(c)$ .

*Proof:* The first part follows from definition, as both sides equal the combined payoff of visitors in  $B$  and  $D$  when the complete set of visitors is given by  $A$ . Next, applying the first part on  $A$ ,  $B'$ ,  $B \setminus B'$ , respectively, we get  $\theta_B^A(c) - \theta_{B'}^A(c) = \theta_{B \setminus B'}^A(c) \geq 0$ . Rearranging gives the result in the second part. For the third part, note that in a game with temporal aspect,

in the case when the visitor set of a cell  $c$  is  $A$ , one possible outcome  $s$  corresponds to all the visitors in  $A$  visiting in the same time step, and thus,  $\theta_{A|_i}^A(c)$  is a possible payoff of player  $i$  from cell  $c$  when its visitor set is fixed to  $A$ . Consequently, the result follows from Lemma 4.  $\square$

To show smoothness for this game, we proceed similarly as in the game with temporal aspect. Let  $s$  and  $s^*$  be any two outcomes. For every player  $i$ , let  $q^i$  denote the outcome  $(s^*, s_{-i})$ . For any cell  $c$ , let multisets  $A_c$  and  $A_c^*$  denote the visitor sets for cell  $c$  when the outcomes are  $s$  and  $s^*$  respectively. Analogous to (16), (17), and (18), we get

$$\begin{aligned} \sum_{i \in \mathcal{P}} \pi_i(s) &= \sum_{c \in \mathcal{C}} \theta_{A_c}^A(c) \\ \sum_{i \in \mathcal{P}} \pi_i(s^*) &= \sum_{c \in \mathcal{C}} \theta_{A_c^*}^A(c) \\ \sum_{i \in \mathcal{P}} \pi_i(q^i) &= \sum_{c \in \mathcal{C}} \sum_{i \in A_c^*} \theta_{A_c|-i \uplus A_c^*|_i}^A(c). \end{aligned}$$

Next, adding the first and third equations, and subtracting the second gives

$$\begin{aligned} \sum_{i \in \mathcal{P}} \pi_i(q^i) + \sum_{i \in \mathcal{P}} \pi_i(s) - \sum_{i \in \mathcal{P}} \pi_i(s^*) \\ \geq \sum_{c \in \mathcal{C}} \left( \theta_{A_c}^A(c) + \sum_{i \in A_c^*} \theta_{A_c|-i \uplus A_c^*|_i}^A(c) - \theta_{A_c^*}^A(c) \right). \end{aligned} \quad (22)$$

Finally, note that  $\theta_{A_c}^A(c) = \pi_{A_c}^\emptyset(c)$  and  $\theta_{A_c^*}^A(c) = \pi_{A_c^*}^\emptyset(c)$  by definition. Furthermore, applying Lemma 6 on  $A_c|-i \uplus A_c^*|_i$ , we get,  $\theta_{A_c|-i \uplus A_c^*|_i}^A(c) \geq \pi_{A_c^*|_i}^{A_c|-i}(c)$ . With this, (22) becomes

$$\begin{aligned} \sum_{i \in \mathcal{P}} \pi_i(q^i) + \sum_{i \in \mathcal{P}} \pi_i(s) - \sum_{i \in \mathcal{P}} \pi_i(s^*) \\ \geq \sum_{c \in \mathcal{C}} \left( \pi_{A_c}^\emptyset(c) + \sum_{i \in A_c^*} \pi_{A_c^*|_i}^{A_c|-i}(c) - \pi_{A_c^*}^\emptyset(c) \right). \end{aligned} \quad (23)$$

As seen in Section IV-C, the right-hand side of the above is nonnegative by Lemma 5, and in turn, so is the left-hand side, which provides the desired result and the corollary using [41].

**Theorem 6:** The multistep game without temporal aspect is  $(1, 1)$ -smooth.

**Corollary 3:** The multistep game without temporal aspect has a PoA at most 2.

Next, we demonstrate the tightness of this bound.

#### E. Tight Example

Consider the game as shown in Fig. 4. The set of players is  $\{1, \dots, p\}$ , with  $\rho_1 = \rho_2 = \dots = \rho_p = 1$ . The number of time steps is  $l$ . The cells in the network and the connectivity among them are as shown. The information initially available in cell  $c$  is 1, whereas in cells  $c_l^2, \dots, c_l^p$ , it is  $v = l/(p-1)l + 1 - \epsilon$ , where  $\epsilon$  is a small positive constant. The information available is 0 in all other cells. Player 1 is initially in cell  $c_{l-1}^1$ , whereas every other player  $i > 1$  is initially in cell  $c_0^i$ . Now, the only strategy for player 1 is the path  $c_{l-1}^1 \rightarrow c_{l-2}^1 \rightarrow \dots \rightarrow c_0^1 \rightarrow c$ . Now, the outcome where every other player  $i > 1$  plays the



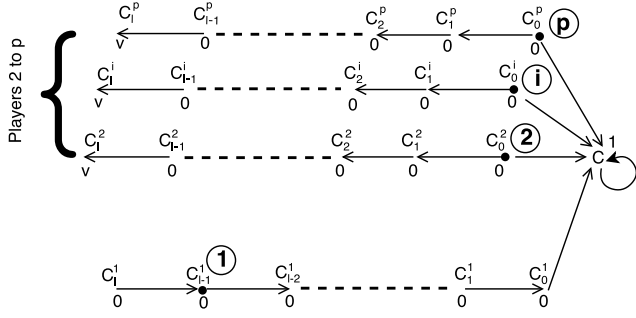


Fig. 4. Game without temporal aspect where PoA can get arbitrarily close to 2 with a high number of players and time steps. Player 1 is initially in cell  $c_{l-1}^1$ , whereas every other player  $i > 1$  is initially in cell  $c_0^i$ .

