


A Heuristic Learning Algorithm for Preferential Area Surveillance by Unmanned Aerial Vehicles

Manickam Ramasamy · Debasish Ghose 

Received: 31 August 2016 / Accepted: 19 January 2017
© Springer Science+Business Media Dordrecht 2017

Abstract A heuristic learning algorithm is presented in this paper to solve the problem of persistent preferential surveillance by an Unmanned Aerial Vehicle (UAV). The algorithm helps a UAV perform surveillance as per quantitative priority specifications over a known area. It allows the specification of regional priorities either as percentages of visitation to be made by a UAV to each region or as percentages of surveillance time to be spent within each. Additionally, the algorithm increases the likelihood of target detection in an unknown area. The neighborhood of a detected target is suspected to be a region of a high likelihood of target detection, and the UAV plans its path accordingly to verify this suspicion. Similar to using the target information, the algorithm uses the risk information to reduce the frequency of visits to risky regions. The technique of using risk map to avoid risky regions is adapted from the existing geometric reinforcement

learning technique. The effectiveness of this algorithm is demonstrated using simulation results.

Keywords Persistent surveillance · Preferential surveillance · Learning · UAV · Target detection

Mathematics Subject Classification 2010 MSC 68T05 · MSC 68T20

1 Introduction

External, internal security and commercial applications such as combat intelligence missions, riot monitoring, and traffic control demand the need for continuous aerial surveillance over a given area. Unmanned Aerial Vehicles (UAVs) are potential candidates for such persistent aerial surveillance applications. The persistent surveillance problems are otherwise called as patrolling or continuous sweep coverage problems. These problems demand a UAV to observe the entire area continuously such that the maximum age, that is, the maximum unobserved time of any part of that area is kept to a minimum over the time of surveillance [1]. In other words, the UAV has to plan its path at every instant of time, based on the age of every part of the surveillance area. Given an optimal solution to this UAV path planning problem, it is not possible to verify the solution in polynomial time. As the persistent surveillance problem is NP-hard, suboptimal solution

This paper is an extension of the work presented by the authors in [24].

D. Ghose (✉)
Department of Aerospace Engineering, Indian Institute of Science, Bangalore, 560012, Karnataka, India
e-mail: dghose@aero.iisc.ernet.in

M. Ramasamy
Aeronautical Development Establishment, Defence Research and Development Organization, Bangalore, 560093, Karnataka, India
e-mail: manickam@ade.drdo.in

techniques have been widely used to solve it. These problems have been solved using reactive policies [1], heuristics [2], approximate dynamic programming [3], exact methods from operations research [4], graph theory [5], optimal control [6], memoryless control design [7], automata-based techniques [8], and bio-inspired strategies [9]. They have been formulated with aircraft performance constraints [6], UAV system health and safety constraints [1–3, 7], communication constraints [3], and battery charging time and temporal logic constraints [8]. A detailed review of the persistent surveillance problems has been presented in [10].

It is often desired to perform persistent and preferential area surveillance by using prior information or on-the-fly estimate acquired about the spatial distribution of targets and threats. In general, the problems mentioned above have been formulated with predefined regions of interest and risk. Practically, the prior information of a surveillance environment is incomplete and has uncertainties. To perform any mission in such environments, the UAV should have the capability to learn about the environment on-the-fly and use the acquired information to plan its future actions. For a persistent surveillance mission, the UAV has to learn the spatial distribution of targets and threats dynamically, identify the regions over the surveillance area as priority regions and risky regions, accordingly, and use this information to plan its path at any instant of time.

Learning techniques find many applications in solving aerial surveillance problems. Supervised, unsupervised, semi-supervised, and reinforcement learning techniques constitute the broad classification of machine learning techniques [11, 12]. They are applied to solve classification, clustering, practical learning and goal-oriented agent problems, respectively. A geometric reinforcement learning technique has been developed to solve the path planning problem of collaborative UAVs, given their source and destination points [13]. This technique formulates an objective function which takes into account, the distance to be traveled by the UAV and a risk map of the surveillance environment. The risk encountered by a UAV may be due to threats such as surface-to-air missile sites or anti-aircraft guns, or another aircraft. But the fact that learning happens through exploration could jeopardize the safety of the UAVs. Integrated

cooperative planning and learning techniques have been developed to curtail the risk associated with such learning process [14].

A machine learner in the name of sequence memoizer has been developed so that the UAVs can track targets in urban environment [15]. It predicts the future location of targets using previous observations. When the UAVs have to track dynamic targets in an urban environment, constraints emerge in their line-of-sight communication capabilities and this problem has been addressed using learning [16]. The health models of a group of heterogeneous agents have been learned in real-time using Decentralized Multi-agent Markov Decision Processes (Dec-MMDP) and incremental Feature Dependency Discovery (iFDD) method [17].

An opportunistic learning method has been developed to combine the performance advantage of centralized learning and the efficiency of decentralized learning [18]. Probability based learning techniques [19] have been developed along with game-theoretic concepts such as negotiation [20] and auction mechanisms to solve UAV cooperative control problems. Motivated learning algorithms have been used with human-inspired autonomous controllers so that the agents can autonomously set goals, improvise, innovate and imagine [21]. A distributed learning and planning method has been developed to learn target locations in an uncertain environment using an online approximation function with modifiable edge weights [22].

The application of learning techniques to persistent surveillance problems is very limited. Active learning is used in a Markov Decision Process (MDP) framework to learn fuel consumption model parameters so as to maintain the required number of UAVs over the persistent surveillance area [23]. This problem focuses on the transition from base station to the surveillance area rather than learning about the surveillance environment itself. A learning-based preferential surveillance algorithm is presented in [24]. The primary objective in this problem is to develop an algorithm using learning to increase the probability of target detection in a surveillance environment with dynamic targets and uncertain threats. The most significant difference in this work compared to [24] is that more attention is dedicated towards the preferential surveillance problem over a known area. The solution to this problem is built considering the different ways

of quantitative priority specifications. The solution methodology is presented more clearly in this paper. In addition to this, simulation results are extended to support a more detailed analysis. We strongly believe that a more detailed analysis of the preferential surveillance problem over a known area would help in improving the learning algorithm for the target detection problem in an unknown area.

In this study, a heuristic learning algorithm is developed to solve the problem. The surveillance area on the ground is discretized into a 2D planar grid. In the mission scenario considered to demonstrate the effectiveness of the algorithm, a target vehicle is assumed to enter the surveillance area, follow a predefined trajectory, exit the surveillance area, and repeat its trajectory. The algorithm works as follows. When a UAV is flying above a particular grid point, if the target is detected, the UAV allocates higher preference to that point and a discounted preference to the 8-connected neighborhood. Then, the UAV plans its path such that it visits these points more often than the rest of the grid points due to the preferential treatment. The absence of target detection in any neighborhood point in the subsequent time steps removes the preference to that point. Over successive target episodes, the algorithm helps in increasing target detection probability. The algorithm motivates the UAV to avoid flying over threats to a similar effect.

This paper makes the following contributions. (1) It proposes a new heuristic learning algorithm that performs preferential surveillance as per the requirement in a known environment and increases the probability of target detection in an unknown environment. (2) It applies the integral risk computation between the current location of UAV and all other grid points to compute a risk map and thereby motivates the UAV to avoid risky regions.

2 Problem Definition

In an $N \times M$ 2D grid world, a UAV has to minimize the maximum age of a region given by,

$$\max_{(i,j) \in R} A_{i,j} \quad (1)$$

$$R : N \times M \in \mathbb{Z}^{+2}, i \in \{1, \dots, N\} \subset \mathbb{Z}^+, j \in \{1, \dots, M\} \subset \mathbb{Z}^+$$

at any time t subject to visitation frequency constraints,

$$F_{R_P} = M_{R_P} F_{R_1} \quad (2)$$

$$M_{R_P} > 1, R_P \subset R, R_1 = R \setminus R_P$$

$$F_{R_T} = M_{R_T} F_{R_2} \quad (3)$$

$$0 < M_{R_T} < 1, R_T \subset R_1, R_2 = R_1 \setminus R_T$$

where, A_R represents the age map, that is, the $N \times M$ matrix of ‘time elapsed since the last visit’ made by the UAV to each grid point [1] (this definition implies that the value of $A_{p,q}$ goes to zero when a UAV is at the point (p, q)), R denotes the surveillance area, (i, j) refers to a grid point belonging to R , R_P is the set of ordered pairs (i, j) belonging to the preferred region which is strictly a subset of the surveillance area R , F_r refers to the visitation frequency to a certain region r , R_1 denotes the surveillance area without the preferred region, M_{R_P} denotes the relative visitation frequency multiplication factor for the preferred (hence, greater than 1) region, R_T is the set of ordered pairs (i, j) belonging to the threat or risky region which is strictly a subset of R_1 , that is, the preferred and risky regions are mutually exclusive, R_2 denotes R_1 without the risky region, and M_{R_T} denotes the relative visitation frequency multiplication factor for the threat (hence, lesser than 1) region. Visitation frequency is defined as the number of visits to a region per grid point of that region within a certain time. M_{R_T} , being lesser than 1 for a threat region, does not imply that high surveillance time over a threat region is impossible. If a UAV has to collect more data from a certain threat region for neutralizing it, this region has to be defined as the preferred region of surveillance.

The decision variable in this problem is the choice of destination to be made by the UAV at every instant of time. Minimization of the maximum age is carried out by choosing a UAV’s destination point, $(i_d, j_d) \in R$, at a time, t , such that the maximum age of the region at the time, $t + \delta t$, is minimum. Here, δt is the time taken by the UAV to travel from its current position, $(x_u, y_u) \in R$, to the destination point, (i_d, j_d) .

The above is the generalized problem formulation for the successively more complex scenarios presented in this paper. A UAV has the objective to minimize the age of any grid point belonging to the discretized planar grid over the surveillance area. Simultaneously,

the UAV has to increase the visitation frequency to a preferred region, R_P , and decrease the visitation frequency to a region with threats, R_T . This problem can be presented under two different conditions as follows. (i) The preferred and threat regions are predefined. The degree of preference of region, R_P , is specified as a percentage of visits to be made to the grid points in that region or as a percentage of surveillance time to be spent within it. (ii) The regions, R_P and R_T , have to be categorized by the UAV on-the-fly. They can be dynamically changing, and uncertainty exists about the risk posed by threats in the region, R_T . The second formulation is more practical than the first one.

3 Solution Methodology

The solution to the considered problem is built progressively. Initially, persistent surveillance problem with no preferred and threat regions is solved using heuristics. Subsequently, a preferential surveillance algorithm is developed to address the problem with predefined fixed target regions. Then, a preferential surveillance problem is considered with predefined fixed target regions and initially unknown threat regions. Finally, a learning algorithm is developed to address the problem with unknown, dynamic target regions and initially unknown threats. This solution method makes the following assumptions. (i) UAV planar motion (that is, at a constant altitude) over the surveillance area has been modeled using kinematic equations of motion as given by Eqs. 4 and 5 without considering turn rate constraint.

$$\dot{x} = V \cos \psi \quad (4)$$

$$\dot{y} = V \sin \psi \quad (5)$$

where, (x, y) is an ordered pair representing the UAV position in the grid world, V is the vehicle speed and ψ is the vehicle heading. Such a simplistic model assumption does not inhibit the demonstration of the use of the learning algorithm, which is the primary objective of this work. (ii) Effects on aircraft motion due to the environmental effects such as winds and turbulence are not considered. (iii) Grid points are spaced at one unit distance from each other as shown in Fig. 1. (iv) The UAV payload is assumed to be electro-optic having a circular footprint over the ground. The

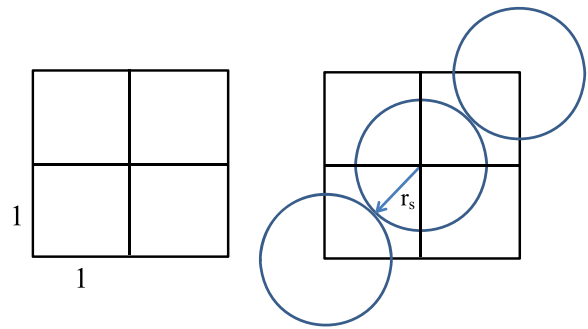


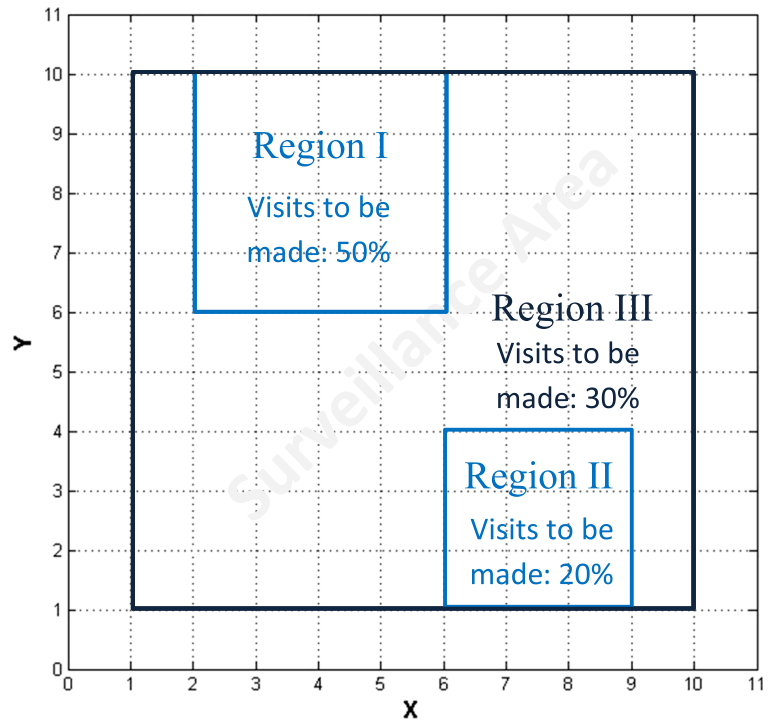
Fig. 1 Assumptions made in the problem

sensing radius, r_s , is taken as $1/\sqrt{2}$ units as shown in Fig. 1. This assumption ensures that the sensor footprint perfectly encompasses an 1×1 square and justifies formulating the problem in a grid world. (v) UAV speed is taken to be 1 unit distance per unit time. (vi) A grid point is said to have been visited by the UAV when the latter passes within a radial distance of 0.05 units measured from that grid point. (vii) The imaging sensor is assumed to detect a target or a threat instantaneously when it is present within the sensor footprint. It can be seen that the problem formulation and the solution approach are realistic with respect to the mission environment and not the vehicle itself.

3.1 Persistent Surveillance Problem

The objective of this problem is to minimize the maximum age of any grid point in the surveillance area as given by Eq. 1. Given an optimal path solution for this problem, it is not possible in polynomial time to verify whether the path is indeed optimal. Heuristics has been used to solve this NP-hard optimization problem. An intuitive solution would be to plan the path of the UAV so as to reach the point with maximum age at every instance. However, when a distant point is set as the destination point, the UAV incurs certain time reaching it, and this increases the maximum age of the surveillance area. Therefore, the heuristic function chosen is such that the UAV decides its destination point by making a tradeoff between the age of that point and distance to that point. The cost function is given in Eq. 6 (which is similar to the method used in [1]). A greedy strategy is adopted by the UAV to set the grid point with maximum cost as the destination point as given by Eq. 7. This solution

Fig. 2 Illustration of preferential surveillance problem with known, fixed target regions



method forms the base for the progressively complex scenarios discussed in Sections 3.2 to 3.4.

$$C_{i,j} = A_{i,j} - K_W W_{i,j} \quad (6)$$

$$p^t = \max_{(i,j)} C_{i,j} \quad (7)$$

$$K_W \geq 0, i \in \{1, ..N\} \subset \mathbb{Z}^+, j \in \{1, ..M\} \subset \mathbb{Z}^+$$

where, p^t is the destination point at time t , $C_{i,j}$ constitutes the cost matrix, $A_{i,j}$ denotes the time elapsed

since the last visit to the point (i, j) , $W_{i,j}$ constitutes the weight matrix, and K_W is the weight multiplication factor that affects the trade-off between age and distance. The destination point, defined in Eq. 7, is determined by a tiebreaker under certain conditions. If two grid points have the maximum cost, the UAV chooses the one that has lesser heading difference with respect to its current heading. If two grid points have the maximum cost and the same heading difference with respect to UAV's current heading, the destination

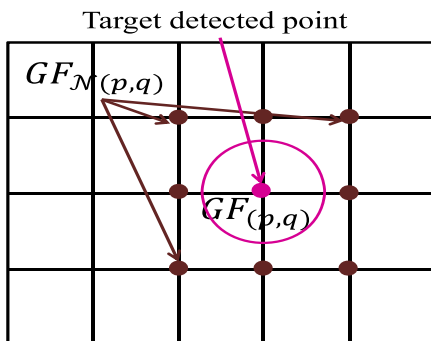


Fig. 3 Neighborhood exploration by UAV through age growth factor

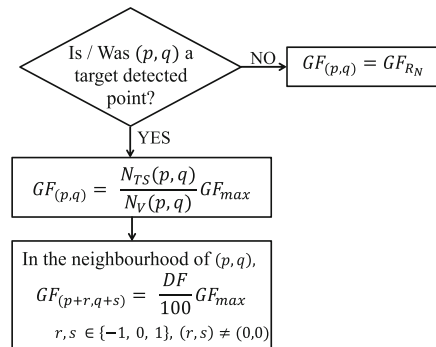


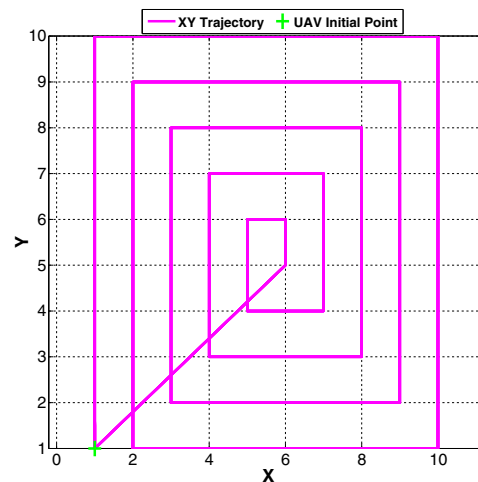
Fig. 4 Heuristic learning algorithm

point is chosen randomly. $W_{i,j}$ is defined as the time taken by the UAV to travel from its current location (x_u, y_u) to a point (i, j) . As the UAV speed is assumed to be 1 unit distance per unit time, $W_{i,j}$ is given by the distance, $d_{i,j}$, from current location of the UAV to a point (i, j) . K_W is non-negative, because, if there exist two grid points (i_1, j_1) and (i_2, j_2) at respective distances d_{i_1, j_1} and d_{i_2, j_2} such that $d_{i_1, j_1} > d_{i_2, j_2}$ and these points have the same age, it is not desired for the UAV to allot higher preference to the grid point (i_1, j_1) which is farther to the UAV than (i_2, j_2) . This follows from the fact that an unnecessary travel to a distant point increases the maximum age of the surveillance area. Additionally, by the definition of age, $A_{i,j} = 0$ at time t , if $(i, j) = (x_u, y_u)$ at time t .

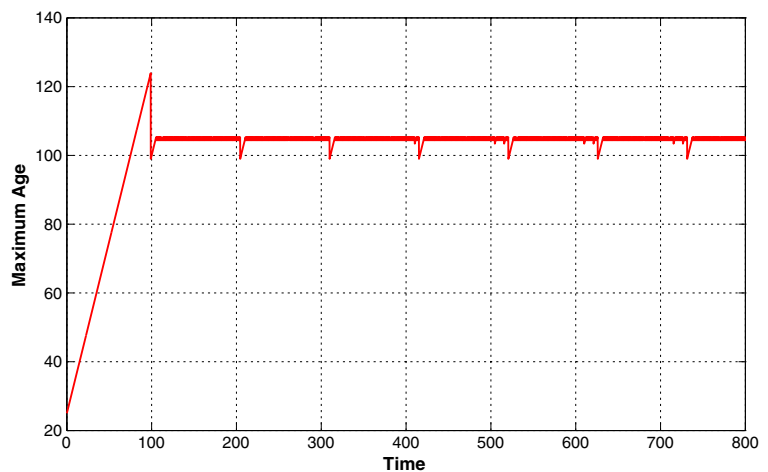
3.2 Preferential Persistent Surveillance with Known, Fixed Target Regions

In this case, the problem objective is to minimize the maximum age given by Eq. 1 subject to the constraint given by Eq. 2. The preferred regions are predefined and stationary. The priority of visitation to the preferred regions is specified in terms of their relative importance. An example problem is shown in Fig. 2. It should be noted that the relative importance can be specified quantitatively as the percentage of visitations to be made to each region or as the percentage of surveillance time to be spent within each region. These two are different as the former requirement specification is more appropriate for a discretized domain

Fig. 5 Results of Test Case 1: Persistent surveillance with no threats and targets for UAV initial position (1, 1)



(a) Path followed by the UAV



(b) Maximum age versus time

problem whereas the latter is more relevant in the continuous domain. However, the performance of the algorithm, which is developed in discrete setting, for a time based surveillance specification is also presented in this paper.

Let the entire surveillance area, preferred region, and threat region be represented by R , R_P , and R_T , respectively. Let R_N be given by $R_N = R \setminus R_P$, which implies that the rest of the surveillance area is non-preferred. For the UAV to perform preferential surveillance over R_P , the rate at which the age grows at the grid points $(i, j) \in R_P$ is made relatively higher than that of other regions. The factor by which the age growth rate is relatively increased depends on the specified relative significance of the preferred region.

The changes in age of different regions over a time interval are given by Δt Eqs. 8, 9, and 10.

$$A_{i,j} = A_{i,j} + GF_{R_N} \Delta t \quad (i, j) \in R_N, (i, j) \neq (x_u, y_u) \quad (8)$$

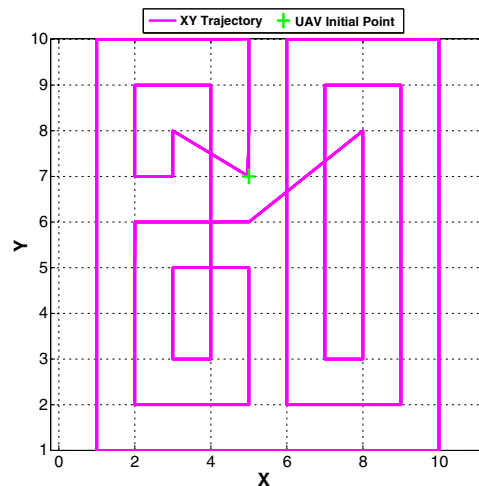
$$A_{i,j} = A_{i,j} + GF_{R_P} \Delta t \quad (i, j) \in R_P, (i, j) \neq (x_u, y_u) \quad (9)$$

$$A_{i,j} = 0 \quad (i, j) = (x_u, y_u) \quad (10)$$

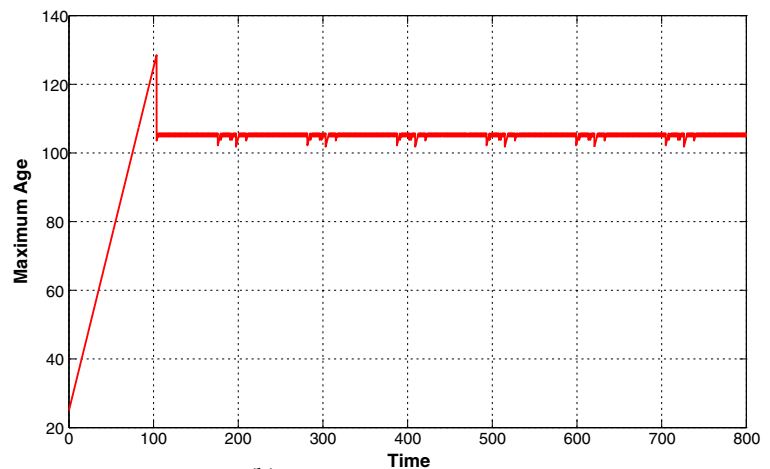
where, GF_{R_N} and GF_{R_P} represent the age growth factors of grid points belonging to non-preferred and preferred regions, respectively and (x_u, y_u) denotes the current location of the UAV.