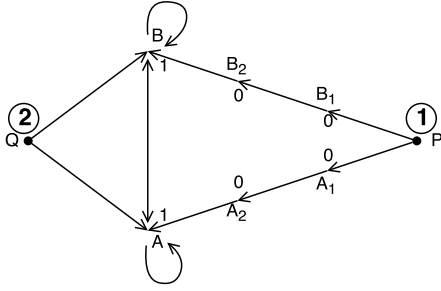


Fig. 6. Counterexample where PNE does not exist. Players 1 and 2, with  $\rho_1 = \rho_2 = 0.8$ , start in cells  $P$  and  $Q$ , respectively, and play the game (without temporal aspect) for 3 time steps.

strategy  $c_0^i \rightarrow c \rightarrow c \rightarrow \dots \rightarrow c \rightarrow c$  gives every player  $i > 1$ , a payoff which evaluates to  $l/(p-1)l + 1$ . Thus, no player  $i$  wants to deviate to the other possible strategy  $c_0^i \rightarrow c_1^i \rightarrow \dots \rightarrow c_l^i$  as it gives a smaller payoff of  $l/(p-1)l + 1 - \epsilon$ . Thus, the aforesaid outcome is a PNE, which has a social welfare of 1. However, it can be seen that the social welfare increases as more and more players switch to the respective alternative strategy, and in the extreme case of every player  $i > 1$  switching to the respective strategy  $c_0^i \rightarrow c_1^i \rightarrow \dots \rightarrow c_l^i$ , the social welfare reaches the maximum value of  $2 - 1/(p-1)l + 1 - (p-1)\epsilon$ , giving the same value of PoA. Thus, as  $\epsilon \rightarrow 0$ , it approaches  $2 - 1/(p-1)l + 1$ , which, in turn, can become arbitrarily close to 2 if the parameters  $p$  or  $l$  become arbitrarily large, showing that the bound of 2 is tight.

**Existence of PNE:** The following example shows that unlike the single-step game the general multistep game without temporal aspect may not always have a PNE.

Consider the game as shown in Fig. 6. The connectivity between cells is given by the directed edges and the information initially available in each cell is shown. The number of time steps is 3. Players 1 and 2, with sensor effectiveness  $\rho_1 = \rho_2 = 0.8$ , are initially in cells  $P$  and  $Q$ , respectively. Thus, player 1 has two strategies: paths  $P \rightarrow A_1 \rightarrow A_2 \rightarrow A$  and  $P \rightarrow B_1 \rightarrow B_2 \rightarrow B$ . Let these be called “ $s_A$ ” and “ $s_B$ ,” respectively. Player 2 has eight strategies, however, since the sequence of the visits does not matter, there are four distinct ones. Let these be called “ $A_3B_0$ ,” “ $A_2B_1$ ,” “ $A_1B_2$ ,” and “ $A_0B_3$ ,” where “ $A_iB_j$ ” denotes a strategy which visits  $A$

$i$  times and  $B$   $j$  times. Then, the payoff matrix for this game is given by

$$\begin{matrix} & A_3B_0 & A_2B_1 & A_1B_2 & A_0B_3 \\ \begin{matrix} s_A \\ s_B \end{matrix} & \begin{bmatrix} (0.24, 0.74) \\ (0.80, 0.99) \end{bmatrix} & \begin{bmatrix} (0.33, 1.46) \\ (0.48, 1.44) \end{bmatrix} & \begin{bmatrix} (0.48, 1.44) \\ (0.33, 1.46) \end{bmatrix} & \begin{bmatrix} (0.80, 0.99) \\ (0.24, 0.74) \end{bmatrix} \end{matrix}$$

It follows that there is no pure equilibrium in this case.

#### F. Maximum Social Welfare Computational Complexity

In Section V, we propose routing policies which achieve near-optimal social welfare and can be computed fast. The need for such policies arises from the hardness of computing the maximum social welfare (MSW) (and a strategy that yields it) in the multistep UAV games. We call this problem UAV-MSW in short. For our experiments, we consider a surveillance environment in the form of a grid of cells—the corresponding graph can be seen to be a more generalized version of the rectangular subgrid graph (RSG) defined as follows [45].

**Definition 9:** Let  $G^\infty$  be the infinite (undirected) graph whose vertex set consists of all points of the plane with integer coordinates and in which two vertices are connected if and only if the (Euclidean) distance between them is equal to 1. An  $m \times n$  RSG is a subgraph of  $G^\infty$  with vertex set (without loss of generality)  $\{v \mid 1 \leq v_x \leq m, 1 \leq v_y \leq n\}$ .

If the edges are directed, we call the graph directed RSG (DRSG). We describe the experimental setup in Section V which shows how DRSGs are a special case of the graphs we consider. Hence, in the interest of showing the computational difficulty, it suffices to show that the following theorem.