The age growth factors have to be computed to meet the visitation or temporal requirements for preferential surveillance. The persistent surveillance problem is solved by allotting a particular grid point to

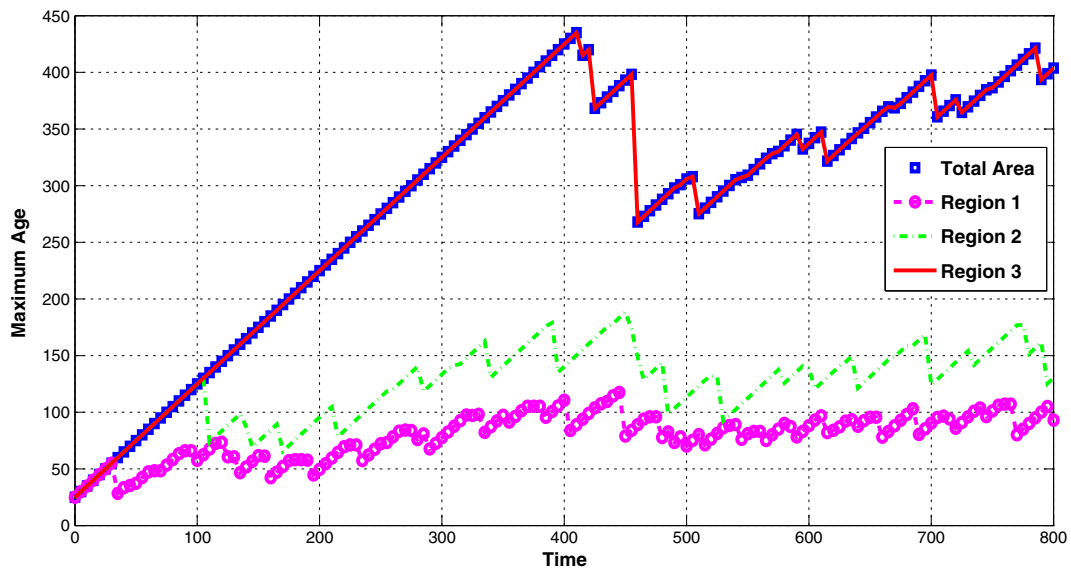
Fig. 6 Results of Test Case 1: Persistent surveillance with no threats and targets for UAV initial position (5, 7)



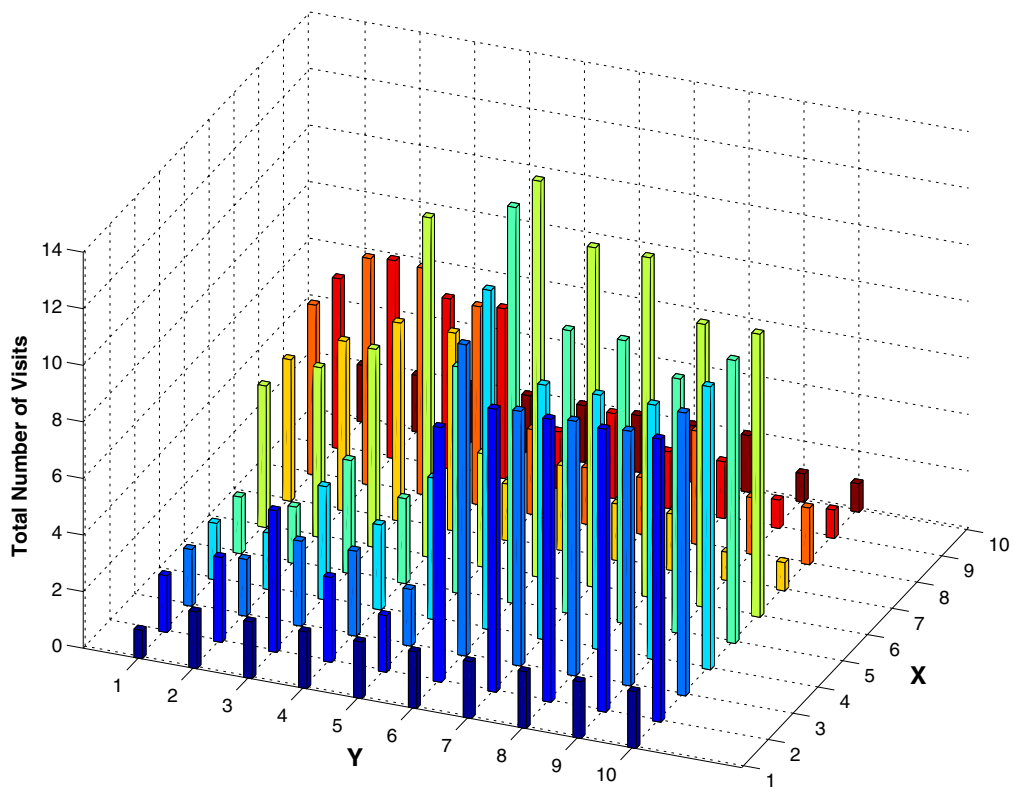
(a) Path followed by the UAV



(b) Maximum age versus time



(a) Maximum age versus time



(b) Total number of visits

Fig. 7 Results of Test Case 2: Preferential surveillance with known, fixed target regions and no threats

Table 1 Number of visits to different regions

	Region I	Region II	Region III	Whole Area
Number of Visits	257	109	140	506
Percentage of Visits	50.79	21.54	27.67	100

the UAV as its destination point at any time instance. Hence, the requirement of visitation percentages to grid points falls in line with the solution approach. Let V_{R_P} and V_{R_N} denote the percentages of visits required to be made to the grid points in the regions, R_P and R_N , respectively. Therefore, the number of grid points in regions, R_P and R_N , become important in computing the age growth factors of those regions. Let these numbers be denoted by N_{R_P} and N_{R_N} , respectively. Then, the ratio of GF_{R_P} to GF_{R_N} is given by Eq. 11. It can be seen that the right-hand side of the equation contains the ratio of required visitation percentages of regions (R_P and R_N) normalized using the number of grid points in the respective regions.

$$\frac{GF_{R_P}}{GF_{R_N}} = \frac{(V_{R_P}/N_{R_P})}{(V_{R_N}/N_{R_N})} \quad (11)$$

Let us take the case where the relative preference between different regions is specified as the percentage of surveillance time to be spent within each region. This case necessitates the UAV flight time within different regions to meet the temporal requirements. The problem is formulated in grid world. Hence, the solution technique adheres to this requirement indirectly. It represents any region through the grid points in that region and allocates those grid points as destination points to the UAV appropriately. In this case, the area of the regions, R_P and R_N , are used to compute the age growth factors of those regions. Let these be denoted by S_{R_P} and S_{R_N} , respectively. Let T_{R_P} and T_{R_N} denote the percentages of surveillance time required to be spent within regions, R_P and R_N , respectively. Similar to Eq. 11, the age growth factor ratio is given by Eq. 12.

$$\frac{GF_{R_P}}{GF_{R_N}} = \frac{(T_{R_P}/S_{R_P})}{(T_{R_N}/S_{R_N})} \quad (12)$$

Table 2 Surveillance time spent in different regions

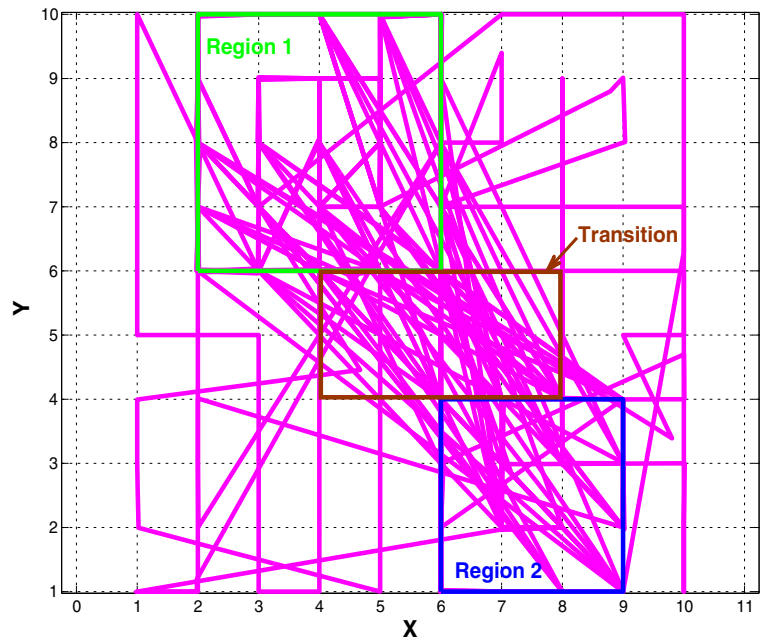
	Region I	Region II	Region III	Whole Area
Time units spent in region	270.7	143	386.4	800
Percentage of Time	33.84	17.87	48.3	100

It should be noted that preferential surveillance is formulated as an additional requirement in both the cases discussed above and is not imposed as a constraint. Additionally, as the relative preference is only important in both the cases, GF_{R_N} is set to a fixed value (say, 1) and GF_{R_P} is then computed. These values are used in Eqs. 8 and 9 to update the age of all grid points. Equations 6 and 7 are then used to determine the next destination point for the UAV to visit.

3.3 Preferential Surveillance Problem with Known, Fixed Target Regions and Initially Unknown Threats

The objective of this problem is to minimize the maximum age given by Eq. 1 subject to the constraints given by Eqs. 2 and 3. The threat positions are initially unknown to the UAV. Once the UAV senses a threat at some point, it has to avoid visiting or even flying over that point. For this purpose, risk posed by the threat is modeled using a normal distribution around that point as in geometric reinforcement learning [13] technique. Risk exposure calculation due to multiple threats in the surveillance environment result in the generation of a risk map. The UAV uses this risk map to avoid threats and the integral risk factor defined in Eq. 15 determines the degree of risk avoidance. In addition to using this existing method of integral risk computation, our algorithm enables the UAV to perform preferential surveillance and to increase target detection probability through learning. The integral risk computation method is discussed in detail in [13]. The basics of this method are presented here for easy reference. The probability of a UAV being at risk due to a threat is modeled using a normal distribution as shown in Eq. 13. The probability distribution is attributed to

Fig. 8 Result of test case 2:
UAV trajectory for
time-based preferential
surveillance



the uncertainties in the threat position or in the risk posed by that threat.

$$f_m(x, y) = \frac{1}{\sigma_m \sqrt{2\pi}} e^{-((x-x_m)^2 + (y-y_m)^2)/2\sigma_m^2} \quad (13)$$

where, m is the threat index, (x_m, y_m) is threat position, σ_m is the standard deviation of probability distribution of risk around (x_m, y_m) , (x, y) is the point where risk exposure is computed and $f_m(x, y)$ is the probability of exposure to risk at (x, y) due to a threat m .

If multiple threats are present in the environment (let the number of threats be denoted by M), the total probability of exposure to risk at (x, y) is given by Eq. 14.

$$F(x, y) = 1 - \prod_{m=1}^M [1 - f_m(x, y)] \quad (14)$$

$W_{i,j}$ in Eq. 6 is given by the sum of two terms in this preferential surveillance problem with initially

unknown threats. The first term, $d_{i,j}$, represents the euclidean distance between a point (i, j) and the current location of the UAV (x_u, y_u) and the second term is the integral of the total probability of exposure to risk along a path C between (x_u, y_u) and (i, j) as shown in Eq. 15.

$$W_{i,j} = d_{i,j} + K_{IR} \int_C F(x, y) ds \quad (15)$$

where, K_{IR} is the integral risk factor. This is a measure of the degree of risk posed by the threats and thereby the degree of avoidance of threats by the UAV. The calculation of $W_{i,j}$ is given by Eq. 15 and that of $A_{i,j}$ by Eqs. 8 and 9. The next destination point for the UAV is then determined using Eqs. 6 and 7. As the risk map is computed at every instant of time, this method works well even in the case of an environment with dynamic threats.

Table 3 Surveillance time spent in different regions including transition region

	Region I	Region II	Region III (without Transition region)	Transition region	Whole Area
Time units spent in region	270.7	143	242.8	143.6	800
Percentage of Time	33.83	17.87	30.35	17.95	100

Table 4 Surveillance time spent in two different regions

	Region I	R_N	Whole Area
Time units spent in region	383.2	416.9	800
Percentage of Time	47.89	52.11	100

3.4 Preferential Surveillance Problem with Unknown, Dynamic Target Regions and Initially Unknown Threats

In this case, the problem objective is for the UAV to identify the regions of interest on-the-fly, avoid

the threat regions, and increase the probability of target detection. This problem is more practical than the cases discussed above. The threat environment is assumed to have uncertainties. When the UAV is at a certain grid point and if a target falls within the UAV's sensor coverage area on the ground, two conditions are possible as follows. (i) The UAV certainly detects a target at that point and (ii) The UAV detects a target at that point based on a specified probability. A new learning algorithm called as a heuristic learning algorithm has been developed to solve this problem under these two conditions. The following steps form a part