**Theorem 7:** UAV-MSW on DRSGs is NP-hard.

**Proof:** The hamiltonian path (HP) problem is the decision problem to determine if there exists a path in a given graph which visits all the vertices exactly once. Solving HP on RSGs is shown to be NP-complete in [45]. This can be reduced to UAV-MSW on RSGs as follows. Let  $G = (C, E)$  be an  $m \times n$  RSG given as input for the HP problem. The vertex set  $C$  consists of all the cells in the grid. Consider a DRSG  $G' = (C, E')$ , where for each edge  $(c, c')$  in  $E$ ,  $E'$  has directed edges  $(c, c')$ ,  $(c', c)$ . Then, we construct  $|C|$  ( $= mn$ ) UAV game instances, one corresponding to each  $c \in C$ . Let each instance has  $G'$  as the surveillance region, number of time steps  $l = |C| - 1$  and let  $\mathcal{P} = \{1\}$ , i.e., only one player whose sensor effectiveness be  $\rho = 1$ . For the instance corresponding to  $c \in C$ , let the initial position of the player be cell  $c$  with value  $v(c) = 0$  and let all other cells have value 1. These parameters completely define a multistep UAV game instance. It is easy to see that player 1 can achieve the maximum payoff of  $|C| - 1$  having started in  $c$  if and only if it can visit all other  $|C| - 1$  cells in as many steps (which is the maximum social welfare since there is only one player). Thus, the maximum social welfare is  $|C| - 1$  for the instance corresponding to cell  $c$  if and only if  $G'$  has a directed HP starting from  $c$ , or equivalently, if and only if  $G$  has an HP starting from  $c$  by construction. Thus, HP outputs YES on input  $G$ , if and only if  $G$  has an HP starting from some cell  $c$ , equivalently, if and only one of the aforementioned UAV-MSW instances outputs

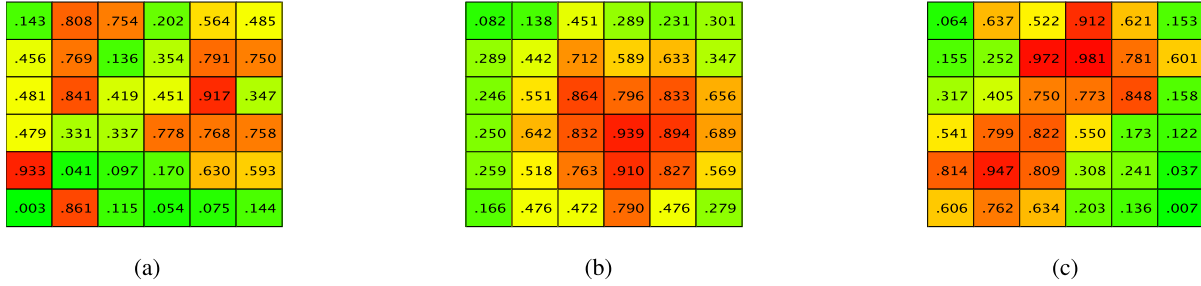


Fig. 5. Randomly generated instances on  $6 \times 6$  grids for the problem scenarios with 0, 1, and 2 peaks respectively. (a) No peaks. (b) One peak. (c) Two peaks.

$|C| - 1$ . This reduction requires polynomially many calls to the UAV-MSW oracle, and this shows that UAV-MSW is NP-hard. (Note that the reduction works regardless of whether the game is with or without the temporal aspect.)  $\square$

## V. EXPERIMENTATION

In a competitive game environment like in the UAV problem, a typical solution is a PNE. As seen from the previous sections, however, a PNE may not exist in the general multistep game. In this section, we devise alternative easily implementable routing policies with reasonable performance guarantees. We propose such policies and simulate plausible problem scenarios in the temporal setting with randomly generated game parameters and evaluate these policies for their performance and study the dependence on various game parameters.

Previous work using similar models, in particular, [4], has studied fundamentally similar policies, however, they do not provide significant insights on the social welfare performance which is the goal of our experimental evaluation. In addition, we provide the analysis for much broader parameter settings in terms of the number of players, time steps, the search arena size and connectivity, and the information distributions. We discuss this setup in detail as follows.

### A. Setup

The randomly generated game instances are in the following setting. The set of cells is in the form of a grid of fixed dimensions (which can be varied across different scenarios to study the dependence on the grid size), which is a common representation, such as in [28]. To allow arbitrary connectivity constraints, we have a cell connectivity parameter  $\delta$  which functions as follows. For each cell, each of the cells within a Chebychev distance of 1 (i.e., rowwise, columnwise, or diagonalwise adjacent cells and the cell itself) is independently chosen to be its out-neighbor with probability  $\delta = 0.5$ .<sup>2</sup>

Since DRSGs can have edges only between cells that are at Euclidean distance 1, it follows that the set of DRSGs is a subset of the grid environments we consider, and consequently, UAV-MSW is NP-hard due to Theorem 7. In every game instance, the sensor effectiveness parameters of the UAVs are

<sup>2</sup>If no cells get chosen after having gone through all the adjacent cells, we repeat the process, until there is at least one out-neighbor for the cell—so that the graph does not have sinks where the UAVs can get stuck.

initialized to values chosen uniformly at random (u.a.r.) from  $[0, 1]$ . The initial positions of the UAVs are chosen u.a.r. from among all the grid cells.

In many real-world problems such as target detection, a few independent cells are of high value, and the value decays as one moves away from these cells. Such cells of the highest value are referred to as the *peaks*. The cell information values for these scenarios are set as follows.

$k$  peaks ( $k > 0$ ): We first fix  $k$  peaks  $p_1, \dots, p_k$  as points from the grid chosen u.a.r. independently. Then, the information value around a peak  $p_i$  as a function of distance  $d$  from it is given by a generalized normal distribution  $f_i(d) = \alpha_i 2^{-\beta_i d^{\gamma_i}}$ , where the parameters  $\alpha_i$ ,  $\beta_i$ , and  $\gamma_i$  are chosen u.a.r. from ranges  $[0.8, 1]$ ,  $[0.1, 0.15]$ , and  $[1, 1.25]$  for each peak  $p_i$ , of which  $\gamma_i$  acts as the shape parameter of the distribution. The aforesaid ranges for the parameters  $\alpha_i$ ,  $\beta_i$ , and  $\gamma_i$  are chosen so as to keep the information values in a good range. Subsequently, each cell is assigned a value  $f_i(d)$  by determining the peak  $p_i$  closest to the cell's center and its distance  $d$  from  $p_i$ .