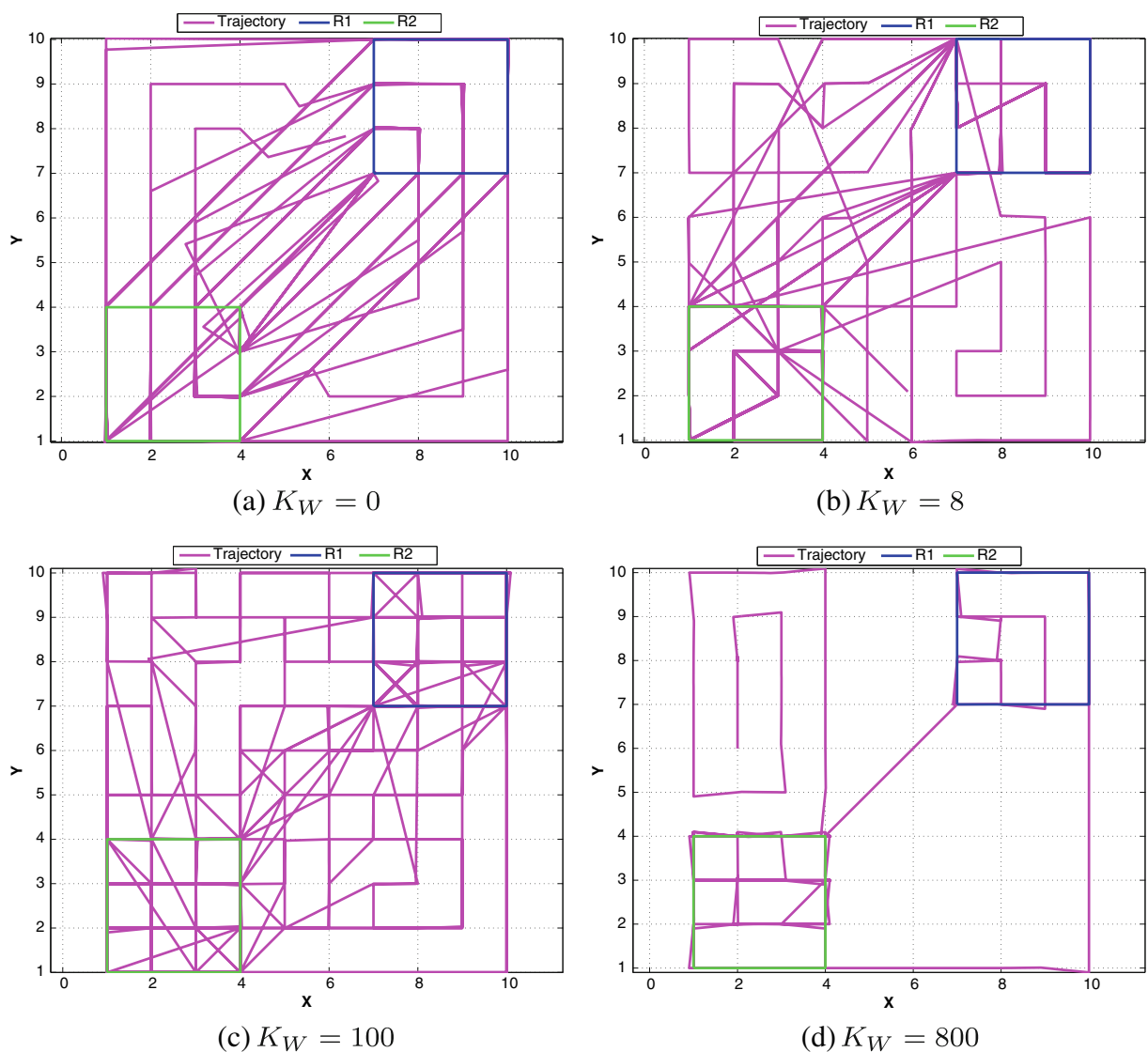
**Fig. 9** Results of Test Case 2: UAV's planar trajectory for different values of K_W

Table 5 Percentage of visits to different regions for different K_W

	Region I	Region II	Region III
$K_W = 0$	26.50	29.20	44.30
$K_W = 0.3$	42.24	42.74	15.02
$K_W = 4$	39.74	41.86	18.40
$K_W = 60$	33.48	41.60	24.92
$K_W = 100$	30.00	41.00	29.00
$K_W = 800$	22.06	50.74	27.20

of this algorithm along with the solution technique described in Section 3.3.

- Step 1: On detection of a target by the imaging sensor of a UAV as it flies over a grid point (p, q) , the algorithm increases the age growth factor of that point to the specified maximum growth factor value, GF_{max} , as shown in Eq. 16.
- Step 2: Suspecting the 8-connected neighborhood of the point, where the target was detected, to possess additional targets, the algorithm increases the age growth factors of those points too. But the difference is that these factors are increased to a discounted percentage, DF , of GF_{max} . This idea is pictorially shown in Fig. 3, where, $\mathcal{N}(p, q)$ denotes the neighborhood of point (p, q) .

$$GF_{(p+r, q+s)} = \begin{cases} GF_{max}, & \text{if } r = s = 0 \\ \frac{DF}{100} GF_{max}, & \text{else} \end{cases} \quad (16)$$

$$r, s \in \{-1, 0, 1\},$$

(p, q) is a point where a target is or was detected

- Step 3: The UAV chooses the grid point with maximum cost as its destination point at any time as given by Eq. 7. The previous step results in an increase in the cost values of the neighborhood points as given by Eq. 6. Hence, a neighborhood point is visited by the UAV eventually. If no target is detected at this point, the age growth factor of this point is reset to GF_{R_N} , as shown in Eq. 17. By renouncing the preference allocated to such a

point, the algorithm enables the UAV to learn about the target environment appropriately.

$$GF_{(p, q)} = GF_{R_N} \quad (17)$$

(p, q) is and has not been a target detection point. The above steps imply that, even if a target has been detected at a point just once, the age growth factor of that point is set to GF_{max} and retained. But there is a need to renounce the significance of a point where target was detected just once out of multiple visits. This is made possible by allocating preference to a point based on the target detection probability at that point inferred from the past. Hence, the age growth factor formulation in Eq. 16 is modified as shown in Eq. 18.

$$GF_{(p+r, q+s)} = \begin{cases} \frac{N_{TS}(p, q)}{N_V(p, q)} GF_{max}, & \text{if } r = s = 0 \\ \frac{DF}{100} GF_{max}, & \text{else} \end{cases} \quad (18)$$

where, (p, q) is the point where target is or was detected, $r, s \in \{-1, 0, 1\}$, $N_{TS}(p, q)$ denotes the number of times any target was sighted at (p, q) , and $N_V(p, q)$ represents the number of visits made to (p, q) . This algorithm is presented in the form of a flowchart in Fig. 4. This algorithm is suitable for a preferential surveillance problem in both static and dynamic environments.

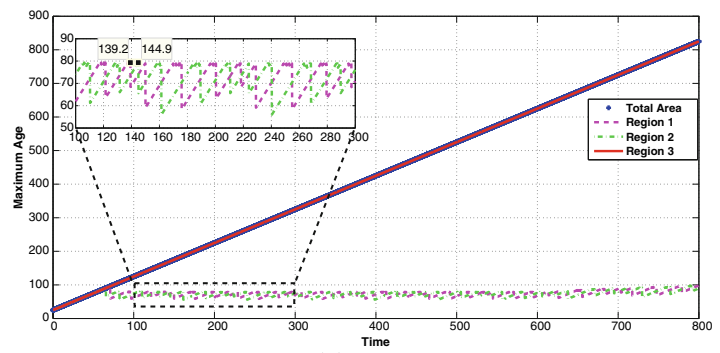
4 Simulation Results and Analysis

The algorithm is implemented and tested on a 10×10 grid using MATLAB® to demonstrate its effectiveness for all the four cases discussed above. The simulation conditions are described as follows. The initial point of UAV is (1,1), unless stated otherwise. Threats are simulated for the cases discussed in Sections 3.3 and 3.4. Initial age of all grid points is set to 25. GF_{R_N} and σ_m are set to 1. Simulation is executed at a time interval of 0.1 unit of time. The simulation conditions specific to each case are explained separately.

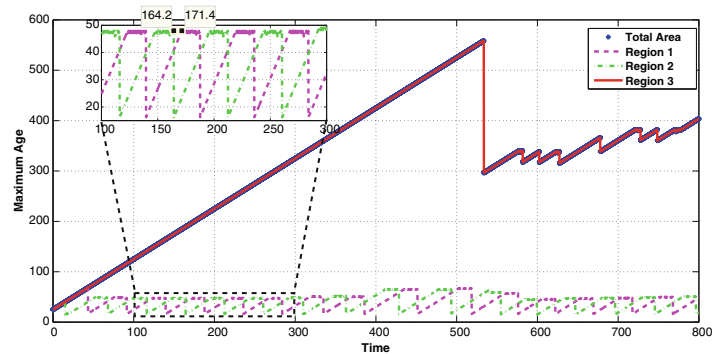
4.1 Test Case 1: Persistent Surveillance with No Targets and Threats

Firstly, the simulation results for a simple persistent surveillance problem with no targets and threats is presented. The simulation was run till 800 units of time.

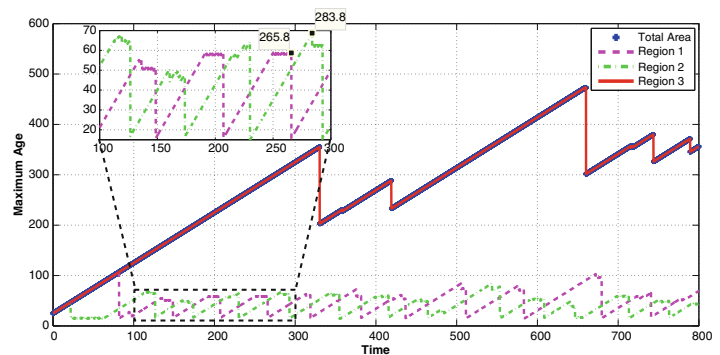
Fig. 10 Results of Test Case 2: Maximum age of regions for different values of K_W



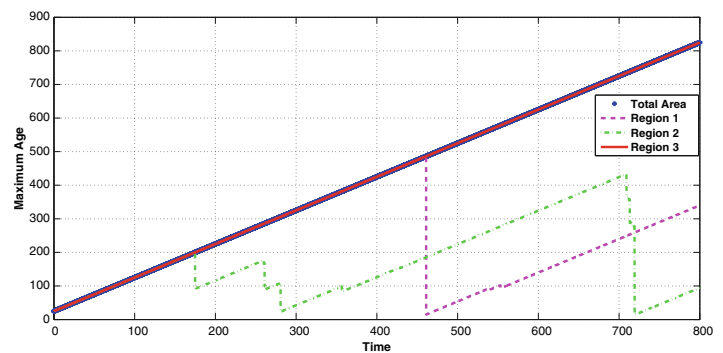
(a) $K_W = 0$



(b) $K_W = 8$



(c) $K_W = 100$



(d) $K_W = 800$

K_W is set to 1. This problem is solved using the method described in Section 3.1. The solution to this problem, that is, the trajectory followed by the UAV, is as shown in Fig. 5a and the maximum age of all grid points with respect to time is shown to stabilize close to 105 units of time in Fig. 5b. The initial peak in maximum age is due to the initial age of all grid points being 25. If the UAV were to jump in zero time from the grid center to initial point during every cycle, the maximum age would have been constantly equal to 99. But, as the UAV has to fly from the center to initial point in finite time, the maximum age increases back to 105 units of time. Additionally, this value of maximum age is dependent on the aspect ratio of the grid. All grid points were found to have a uniform number of visits for each visitation cycle. The dependency of the solution on the trade-off factor between age and distance, K_W , is discussed in the next section. The trajectory followed by the UAV changes based on its initial point. For example, the solution to this problem for the UAV initial point at (5, 7) is shown in Fig. 6a. Invariably, the maximum age of all grid points with respect to time stabilizes close to 105 units of time as depicted in Fig. 6b.

4.2 Test Case 2: Preferential Surveillance with Known, Fixed Target Regions and No Threats

Two persistent surveillance scenarios with known preferred regions and the required percentage of visits to be made to each region are depicted in Figs. 2 and 9. In the first scenario, Region I, Region II, and Region III have 25, 16, and 59 grid points, and the required percentage of visits to be made to these regions are 50 %, 20 %, and 30 %, respectively. This problem has been solved using Eq. 11 discussed in Section 3.2. K_W was set to 1 in this simulation. The simulation was run till 800 units of time. The maxima of the maximum age of the grid points for the different regions are as shown in Fig. 7a. The maximum age of the highest priority region is the least, and that of the least priority region is the highest with respect to time. Hence, the maximum age of the entire region is the same as the maximum age of the least priority region. The simulation results in Table 1 show that the percentage of visits made to each region meets the requirement very closely. However, as seen in Fig. 7b, all points in a certain preferred region are not visited uniformly. Those points on the path of transition from one preferred