### B. Zero Peaks

In this scenario, cell information values are simply chosen u.a.r. from  $[0, 1]$ .

Fig. 5 shows randomly generated instances on  $6 \times 6$  grids for these scenarios. To study any of the scenarios, we generate a large number of instances  $n$  with all the parameters randomly chosen as described earlier. We let the UAVs apply a routing policy to move in the search space and collect information accordingly. Our experiments evaluate the performance of these policies and the dependence on various game parameters.

### C. Routing Policies

We let the routing policies to be defined with a parameter  $\tau$  called as the horizon, which denotes the length of the walks that the players consider as strategies. This is often a case in the UAV routing domain, due to limitations on computational power and/or the visual sensor span, etc. Thus, the total number of time steps  $l$  could be much larger and the UAVs route by successively committing to walks of length  $\tau$ .<sup>3</sup> We first describe some rudimentary policies which are later used

<sup>3</sup>If the number of time steps  $l$  remaining is smaller than  $\tau$ , then naturally,  $\tau$  is reduced and set to  $l$ .

TABLE III  
DEPENDENCE OF  $\tau'$ , THE LONGEST HORIZON AT WHICH A PNE IS FOUND BY THE  $NE_5$  POLICY, ON THE NUMBER OF PEAKS AND PLAYERS. THE OTHER PARAMETERS ARE SET TO  $n = 1000$ ,  $l = 5$ , AND GRID SIZE  $6 \times 6$

$\tau'$		Number of peaks														
		0					1					2				
		5	4	3	2	1	5	4	3	2	1	5	4	3	2	1
Number of players	2	894	79	23	3	1	737	162	69	30	2	790	146	47	16	1
	3	697	193	78	29	3	475	272	192	56	5	490	243	184	73	10
	4	490	266	166	62	16	290	281	253	54	22	274	291	277	138	20
	5	281	273	272	150	40	183	266	303	207	41	156	236	327	236	45

TABLE IV  
RUNTIME (IN SECONDS) COMPARISON AND DEPENDENCE ON GAME PARAMETERS (GRID:  $6 \times 6$ )

		Number of peaks								
		0			1			2		
		GR	NE	SW	GR	NE	SW	GR	NE	SW
Number of players; ( $l = 6$ )	2	0.12	0.49	0.12	0.12	0.40	0.45	0.11	0.68	0.94
	3	0.16	0.68	1.38	0.12	0.61	11.60	0.19	1.59	6.39
	4	0.18	1.92	19.42	0.15	1.54	82.77	0.16	1.93	51.03
Number of time-steps; ( $p = 4$ )	4	0.01	0.05	0.12	0.01	0.04	0.28	0.01	0.05	0.20
	5	0.04	0.26	2.54	0.04	0.30	14.29	0.05	0.48	9.84
	6	0.18	1.92	19.42	0.15	1.54	82.77	0.16	1.93	51.03

TABLE V  
EMPIRICAL PoA AND ITS DEPENDENCE ON GAME PARAMETERS

		Number of peaks								
		0			1			2		
		min	avg	max	min	avg	max	min	avg	max
Number of players; Grid: $5 \times 5$	2	1	1.025	1.401	1	1.033	1.271	1	1.026	1.291
	3	1	1.052	1.462	1	1.059	1.437	1	1.052	1.304
	4	1	1.078	1.631	1	1.065	1.356	1	1.069	1.440
Grid size; $p = 4$	$4 \times 4$	1	1.083	1.367	1	1.075	1.303	1	1.079	1.586
	$5 \times 5$	1	1.075	1.261	1	1.063	1.245	1	1.067	1.514
	$6 \times 6$	1	1.054	1.274	1	1.054	1.200	1	1.057	1.179

in describing our proposed policies and are also useful for comparison as baselines.

1) *Greedy Parameter  $\tau$* : In this policy, the strategies considered at each step are walks of length  $\tau$ . The strategy which yields the highest payoff to the player disregarding other UAVs in the environment is chosen under this policy.

2) *Stepwise Social Welfare Maximizer parameter  $\tau$* : This policy also considers walks of length  $\tau$  as strategies. Each player computes the joint strategy which induces maximum social welfare<sup>4</sup> and plays the individual strategy corresponding to this outcome.

Next, we propose routing policies which are based on a pure-strategy NE.

3) *Single-Step NE*: In this policy, a UAV computes and plays a PNE of the single-step game at every time step. Such a PNE is guaranteed to exist in single-step games via Theorem 1. To ensure that all the players compute the same equilibrium, we require that the procedure of computing it is mutually agreed upon before the game begins, as follows: the equilibrium is obtained by applying the best-response (BR)

<sup>4</sup>Note that since the horizon  $\tau$  can be a lot smaller than the total number of time steps  $l$ , hence, greedily maximizing the social welfare in each step need not lead to the true maximum social welfare of the game. (In fact, any policy with a limited horizon can be made to yield arbitrarily bad social welfare, on adversarially constructed game instances.)

TABLE VI  
DEPENDENCE OF THE DEGREE OF OVERLAP ON THE GRID SIZE ( $p = 2$ ).

		Number of peaks		
		0	1	2
Grid Size	$4 \times 4$	0.471	0.431	0.428
	$5 \times 5$	0.309	0.292	0.286
	$6 \times 6$	0.301	0.279	0.282

dynamics, the initial outcome for the BR dynamics is taken as the one with every player playing as per the greedy myopic policy, and the order of players for sequentially computing the BRs is taken to be the descending order of  $\rho_i$ s.

4) *Multistep NE Parameter  $\tau$* : In this policy as well, a UAV tries to compute a PNE via a mutually agreed BR dynamics procedure as in the single-step NE case, except that the walk is of length  $\tau$ . Since such a PNE is not guaranteed to exist, if it is not found within  $2p\tau$  rounds<sup>5</sup> of BR dynamics, the player reduces the horizon to  $\tau - 1$  and attempts to find a PNE. The horizon is repeatedly lowered until a PNE is found, for horizon (say)  $\tau'$ . Note that  $\tau' \geq 1$  since a PNE always exists in the single-step game. Having computed such a PNE, the players choose the respective strategies which are walks of length  $\tau'$ .

<sup>5</sup> $2p\tau$  is chosen as a conservative threshold—if the BR dynamics does not converge in so many rounds, it is highly likely to never converge.



TABLE VII  
WELFARE QUOTIENT OF THE  $NE_\tau$  POLICY, AND ITS DEPENDENCE ON THE HORIZON  $\tau$

		Number of peaks								
		0			1			2		
		min	avg	max	min	avg	max	min	avg	max
Horizon	1	0.795	0.995	1.164	0.932	1.007	1.172	0.883	1.005	1.266
	2	0.746	1.005	1.214	0.859	1.009	1.436	0.914	1.011	1.279
	3	0.819	1.007	1.270	0.939	1.010	1.446	0.820	1.017	1.187
	4	0.888	1.012	1.330	0.958	1.011	1.203	0.848	1.017	1.274
	5	0.794	1.017	1.317	0.938	1.011	1.194	0.844	1.019	1.500

Table III shows the multistep NE policy in operation for various scenarios. When operated with  $\tau = 5$ , a PNE was found at the highest horizon in 894 out of 1000 times for the two-player game with no peaks. However, this proportion drops as the number of players increases (for the same number of peaks). Also, it drops as the number of peaks increases (for the same number of players). Indeed, for the game with two peaks and five players, it is a modest 156 out of 1000 times. These values and trends critically suggest limiting  $\tau$  to as high as 5 in our experiments.

For convenience, we denote the Greedy policy with horizon  $\tau$  by  $GR_\tau$ , and the stepwise social welfare (SW) maximizer policy with horizon  $\tau$  as  $SW_\tau$ . Similarly, we denote the multistep NE policy with horizon  $\tau$  by  $NE_\tau$ , and thus, the single-step NE policy in particular by  $NE_1$ .

#### D. Evaluation and Results

In this section, we discuss how the policies are evaluated. We compute the social welfare optimality of the  $NE_\tau$  policy by comparing it against  $SW_\tau$ . As explained before using Table III, the marginal utility of having a high horizon sharply decreases while the computation time naturally increases rapidly with horizon as the strategy space for each player increases exponentially with it. Similarly, computing the maximum social welfare for a large horizon becomes computationally intractable. Even for the scenarios studied below with low horizon and number of players, the brute force search takes time exponential in these parameters, and hence, we compute it using an elaborate branch-and-bound technique. We conduct experimental runtime analysis on a small grid size  $6 \times 6$  yielding results as shown in Table IV. For  $l = 6$  time steps, comparing policies  $SW_6$ ,  $GR_6$ , and  $NE_6$  shows that the runtime (in seconds) increases with the number of players much more rapidly for  $SW_6$  while it is relatively stable and small for  $GR_6$  and  $NE_6$ . Similarly, for a fixed number of players  $p = 4$ , the runtime increases very rapidly with  $l$  for the SW policy while it is relatively stable for the greedy and the NE policies. Also, the runtimes are seen to be greater for the scenario with one peak as compared to others.

1) *Empirical PoA*: We first note that  $SW_\tau$  is guaranteed to achieve the maximum social welfare only when  $\tau = l$ . Hence, in our first experiment, we set  $l = \tau = 5$  and compare  $NE_5$  against  $SW_5$ . To evaluate the performance of  $NE_5$  on a game instance, we compute the social welfare obtained when all the UAVs employ  $NE_5$  and the one obtained when all employ  $SW_5$ , and then obtain their ratio (that of latter to former). To obtain the *empirical PoA*, we average this ratio

over  $n = 1000$  randomly generated game instances. In our results below, we also include the minimum and the maximum ratios obtained across the  $n$  instances, for more insight.

We analyze the dependence of the empirical PoA on the number of players by varying  $p$  between 2 and 4, fixing the grid size to  $5 \times 5$ . Similarly, we analyze the dependence on grid size by varying it between  $6 \times 6$ ,  $5 \times 5$ , and  $4 \times 4$ , while fixing  $p = 4$ . We consider scenarios with up to two peaks. The results are shown in Table V.

Observe that, in each scenario, the empirical PoA increases with the number of players and decreases with the grid size. A plausible explanation for this is that, with only a few players, or with a large grid, players are more likely to be far from each other, thus with little overlap between their strategies. As a result, their BRs to each other are more likely to be dominant strategies which yield the individual maximum payoffs possible when there are no other players. Thus, the  $NE_5$  policy is more likely to get a social welfare close to the maximum, resulting in a low PoA. To capture this more closely, we define and analyze the *degree of overlap* as follows.