Fig. 11 Test Case 3: Visitation requirements and simulation conditions

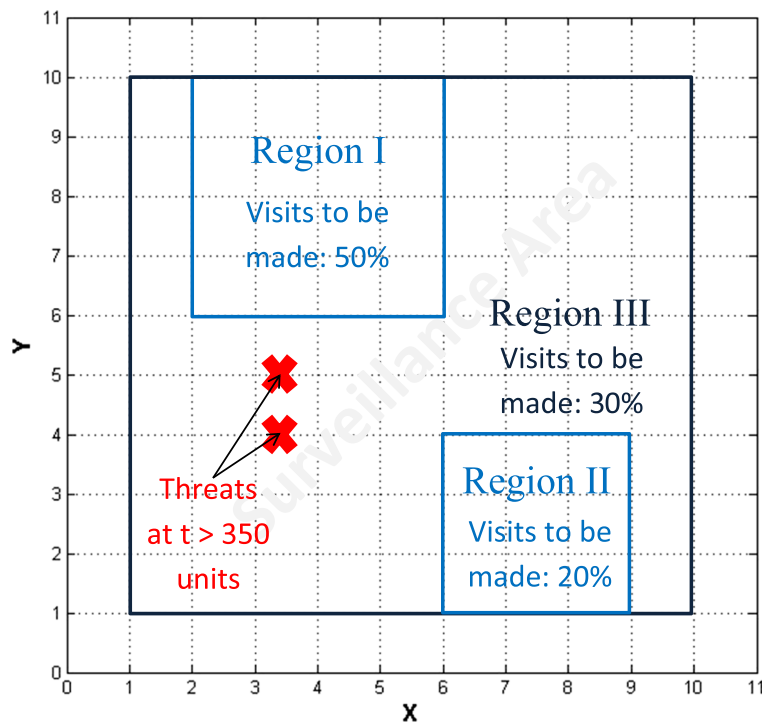
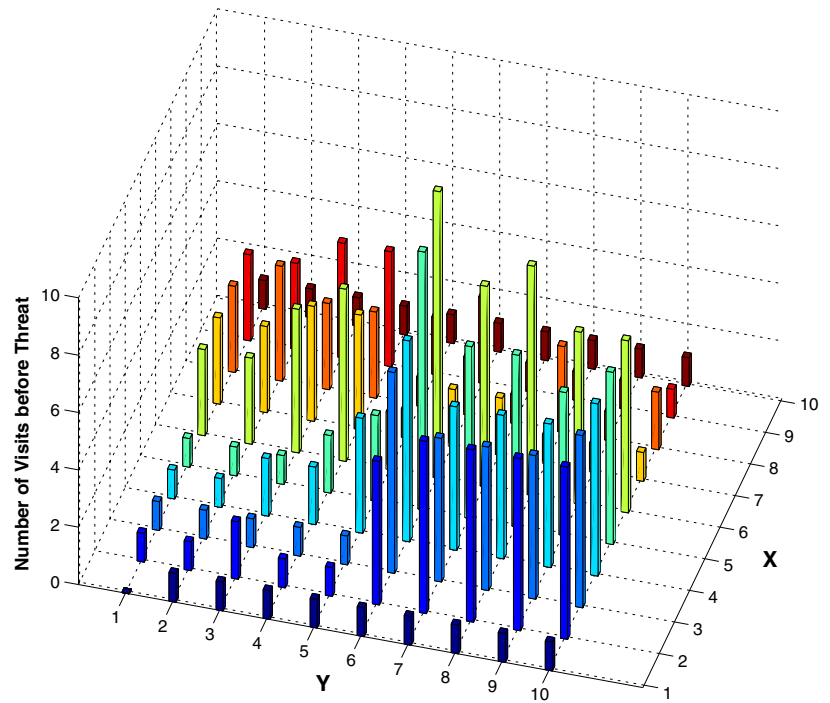
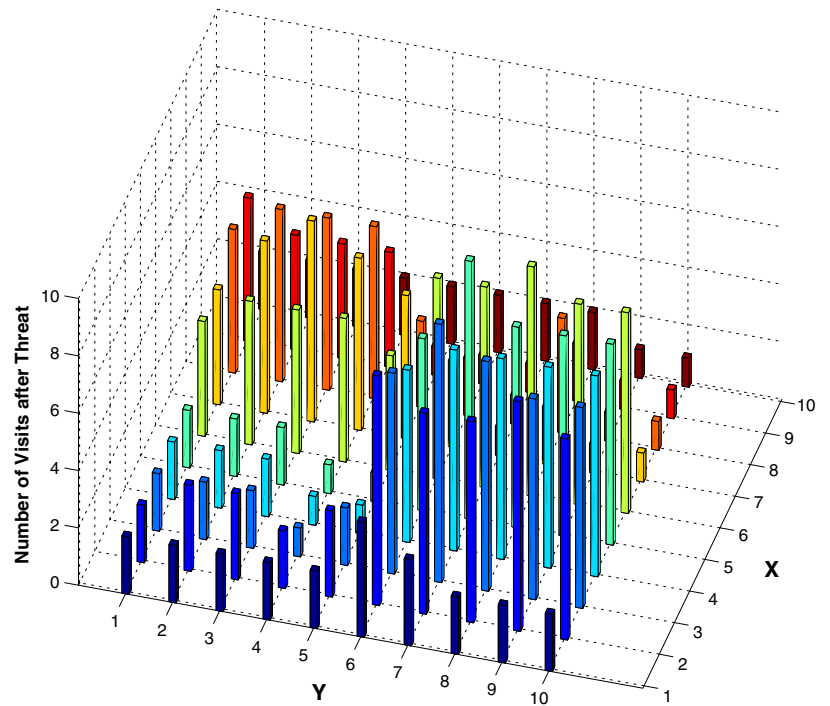


Fig. 12 Results of Test Case 3: Preferential surveillance with known, fixed target regions and initially unknown threats



(a) UAV visits before uncertain threat $K_{IR} = 200$



(b) UAV visits after uncertain threat $K_{IR} = 200$

Table 6 Number of UAV visits to different regions after threat introduction

	Region I	Region II	Region III	Whole Area
Number of visits	181	82	115	378
Percentage of visits	47.89	21.69	30.42	100

region to another, record a higher number of visits by the UAV.

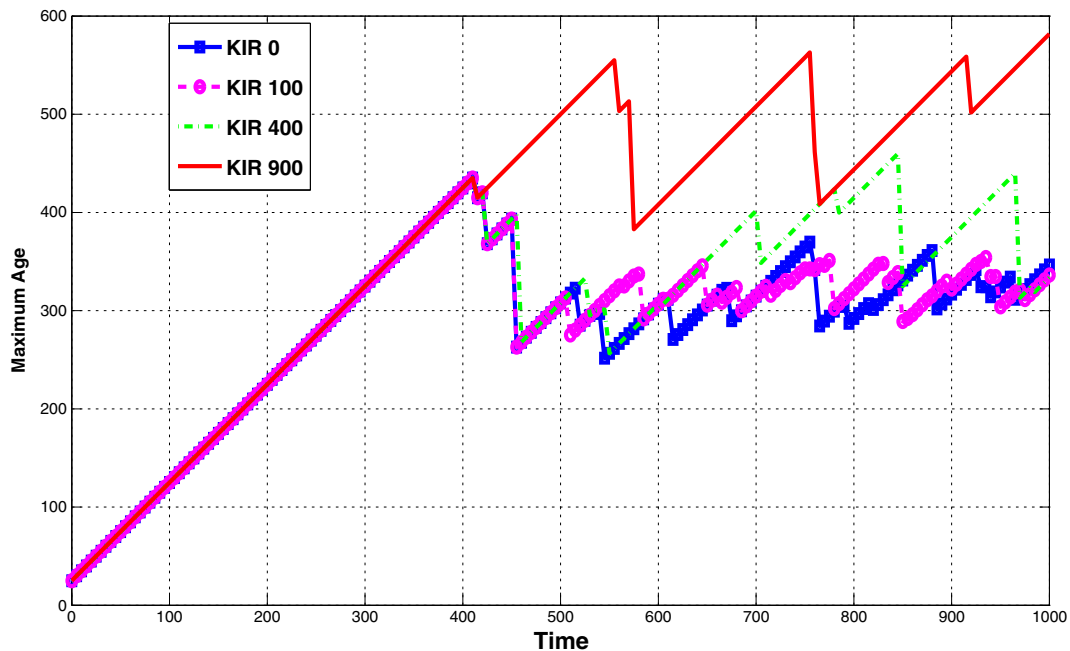
Let us take the case where the relative preference between regions is specified as the percentage of surveillance time to be spent within each region. As shown in Fig. 2, Region I, Region II, and Region III have respective areas of 16, 9, and 56 square units. Let the required percentage of surveillance time to be spent within those regions be 50 %, 20 %, and 30 %, respectively. This problem is solved using Eq. 12 discussed in Section 3.2. The simulation results are tabulated in Table 2.

It can be seen that the surveillance time spent within each region differs much from the requirement, especially, in Regions I and III. As the results shown in Table 1 closely meet the visitation requirement, it can be perceived that discretizing the regions into an infinitesimally spaced grid and using Eq. 11 might solve this problem. But this method is not practical. Hence, the simulation results for this case are

analyzed to determine the reason for the huge difference between the obtained results and the specified requirements.

The path followed by the UAV over the simulation time of 800 units is as shown in Fig. 8. The UAV travels between one preferred region to another mostly through the transition region, the boundary of which is marked in Fig. 8. Now, the time spent by the UAV in the regions is tabulated along with the time spent in transition region as shown in Table 3. The UAV surveillance time within Region II and Region III (without Transition region) meet the requirements closely. But, the UAV has spent 17.95 % of its flight time within Transition region alone, which is the primary reason for the flight time within Region I to be just 33.83 % instead of the desired 50 %.

A simplified test case was simulated to confirm that the transition between two preferred regions is the primary reason for the difference in the actual and the required percentages of surveillance time. In this case,

**Fig. 13** Results of Test Case 3: Comparison of maximum ages for various values of K_{IR}

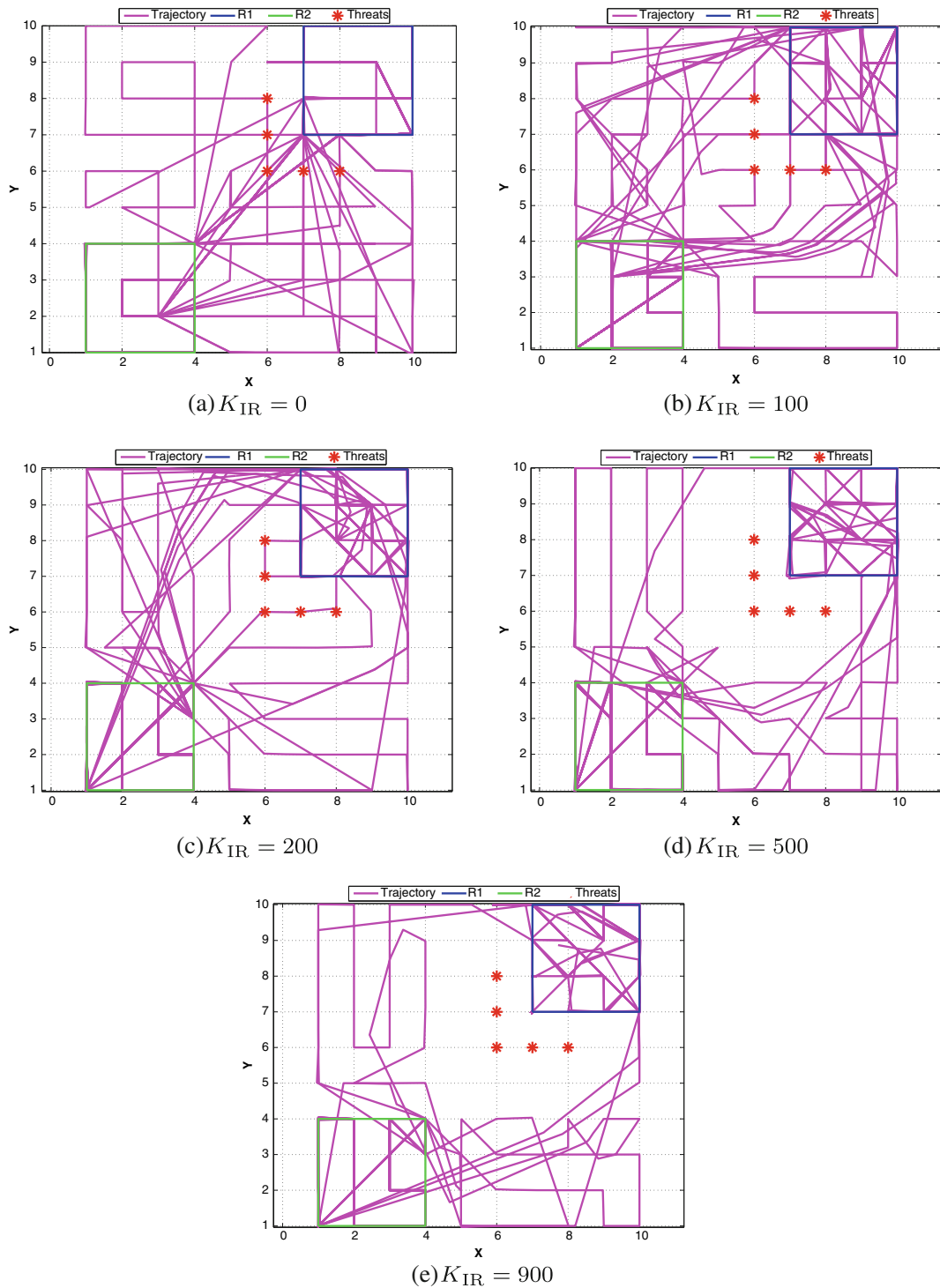


Fig. 14 Results of Test Case 3: UAV's planar trajectory for different values of K_{IR}

Region II is disregarded. Region I, with an area of 16 square units, retains the percentage of surveillance

time required as 50 %. Let the non-preferred region, R_N , represent the entire area without the preferred

Table 7 Percentage of visits to different regions for different K_{IR}

	Region I	Region II	Region III	Entry into Region I via Threat Region	Other entries into Region I
$K_{IR} = 0$	39.74	41.86	16.61	1.79	0.98
$K_{IR} = 100$	36.02	43.22	19.38	1.38	1.72
$K_{IR} = 200$	31.30	50.00	17.80	0.9	1.62
$K_{IR} = 500$	29.14	49.63	21.23	0	0.99
$K_{IR} = 900$	31.41	49.00	19.59	0	0.87
$K_{IR} = 200$ [Initial position (5,5)]	32.64	45.54	20.87	0.95	2.2

region, Region I. The simulation results are as shown in Table 4.

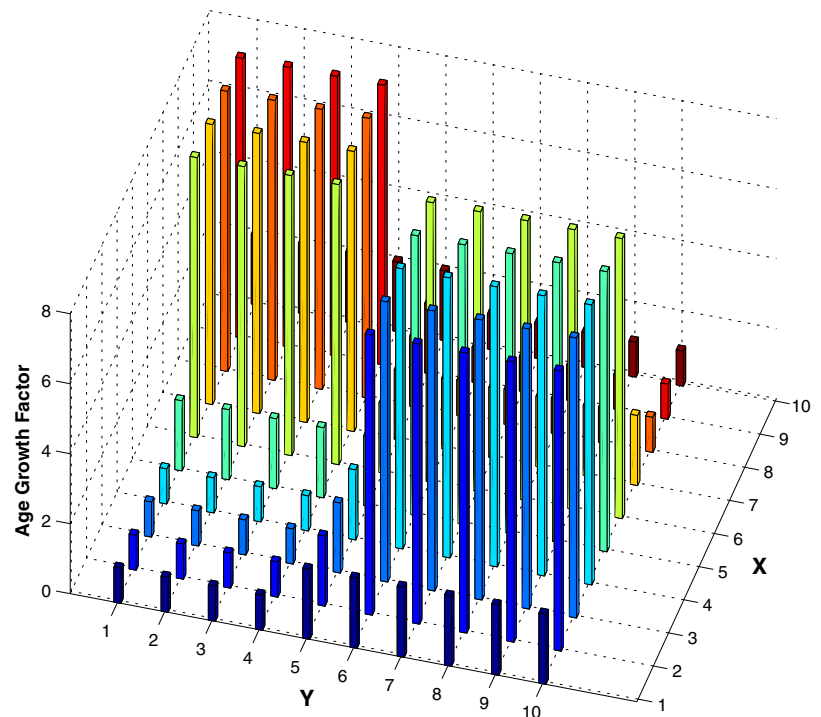
Table 4 shows that the time spent in each region closely satisfies the requirement. It is thus clear that the UAV spends much time and energy on switching destination points between preferred regions. The algorithm enables the UAV to meet the preference requirements specified as visitation percentages to different regions (Table 1) but the UAV does not meet those requirements specified as percentages of surveillance time within different regions. This difference is because the surveillance problem is formulated and solved in a discretized domain in this work, and the requirement to visit grid points, preferentially, conforms with this problem formulation. Additionally, as mentioned in assumption (vi) in Section 3, a grid point in the transition region is not said to be visited unless the UAV is within 0.05 units of radial distance of that point. Visitations are calculated as per this assumption and analyzed for a problem with visitation frequency specification. Whereas, the time spent within each region is calculated and analyzed for the problem with time based specification. Hence, the problem with visitation frequency specification circumvents the transition region phenomenon. In cases where the percentage requirement of surveillance time within different regions is to be satisfied, and the preferred regions are at far distances from each other, a better option would be to use multiple UAVs, so that much time is not lost in the transition from one preferred region to another. The same argument of using multiple UAVs holds good for a high detection probability of highly dynamic targets on the ground. This analysis concludes the simulation studies for the first scenario of Test Case 2. Only visitation

percentage requirements are considered for simulation in the subsequent cases.

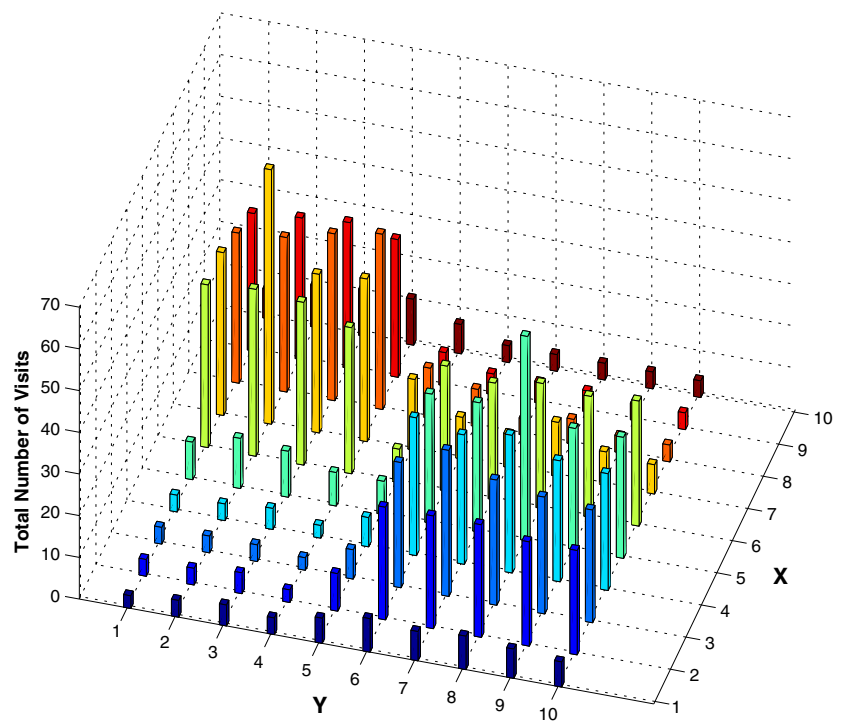
The regions in second scenario used to evaluate the dependency of the simulation results on trade-off factor, K_W , are as shown in Fig. 9. Region I, Region II, and Region III have 16, 16, and 68 grid points, and the required percentage of visits to be made to these regions are 40 %, 40 %, and 20 %, respectively. The visitation percentages to these regions for different K_W are shown in Table 5. $K_W = 0$ would make the UAV to ignore the distance to any grid point. Hence, the UAV would consume more fuel by shuttling between grid points across Region I and Region II, and hence Region III has recorded 44.3 % visitation against the required 20 %. Values of K_W between 0.3 and 60 caused the visitation percentage of Region III to vary only between 15 % and 25 %. Hence, the algorithm is not highly sensitive to the value of K_W . K_W has to be set by simulation and analysis for the given operational scenario.

It can be observed from Table 5 that high values of K_W such as 100 and 800 would result in uneven visitation percentages. The reason is that, once a UAV enters a certain region, it prefers to perform surveillance in the neighboring area till the age of a farther point is high enough. Figure 9 shows the UAV trajectories for different values of K_W . Localization of surveillance is observed to happen with increase in K_W . Figure 10 shows the maximum age of various regions for different values of K_W . It can be seen that the variation in maximum age of Region I and Region II with respect to time is dependent on K_W . The time between the drop in maximum age of Region I and Region II signify the UAV departing from one region and reaching another. This time is observed to increase

Fig. 15 Results of Test Case 4: Deterministic appearance of targets in preferred regions with $GF_{max} = 8$ and $DF = 25\%$

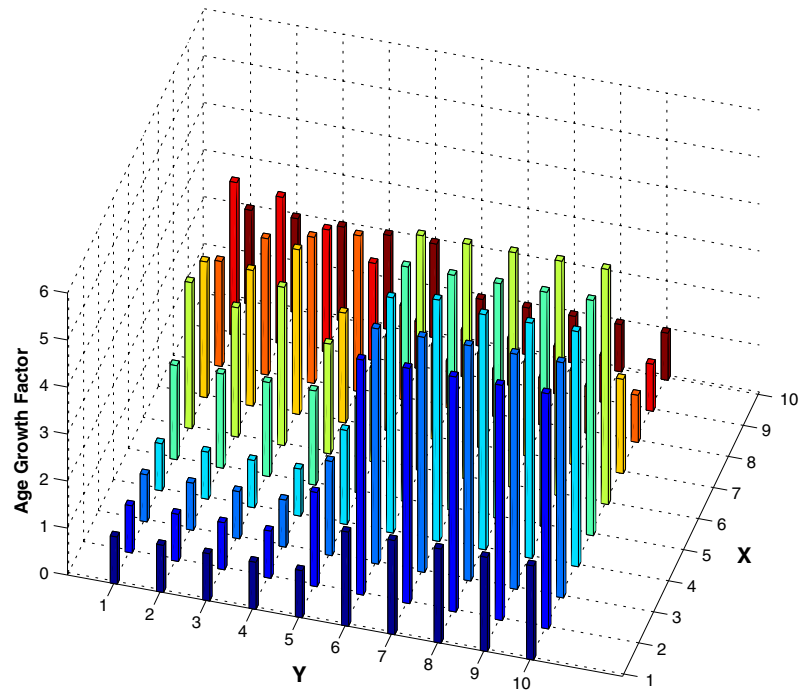


(a) Age growth factor

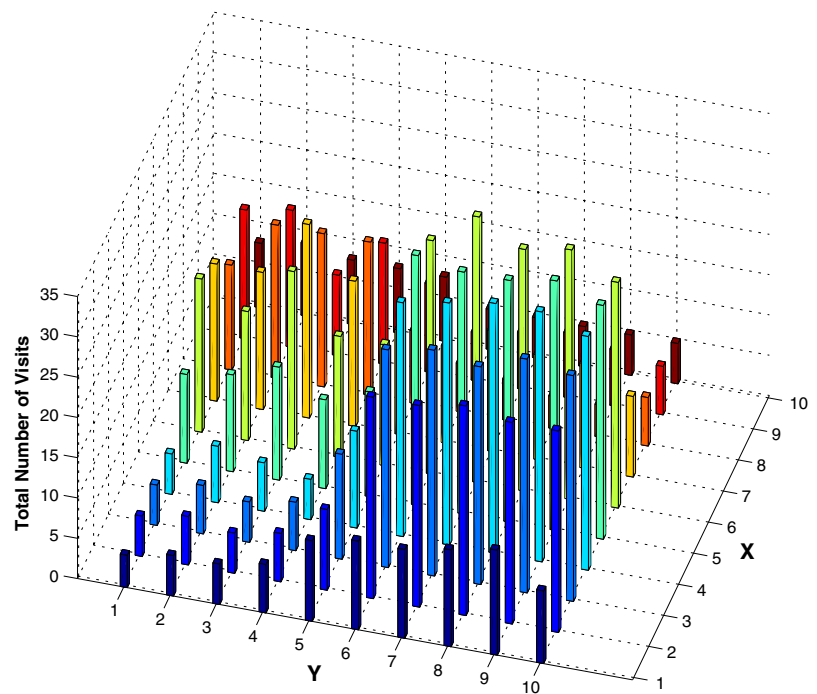


(b) Total number of visits

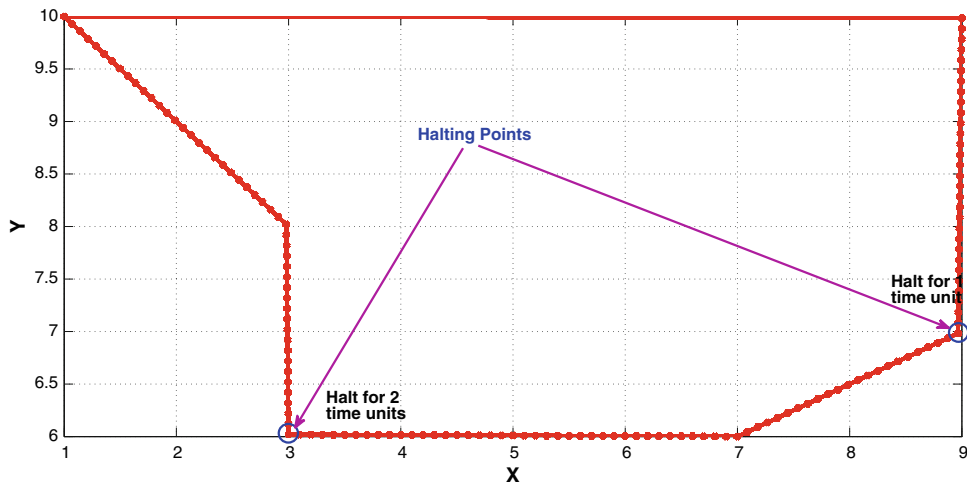
Fig. 16 Results of Test Case 4: Probabilistic appearance of targets in preferred regions with $GF_{max} = 5$ and $DF = 40\%$