We fix the number of players to 2 and for  $n = 1000$  instances, we compute the average probability that two strategies, one each drawn u.a.r. from the set of strategies of the two players, respectively, overlap. Observe that when the two players are in the same location or very close, this probability is very high, whereas, when the players are far away, their strategies are nonoverlapping walks, making this probability 0. Table VI shows the dependence of this degree of overlap on the grid size. As we intuitively expect, this dependence does align with that of the empirical PoA in Table V, as both the quantities are higher for the smaller grid size and for the scenario with no peaks.

Overall, the results indicate that the  $NE_5$  policy does well as compared to the optimal  $SW_5$  policy, with the empirical PoA smaller than a meager 1.1 on average. Also, the worst instances found in all the scenarios have a ratio well below the theoretical bound of 2 while there are always instances where the policy does achieve the maximum social welfare.

2) *Experiments on Large Problem Instances*: Here, we consider larger game parameters by setting the grid size to  $10 \times 10$ , the number of time steps  $l = 25$ , and the number of players  $p = 4$ . We compare  $NE_\tau$  against  $SW_\tau$ , keeping the horizon equal for the policies being compared. We vary  $\tau$  from 1 through 5, to study the dependence on the horizon. The evaluation is done by obtaining the ratio of social welfare averaged over  $n = 1000$  randomly generated game instances,

as described before. We call this the *welfare quotient* of  $NE_\tau$  (and not empirical PoA, since it is not compared against the true maximum social welfare). Naturally, lower the quotient, better is the performance. The results are shown in Table VII.

We observe that the average welfare quotient of  $NE_\tau$  for all horizons (up to 5) is very close to 1. As pointed out earlier,  $SW_\tau$  does not yield the true maximum social welfare, and hence,  $NE_\tau$  could yield a higher welfare than  $SW_\tau$  at times causing the quotient to be less than 1, which is prominently reflected in many instances, such as one in the scenario with no peaks where  $SW_2$  achieves only about 75% of the welfare achieved by  $NE_2$ . Thus, having a horizon much lower than the number of time steps critically affects the performance of  $SW_\tau$ , and as a result,  $NE_\tau$  does almost as well in comparison, with a mild rise in the welfare quotient as  $\tau$  increases.

## VI. CONCLUSION

UAVs are becoming increasingly popular in search and reconnaissance missions due to their ability to conduct tedious tasks, inexpensively and without risk to human life. When the area of operations is large, a fleet of UAVs might be deployed. This paper addressed the problem of autonomous routing of a UAV fleet in a communication denied area where the UAVs are not allowed to exchange information or negotiate trajectories. The problem is modeled in a game theoretic framework. More specifically, we formulated the multiagent UAV routing problem as a game where UAVs are players and their strategies are the different routes they can take. We have considered many useful and interesting concepts including temporal aspect and fusion.

This paper develops many research results on methodological and practical sides. For the commander or manager of the fleet, it is reassuring to know that the multiagent routing policies developed in this paper can produce outcomes that are theoretically bounded by one-half of the optimal centralized policy if, in fact, one could come up with it. This bound is independent of the length of the routes and independent of the fleet size. Furthermore, we have categorized the single-step game as a new variant of weighted congestion-type games. For this class, and for the general formulations of the multistep UAV game, we prove the existence or nonexistence of the pure equilibrium. Along with this, tight bounds on PoA and PoS for all the various formulations, constitute the majority of the technical results of this paper.

To learn the actual PoA and PoS values in practice, we juxtapose our theoretical results with empirical studies. We propose a multiagent UAV routing policy, which, on the average, produces total information gain much closer to the centralized solution than to the theoretical bound. We also find that PoA increases with the congestion level when the UAVs compete for a smaller area or more UAVs are added to the search area. We also highlight the limitations of having a small horizon, as an upshot of which, our proposed routing policy performs almost as well as the policy which greedily maximizes the social welfare at each horizon.

The future effort could be directed at studying these policies for longer rolling horizons and for persistent surveillance

scenarios. It is also natural to think about the obsolescence of information where the collected value can decay due to the passage of time, and a UAV has to be routed back to collect the information being “built up.” When intermittent communication is permitted, new cooperation strategies can be conceived and reduction in PoA can be studied. Routing to avoid an adversarial team is also an interesting area that can benefit from the framework developed in this paper.

## ACKNOWLEDGMENTS

The authors would like to thank the anonymous reviewers and Associate Editor for providing valuable comments that helped to improve this paper.