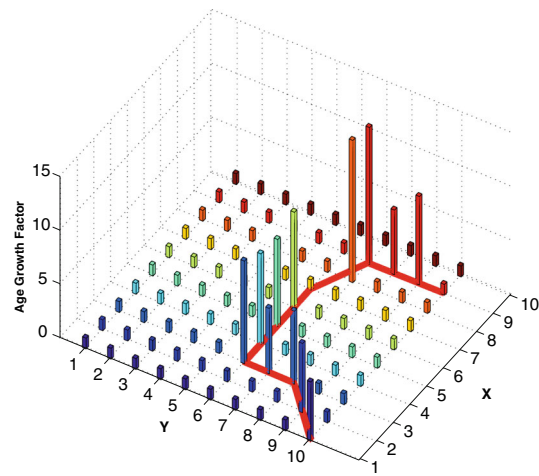
(a) Age growth factor



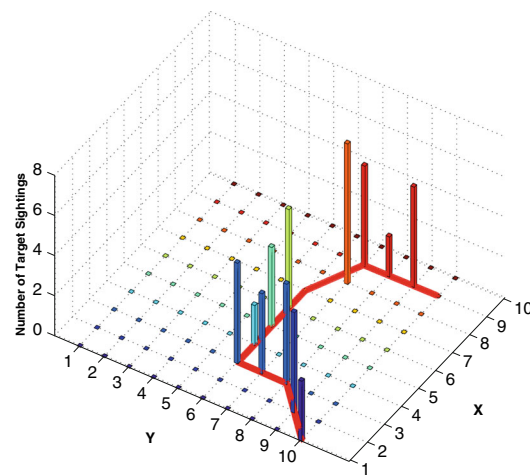
(b) Total number of visits



(a) Target Trajectory



(b) Age Growth Factor



(c) Number of Target Sightings

Fig. 17 Results of Test Case 4: Target following a repetitive Pattern with $GF_{max} = 75$ and $DF = 25\%$

with increase in K_W , implying that the UAV prefers to stay in the neighborhood of the current surveillance point as K_W increases.

4.3 Test Case 3: Preferential Surveillance with Known, Fixed Target Regions and Initially Unknown Threats

Two persistent surveillance scenarios with known preferred regions, the required percentage of visits to

be made to each region, and the threat positions are depicted in Figs. 11 and 14. In the first scenario, as shown in Fig. 11, threats are simulated to be present at (4,4) and (4,5) after 350 units of simulation time. When the UAV flies over these points after 350 units of time, it detects the presence of threats. Hence, R is a region containing 10×10 grid and $R_T = \{(4,4), (4,5)\}$, if $t > 350$. This problem is solved using the method explained in Section 3.3. K_W is set to 1 in the simulation for first scenario. Setting

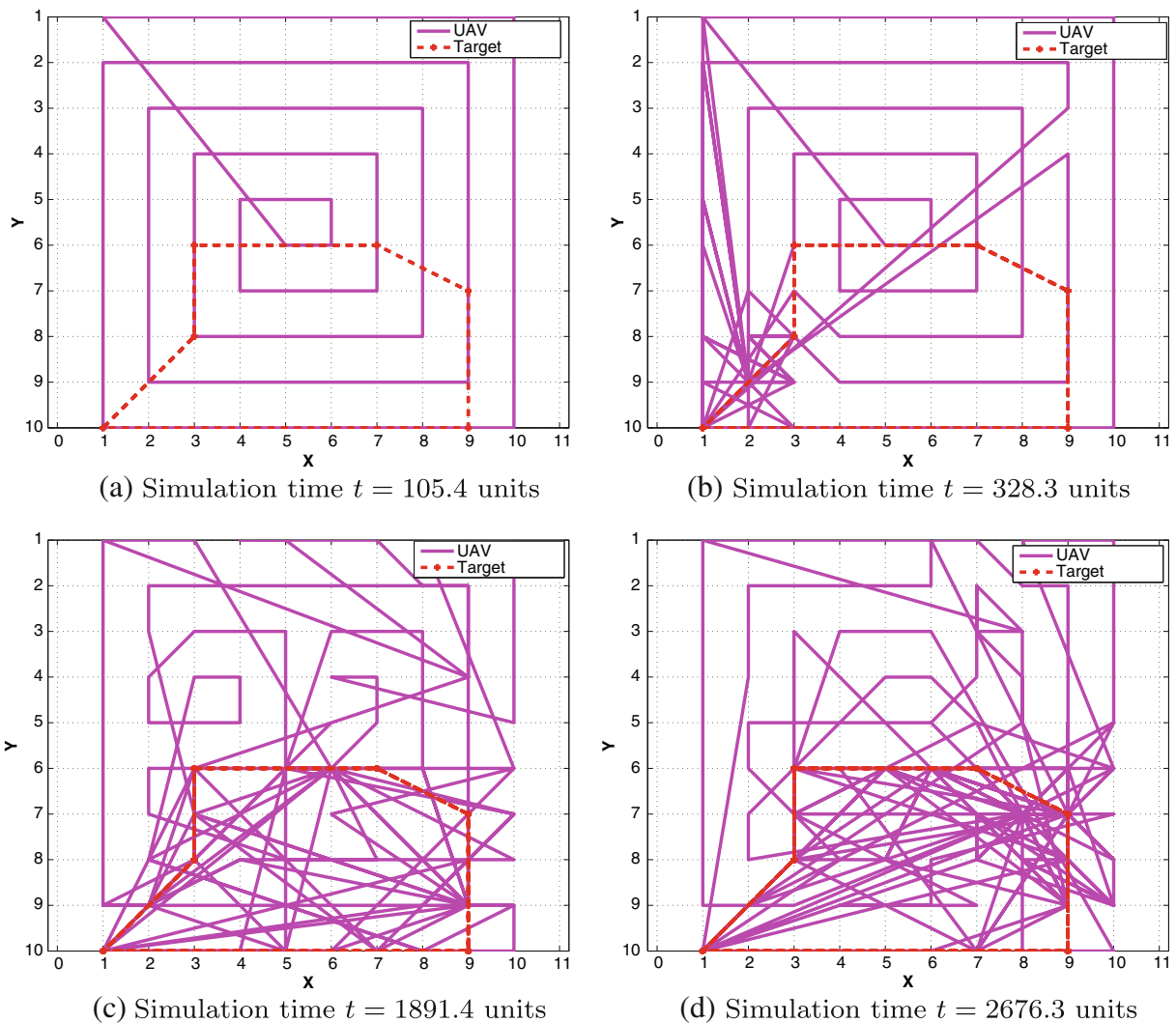


Fig. 18 Results of Test Case 4: Evolution of UAV's planar trajectory with time

the integral risk factor, $K_{IR} \neq 0$, affects the weight map so as to help the UAV to adapt to threats in the environment.

The simulation was run till 1000 units of time. K_{IR} is set to 200. The effect of variation in K_{IR} on the simulation results is discussed later. The UAV senses the threats at $R_T = \{(4, 4), (4, 5)\}$ as it flies over them at $t > 350$ units. The algorithm computes the risk map at every instant of time and reduces the number of visits to the threat positions by devaluing their costs. The visits made to the entire surveillance before and after threat introduction are shown in Fig. 12a and b. The number of visits to the R_T reduced to 2 out of 378 total number of visits from 6 out of 289 visits made by the UAV before the detection of threats. This problem involves performing preferential surveillance while avoiding the risky regions. The visitation percentages shown in Table 6 meet the specified requirement (Fig. 11).

Similar simulation results closely satisfying the visitation percentage requirements are obtained for other values of K_{IR} . K_W determines the cost matrix by affecting the trade-off between age and weight matrices whereas K_{IR} determines the weight matrix by affecting the trade-off between distance and integral risk. Figure 13 shows the effect of K_{IR} on the maximum age. Increasing the K_{IR} to very high values such as 900, as seen in Fig. 13, increases the risk posed by such threats and devalues the cost of certain points around the threat too. Then, the UAV avoids visiting these points because of their lower cost values, and the maximum age increases. It can be observed from Eqs. 6, 15 and 10 that maximum age is much more sensitive to changes in K_W . This concludes the analysis of simulation results for the first scenario with threats.

The regions in second scenario used to evaluate the dependency of the simulation results on integral risk factor, K_{IR} , are as shown in Fig. 14. Region I, Region II, and Region III have 16, 16, and 68 grid points, and

the required percentage of visits to be made to these regions are 40 %, 40 %, and 20 %, respectively. In addition to this, threats are positioned very close to preferred region, Region I, in order to study its effect on the algorithm performance. As shown in Fig. 14, threats are located at (6,8), (6,7), (6,6), (7,6), and (8,6), covering more than half of the entry points into Region I. K_W is set to 4 in this simulation and the simulation is run for 800 time units. It can be observed how the UAV avoids the threats based on the value of K_{IR} . While avoiding the threats, the algorithm helps the UAV to adhere to the visitation percentage requirements as much as possible. The visitation percentages to the various regions for different K_{IR} are shown in Table 7.

It can be observed through the decreasing entries into Region I via threat region that the increase in K_{IR} makes the UAV to approach Region I via the grid points in Region III without threats. Though high values of K_{IR} such as 500 and 900 make the UAV completely avoid threat regions, it also reduces the visitation percentage in Region I to close to 30 % against the required 40 %. This reduction leads to an increase in visitation percentage in Region II to close to 50 % as shown in Table 7. It can also be observed that change in the initial position of UAV does not significantly affect the simulation results.

4.4 Test Case 4: Preferential Surveillance with Unknown, Dynamic Target Regions and Initially Unknown Threats

The method discussed in Section 3.4 has been used to solve this problem. The simulation was run till 3000 units of time. Three target scenarios are simulated to evaluate the algorithm. K_W is set to 1 and K_{IR} is set to 100, unless stated otherwise. The first scenario simulates the target to appear certainly within the UAV sensor coverage when the UAV flies over any grid point belonging to a priority region (initially unknown to the UAV). The priority region is representative of a fixed target region with 100 % target detection probability. The UAV has to discover this region using learning. GF_{max} is set to 8 and the discounted factor of 8-connected neighborhood, DF , is set to 25 % (that is, GF value of 2). Region I and Region II, as shown in Fig. 11, are considered as priority regions in the simulation. When the UAV flies over a grid point

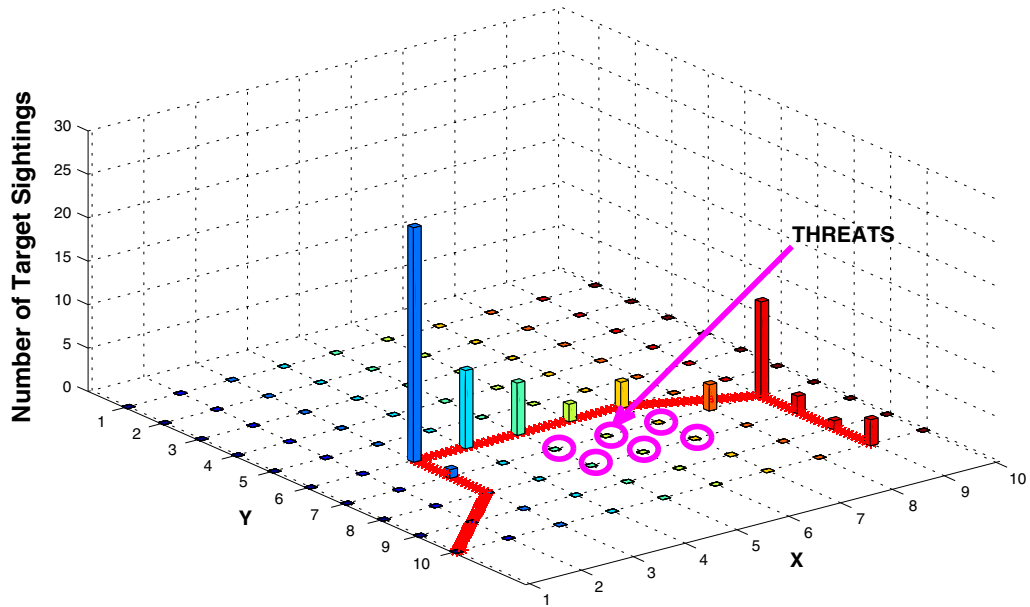
Table 8 Number of target sightings and threat visits for different K_{IR}

	Visit percentage to Threat Region	Number of target sightings
$K_{IR} = 0$	6.17	68
$K_{IR} = 900$	4.63	368

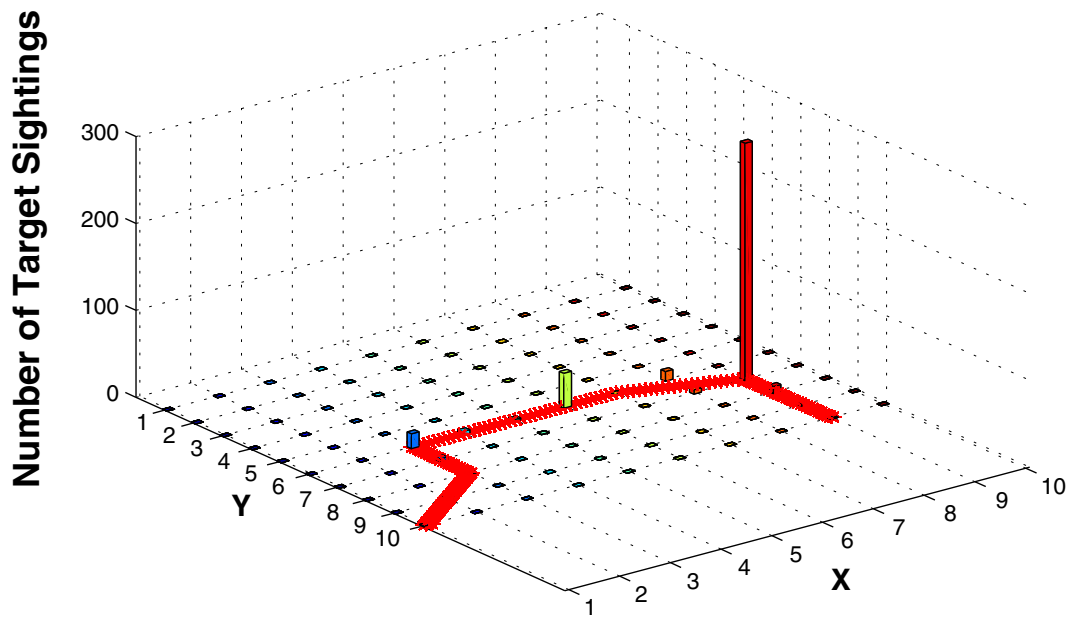
belonging to these regions, it senses a target, increases the age growth factor of that point and neighborhood.