## REFERENCES

- [1] C. B. P. Tice, “Unmanned aerial vehicles: The force multiplier of the 1990s,” *Airpower J.*, vol. 5, no. 1, pp. 41–55, Spring 1991.
- [2] L. Gupta, R. Jain, and G. Vaszun, “Survey of important issues in UAV communication networks,” *IEEE Commun. Surveys Tuts.*, vol. 18, no. 2, pp. 1123–1152, 2nd Quart., 2016.
- [3] I. Khoufi, P. Minet, and N. Achir, “Unmanned aerial vehicles path planning for area monitoring,” in *Proc. Int. Conf. Perform. Eval. Model. Wired Wireless Netw. (PEMWN)*, Nov. 2016, pp. 1–5.
- [4] P. B. Sujit and D. Ghose, “Multiple agent search of an unknown environment using game theoretical models,” in *Proc. Amer. Control Conf.*, Jun./Jul. 2004, vol. 6, pp. 5564–5569.
- [5] H. Ortiz-Peña, M. Hirsch, M. Karwan, R. Nagi, and M. Sudit, “Representation of potential information gain to measure the price of anarchy on ISR activities,” *Proc. SPIE*, vol. 8758, May 2013, Art. no. 87580M.
- [6] D. L. Hall and J. Llinas, “An introduction to multisensor data fusion,” *Proc. IEEE*, vol. 85, no. 1, pp. 6–23, Jan. 1997.
- [7] R. R. Brooks and S. S. Iyengar, *Multi-Sensor Fusion: Fundamentals and Applications With Software*. Upper Saddle River, NJ, USA: Prentice-Hall, 1998.
- [8] J. Nash, “Non-cooperative games,” *Ann. Math.*, vol. 54, no. 2, pp. 286–295, 1951.
- [9] T. Roughgarden, *Twenty Lectures on Algorithmic Game Theory*. Cambridge, U.K.: Cambridge Univ. Press, 2016.
- [10] O. Johannes and K. M. Royset, “Aerial search optimization model (ASOM) for UAVs in special operations,” *Mil. Oper. Res.*, vol. 13, no. 1, pp. 23–33, 2008.
- [11] K. P. Tognetti, “An optimal strategy for a whereabouts search,” *Oper. Res.*, vol. 16, no. 1, pp. 209–211, 1968.
- [12] J. B. Kadane, “Discrete search and the Neyman-Pearson lemma,” *J. Math. Anal. Appl.*, vol. 22, no. 1, pp. 156–171, Apr. 1968.
- [13] J. B. Kadane, “Optimal whereabouts search,” *Oper. Res.*, vol. 19, no. 4, pp. 894–904, 1971.
- [14] L. D. Stone, *Theory of Optimal Search*. New York, NY, USA: Academic Press, 1975.
- [15] R. Ahlswede and I. Wegener, *Search Problems*. Hoboken, NJ, USA: Wiley, 1987.
- [16] Y. Xia, R. Batta, and R. Nagi, “Controlling a fleet of unmanned aerial vehicles to collect uncertain information in a threat environment,” *Oper. Res.*, vol. 65, no. 3, pp. 674–692, 2017.
- [17] B. Alidaee, H. Wang, and F. Landram, “On the flexible demand assignment problems: Case of unmanned aerial vehicles,” *IEEE Trans. Autom. Sci. Eng.*, vol. 8, no. 4, pp. 865–868, Oct. 2011.
- [18] K. Sundar and S. Rathinam, “Algorithms for routing an unmanned aerial vehicle in the presence of refueling depots,” *IEEE Trans. Autom. Sci. Eng.*, vol. 11, no. 1, pp. 287–294, Jan. 2014.
- [19] C. Schumacher, P. R. Chandler, M. Pachter, and L. S. Pachter, “Optimization of air vehicles operations using mixed-integer linear programming,” *J. Oper. Res. Soc.*, vol. 58, no. 4, pp. 516–527, 2007.
- [20] C. C. Murray and M. H. Karwan, “An extensible modeling framework for dynamic reassignment and rerouting in cooperative airborne operations,” *Nav. Res. Logistics*, vol. 57, no. 7, pp. 634–652, 2010.
- [21] S. G. Manyam, S. Rasmussen, D. W. Casbeer, K. Kalyanam, and S. Manickam, “Multi-UAV routing for persistent intelligence surveillance reconnaissance missions,” in *Proc. Int. Conf. Unmanned Aircr. Syst. (ICUAS)*, Jun. 2017, pp. 573–580.

[22] M. Mozaffari, W. Saad, M. Bennis, and M. Debbah, "Efficient deployment of multiple unmanned aerial vehicles for optimal wireless coverage," *IEEE Commun. Lett.*, vol. 20, no. 8, pp. 1647–1650, Aug. 2016.

[23] D. Sikeridis, E. E. Tsiropoulou, M. Devetsikiotis, and S. Papavassiliou, "Wireless powered Public Safety IoT: A UAV-assisted adaptive-learning approach towards energy efficiency," *J. Netw. Comput. Appl.*, vol. 123, pp. 69–79, Dec. 2018.

[24] E. E. Tsiropoulou, P. Vamvakas, and S. Papavassiliou, "Supermodular game-based distributed joint uplink power and rate allocation in two-tier femtocell networks," *IEEE Trans. Mobile Comput.*, vol. 16, no. 9, pp. 2656–2667, Sep. 2017.

[25] E. E. Tsiropoulou, P. Vamvakas, and S. Papavassiliou, "Energy efficient uplink joint resource allocation non-cooperative game with pricing," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, Apr. 2012, pp. 2352–2356.

[26] D. Sikeridis, E. E. Tsiropoulou, M. Devetsikiotis, and S. Papavassiliou, "Self-adaptive energy efficient operation in UAV-assisted public safety networks," in *Proc. IEEE 19th Int. Workshop Signal Process. Adv. Wireless Commun. (SPAWC)*, Jun. 2018, pp. 1–5.

[27] C. Shi, F. Wang, M. Sellathurai, and J. Zhou, "Non-cooperative game theoretic power allocation strategy for distributed multiple-radar architecture in a spectrum sharing environment," *IEEE Access*, vol. 6, pp. 17787–17800, 2018.

[28] J. Scherer and B. Rinner, "Persistent multi-UAV surveillance with energy and communication constraints," in *Proc. IEEE Int. Conf. Autom. Sci. Eng. (CASE)*, Aug. 2016, pp. 1225–1230.

[29] S. Rasmussen, K. Kalyanam, S. Manyam, D. Casbeer, and C. Olsen, "Practical considerations for implementing an autonomous, persistent, intelligence, surveillance, and reconnaissance system," in *Proc. IEEE Conf. Control Technol. Appl. (CCTA)*, Aug. 2017, pp. 1847–1854.

[30] A. Hayrapetyan, E. Tardos, and T. Wexler, "A network pricing game for selfish traffic," *Distrib. Comput.*, vol. 19, no. 4, pp. 255–266, 2007.

[31] P. B. Charlesworth, "A non-cooperative game to coordinate the coverage of two communications UAVs," in *Proc. IEEE Mil. Commun. Conf.*, Nov. 2013, pp. 668–673.

[32] P. B. Charlesworth, "Using non-cooperative games to coordinate communications UAVs," in *Proc. IEEE Globecom Workshops (GC Wkshps)*, Dec. 2014, pp. 1463–1468.

[33] T. Roughgarden and E. Tardos, "How bad is selfish routing?" *J. ACM*, vol. 49, no. 2, pp. 236–259, Mar. 2002.

[34] R. W. Rosenthal, "A class of games possessing pure-strategy Nash equilibria," *Int. J. Game Theory*, vol. 2, no. 1, pp. 65–67, 1973.

[35] I. Milchtaich, "Congestion games with player-specific payoff functions," *Games Econ. Behav.*, vol. 13, no. 1, pp. 111–124, 1996.

[36] H. Ackermann and H. Röglin, and B. Vöcking, "Pure Nash equilibria in player-specific and weighted congestion games," in *Proc. Int. Workshop Internet Netw. Econ.* Patras, Greece: Springer, 2006, pp. 50–61.

[37] K. Bhawalkar, M. Gairing, and T. Roughgarden, "Weighted congestion games: Price of anarchy, universal worst-case examples, and tightness," *ACM Trans. Econ. Comput.*, vol. 2, no. 4, p. 14, 2014.

[38] T. Harks, M. Klimm, and R. H. Möhring, "Characterizing the existence of potential functions in weighted congestion games," in *Proc. Int. Symp. Algorithmic Game Theory*. Paphos, Cyprus: Springer, 2009, pp. 97–108.

[39] V. Gkatzelis, K. Kollias, and T. Roughgarden, "Optimal cost-sharing in general resource selection games," *Oper. Res.*, vol. 64, no. 6, pp. 1177–1565, 2016.

[40] J. Kleinberg and S. Oren, "Mechanisms for (mis)allocating scientific credit," in *Proc. 43th Annu. ACM Symp. Theory Comput. (STOC)*, 2011, pp. 529–538.

[41] T. Roughgarden, "Intrinsic robustness of the price of anarchy," *J. ACM*, vol. 62, no. 5, p. 32, 2015.

[42] D. Monderer and L. S. Shapley, "Potential games," *Games Econ. Behav.*, vol. 14, no. 1, pp. 124–143, 1996.

[43] X. Chen, X. Deng, and S.-H. Teng, "Settling the complexity of computing two-player Nash equilibria," *J. ACM (JACM)*, vol. 56, no. 3, p. 14, 2009.

[44] D. S. Johnson, C. H. Papadimitriou, and M. Yannakakis, "How easy is local search?" *J. Comput. Syst. Sci.*, vol. 37, no. 1, pp. 79–100, 1988.

[45] A. Itai, C. H. Papadimitriou, and J. L. Szwarcfiter, "Hamilton paths in grid graphs," *SIAM J. Comput.*, vol. 11, no. 4, pp. 676–686, 1982.



**Omkar Thakoor** received the B.Tech. degree from IIT Bombay, Mumbai, India, in 2014, and the M.S. degree in computer science from the University of Illinois at Urbana–Champaign, Urbana, IL, USA, in 2017. He is currently pursuing the Ph.D. degree with the Teamcore Research Group, University of Southern California, Los Angeles, CA, USA.

His current research interests include artificial intelligence (AI) and theoretical computer science, algorithmic and computational game theory, and a wide variety of applications such as unmanned

aerial vehicle routing, learning in markets, and cyber security, all employing game-theoretic modeling.



**Jugal Garg** received the Ph.D. degree from IIT Bombay, Mumbai, India, in 2012.

He was a Post-Doctoral Fellow with Max-Planck-Institute for Informatics, Saarbrücken, Germany, and Georgia Tech, Atlanta, GA, USA. He is currently an Assistant Professor of industrial and enterprise systems engineering and an Affiliate Assistant Professor with the Department of Computer Science, University of Illinois at Urbana–Champaign, Urbana, IL, USA. His current research interests include computational and strategic aspects of equilibria in game

theory and economics, and their connections with dynamical systems and learning, and design and analysis of algorithms, optimization, and mathematical programming.



**Rakesh Nagi** received the B.E. degree in mechanical engineering from IIT Roorkee, Roorkee, India, in 1987, and the M.S. and Ph.D. degrees in mechanical engineering from the University of Maryland at College Park, College Park, MD, USA, in 1989 and 1991, respectively.

He served as the Chair from 2006 to 2012 and a Professor of industrial and systems engineering with the University at Buffalo, the State University of New York, Buffalo, NY, USA, from 1993 to 2013.

He was with the Institute for Systems Research at the University of Maryland, and INRIA-Lorraine, France. He is currently a Willett Professor and the Head of the Industrial and Enterprise Systems Engineering Department, University of Illinois at Urbana–Champaign, Urbana, IL, USA. He has authored or co-authored more than 200 journal and conference publications. His current research interests include production systems, applied/military operations research, and data fusion using graph theoretic models.