Simulation results for the first scenario are shown in Fig. 15. The average number of visits made to the

region per grid point is relatively higher for Region II (38.75) compared to Region I (30.04). This is because the latter has a higher number of grid points than the former. If no target is detected in the subsequent



(a) $K_{IR} = 0$



(b) $K_{IR} = 900$

Fig. 19 Results of Test Case 4: Number of target sightings for different values of K_{IR}

time steps in the neighborhood of the point where a target was detected, the age growth factor of that neighborhood point is reset to GF_{R_N} .

The second scenario simulates targets that the UAV can detect with a specified probability when the latter flies over priority regions (initially unknown to the UAV). Here, $GF_{max} = 5$ and $DF = 40\%$. Target detection probability in Region I is set to 100% and that in Region II to 50%. The average number of visits per point in Region II (17.25) is lower than that of Region I (28.68). This is because the age growth factor of Region II is lower than that of Region I. Simulation results are as shown in Fig. 16.

The third scenario is more practical than the first two. A target exhibits a repetitive pattern on the ground, and it halts for a particular time at two intermediate points. The target enters the surveillance area at (1, 10), follows the trajectory shown in Fig. 17a, exits the area at (9, 10), and repeats this motion. It halts at (9, 7) and (3, 6) for 1 and 2 time units, respectively. The halt points signify containment loading or unloading points, or a major traffic junction. The speed of the target is assumed to be 1 unit distance per unit time. The number of episodes the target completed is 166 for a simulation time of 3000 units. Here, $GF_{max} = 75$ and $DF = 25\%$. The age growth factor and the number of target sightings are shown in Fig. 17b and c, respectively. The evolution of UAV trajectory is shown in Fig. 18.

Figure 18 shows the effect of age growth factor on the UAV path planning. The vertical axis is reversed in Fig. 18 for easier correlation with Fig. 17. Figure 18a shows the path followed by the UAV till it completes one cycle of visitation to all the grid points. This UAV path is the same as that shown in Fig. 5a, the solution for a simple persistent surveillance problem. As the simulation time increases, Fig. 18 implies that the algorithm enables the UAV to discover the region with a high likelihood of target detection through learning and to visit that region frequently.

In the third scenario discussed above, threats are added at (5,7), (5,8), (6,7), (6,8), (7,7), and (7,8). K_W is set to 4 for similarity with the second threat scenario discussed in Section 4.3. The effects of these threats on the simulation results are studied and tabulated for different values of K_{IR} in Table 8. It can be seen that the visitation percentage to a threat region, in this case, does not vary much with K_{IR} . This is because the heuristic learning algorithm, explained in

Section 3.4, inherently motivates the UAV to navigate to grid points where the target was frequently detected in the recent past. Hence, the visitation percentage to the threat regions is less. However, avoiding the threat region in the surveillance area helps the UAV to focus more towards the points of earlier target detection. This causes an increase in the number of target sightings. But the drawback is that the UAV focuses more on certain grid points, that is, the UAV visits target detected grid points much more frequently than other points. This results in lesser exploratory behavior of the UAV. The number of target sightings for different values of K_{IR} are shown in Fig. 19.

The vehicular kinematic and dynamic constraints have not been included in the test four cases discussed above. When these constraints are enforced, it costs more time and energy for the UAV to maneuver to the destination point. It is essential to understand how the time taken to perform such successive maneuvers to reach destination points affects the learning and path planning capabilities of the developed algorithm. In addition to this, the effectiveness of this algorithm has to be studied under energy (fuel or battery) constraints and health (subsystem failure) constraints.

5 Conclusions and Future Work

A heuristic learning algorithm for a preferential surveillance problem in an environment with dynamic targets and uncertain threats is proposed and evaluated through simulations. Furthermore, this work presents a technique for quantitative preference allocation to different regions of a known surveillance area. The benefits of the algorithm in increasing the target detection probability in an unknown surveillance area have also been demonstrated. UAV dynamics has to be implemented for a realistic study. Analysis of the results obtained for preferential surveillance problem indicates the need to use multiple UAVs to solve the problem more effectively. It is necessary to estimate the computational complexity of this method and to study the methods of reducing the computational load, which would help in applying the algorithm for large-scale problems. The learning algorithm can be further improved to locate a target in minimum time.

Acknowledgments The authors thank the Defence Research and Development Organization (DRDO), India and Indian

Institute of Science (IISc), Bangalore, India for the Research & Training scheme through which Manickam Ramasamy could pursue this work as a part of his Masters degree at the Department of Aerospace Engineering, Indian Institute of Science, India.

References

1. Nigam, N., Bieniawski, S., Kroo, I., Vian, J.: Control of multiple UAVs for persistent surveillance: Algorithm and flight test results. *IEEE Trans. Control Syst. Technol.* **20**(5), 1236–1251 (2012)
2. Mersheeva, V., Friedrich, G.: Multi-UAV monitoring with priorities and limited energy resources, Jerusalem, Israel (2015)
3. Bethke, B., How, J.P., Vian, J.: Multi-UAV persistent surveillance with communication constraints and health management. In: *AIAA Guidance, Navigation, and Control Conference*, Chicago, Illinois (2009)
4. Stump, E., Michael, N.: Multi-robot persistent surveillance planning as a vehicle routing problem. In: *Proceedings of IEEE Conference on Automation Science and Engineering (CASE)*, pp. 569–575, Trieste (2011)
5. Alamdari, S., Fata, E., Smith, S.L.: Persistent monitoring in discrete environments: Minimizing the maximum weighted latency between observations. *Int. J. Robot. Res.* **33**(1), 138–154 (2014)
6. Cassandras, C.G., Lin, X.: Optimal control of multi-agent persistent monitoring systems with performance constraints. In: *Proceedings of the 50th IEEE Conference on Decision and Control*, pp. 2907–2912 (2011)
7. Arvelo, E., Kim, E., Martins, N.C.: Memoryless Control Design for Persistent Surveillance under Safety Constraints. *Cornell University Library Article*. arxiv: [1209.5805](https://arxiv.org/abs/1209.5805). Accessed 20 May 2016
8. Leahy, K., Zhou, D., Vasile, C.-I., Oikonomopoulos, K., Schwager, M., Belta, C.: Provably correct persistent surveillance for unmanned aerial vehicles subject to charging constraints. In: *Proceedings of the International Symposium on Experimental Robotics (ISER 14)*, Marrakech, Morocco (2014)
9. Burman, J., Hespanha, J., Madhow, U., Isaacs, J., Venkateswaran, S., Pham, T.: Autonomous UAV persistent surveillance using bio-inspired strategies. In: *Proceedings of the SPIE 8389, Ground/Air Multisensor Interoperability, Integration, and Networking for Persistent ISR III 838917* (2012)
10. Nigam, N.: The multiple unmanned air vehicle persistent surveillance problem: a review. *Machines* **20**(1), 13–72 (2014)
11. Russell, S.J., Norvig, P.: Learning from Examples. In: Russell, S.J., Norvig, P. (eds.) *Artificial Intelligence: A Modern Approach*. 3rd edn, pp. 695–696. Pearson Education Inc (2010)
12. Sutton, R.S., Barto, A.G.: Introduction. In: Sutton, R.S., Barto, A.G. (eds.) *Reinforcement Learning: An Introduction*, pp. 3–9. The MIT Press, Cambridge (1998)
13. Zhang, B., Mao, Z., Liu, W., Liu, J.: Geometric reinforcement learning for path planning of UAV. *J. Intell. Robot. Syst.* **77**, 391–409 (2015)
14. Geramifard, A., Redding, J., How, J.P.: Intelligent cooperative control architecture: a framework for performance improvement using safe learning. *J. Intell. Robot. Syst.* **72**, 83–103 (2013)
15. Cook, K., Bryan, E., Yu, H., Bai, H., Seppi, K., Beard, R.: Intelligent cooperative control for urban tracking. *J. Intell. Robot. Syst.* **74**, 251–267 (2014)
16. Hirsch, M.J., Ortiz-Pena, H.J., Eck, C.: Cooperative tracking of multiple targets by a team of autonomous UAVs. doi:[10.4018/joris.2012010104](https://doi.org/10.4018/joris.2012010104) (2012)
17. Ure, N.K., Chowdhary, G., Chen, Y.F., Cutler, M., How, J.P., Vian, J.: Decentralized learning-based planning for multi-agent missions in the presence of actuator failures. In: *International Conference on Unmanned Aircraft Systems*, GA Atlanta (2013)
18. Polycarpou, M.M., Yang, Y., Passino, K.M.: Multi-UAV cooperative search using an opportunistic learning method. *J. Dyn. Syst. Meas. Contr.* **129**(5), 716–728 (2007)
19. Yu, H., Beard, R., Argyle, M., Chamberlain, C.: Probabilistic path planning for Cooperative Target Tracking using aerial and ground vehicles. In: *American Control Conference*, San Francisco CA (2011)
20. Sujit, P.B., Ghose, D.: Self Assessment-Based decision making for multiagent cooperative search. *IEEE Trans. Autom. Sci. Eng.* **8**(4), 705–719 (2011)
21. Zhu, J.J., Xu, X.: Biopsychically inspired cognitive control for autonomous agents based on motivated learning. In: *17th International Conference on Methods and Models in Automation and Robotics, Miedzyzdroje* (2012)
22. Polycarpou, M.M., Yang, Y., Passino, K.M.: Cooperative Control of Distributed Multi-agent Systems. *IEEE Control Syst. Mag.* (2001)
23. Redding, J., Bethke, B., Bertuccelli, L.F., How, J.P.: Active learning in persistent surveillance uav missions. In: *AIAA Infotech@Aerospace Conference*, Seattle, Washington (2009)
24. Ramasamy, M., Ghose, D.: Learning-based preferential surveillance algorithm for persistent surveillance by unmanned aerial vehicles. In: *2016 International Conference on Unmanned Aircraft Systems (ICUAS)*, Arlington, VA (2016)

Manickam Ramasamy obtained his Bachelor's degree in Aeronautical Engineering from the Madras Institute of Technology, Chennai, India in 2009. He is a scientist at the Aeronautical Development Establishment, Defence Research and Development Organization, Bangalore, India. He has completed his Master's in Aerospace Engineering at the Department of Aerospace Engineering, Indian Institute of Science, Bangalore in 2016. His research interests include modeling, simulation, and analysis of flight dynamics, path planning, multi-agent learning, and joint learning-planning techniques with applications to autonomous aerial vehicles.

Debasish Ghose obtained his Bachelor's degree in Electrical Engineering from the Regional Engineering College, Rourkela, India in 1982. He received his Master's and Ph.D. degrees in engineering from the Indian Institute of Science, Bangalore, India in 1984 and 1990, respectively. He is a Professor in the Department of Aerospace Engineering, Indian Institute of Science, Bangalore, India. His main area of interest is the guidance and control of autonomous vehicles, with a special focus on aerospace and underwater vehicles. Guidance theory, vision based guidance, sliding mode control based guidance, collision avoidance algorithms, path planning, trajectory optimization, and multi-agent systems are some of his major current research areas.