

Research Article

Cooperative Loitering Munition Swarm Online Patrolling Route Planning with Nonlinear Seeker Measurement

Haifeng Ling,¹ Hongchuan Luo ,¹ Linyuan Bai,¹ Tao Zhu ,¹ Qing Wang,¹ and Lidong Yu²

¹Field Engineering College Army, Engineering University of PLA, Nanjing, China

²Unite of 94782 of PLA, Hangzhou, China

Correspondence should be addressed to Hongchuan Luo; luohc2004@outlook.com

Received 17 August 2020; Revised 31 August 2020; Accepted 8 September 2020; Published 14 October 2020

Academic Editor: Xingling Shao

Copyright © 2020 Haifeng Ling et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

With the development of autonomous systems, the operational use of loitering munition is shifting from the following of a preplanned fixed route without communication to smart decision-making and collaborative cooperation with sharing information. In this paper, we study the autonomous decision-making and cooperative control strategy of online patrolling for a swarm of loitering munitions using communication to coordinate their route based on maximizing the information they gathered in the operation region. Taking the non-Gaussian nonlinear property of airborne radar seeker into account, we utilized a particle filter-based method to evaluate or to predict the information quality of each action candidate. We also implemented a coordinate descent scheme to enable a distributed and scalable swarm decision-making. Simulation results show that the proposed method provides a better estimation than baselines without the need for external or centralized decision agent.

1. Introduction

Loitering munitions are mainly used to destroy and suppress the enemy's air defense system and to create opportunities for attackers. The loitering munition is usually equipped with a radar seeker a.k.a. an airborne radar signal receiver, guiding devices and a warhead. After launching into the operation air zone, it circulates and searches for enemy's radar radiation signals and conducts strikes to the radar target positioned by its radar seeker. As early as in the air battle in the Syria-Israel Bekaa Valley, the Israeli loiter munition Harpy has achieved remarkable success in suppressing and destroying Syrian air defense forces and has become famous all over the world. Taken the Harpy as an example, a typical composition of a loiter munition is shown in Figure 1. Modern military powers are ubiquitously equipped with various types of loitering munitions.

In general, equipped with an external positioning system and inertial guidance devices, traditional loitering munitions fly along a preplanned route (usually, this route is circular or figure-eight) over the assigned combat air zone and discard any control signal until a radar target is found, the strike

operation is conducted, or they run out of fuels. This phase is known as patrolling stage. The main task of the patrolling stage is to maximize the information gathered about the target in the combat air zone and to prepare for the strike phase. Therefore, the effectiveness of patrolling route planning plays a significant role in the combat performance of loitering munitions.

1.1. Related Work. Existing research on loitering munitions patrolling route planning mostly uses coverage as the indicator and conducts offline route planning based on the geometry of the combat air zone. The authors of [1–3] use “field coverage” as the route planning rewards indicator, a consensus control law and a modified artificial potential field method to improve the performance of the planned patrolling route. However, since these methods are open loop which only carry out the fixed route planned in advance, they neglect the real-time measurement of the radar seeker, neither replan online nor exert the advantage of autonomous decision-making ability of the unmanned system. And also, when there are multiple loitering munitions operating in the

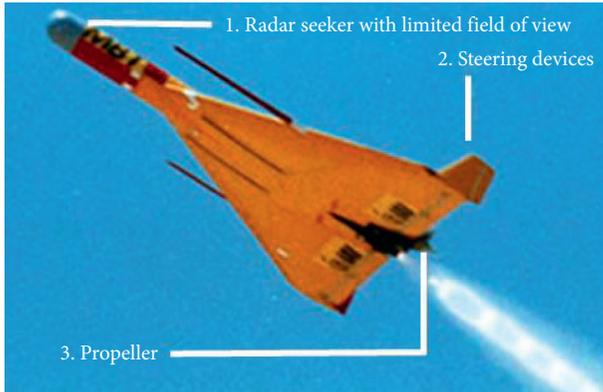


FIGURE 1: Typical composition of a loiter munition. Original picture excerpt from [4].

same air zone at the same time, these methods can solely rely on the radar seeker carried by each loitering munition to search for the target rather than fully using the measurements from other UAVs to update the information. Thus, the performance is limited.

With the development of unmanned swarm system technology in recent years [5], through the interconnection and transmission and information exchange between individuals in the swarm, improving search efficiency using cooperative control has aroused the research interest of experts and scholars at home and abroad. The literature of this type is often referred to as the cooperative search problem. The research on the target search problem can be traced back to the research of antisubmarine warfare by Koopman and others [6] during World War II. Derived from the probability theory, the authors of [7, 8] considered a single agent search problem including cost of switching search location and failure rate of search due to imperfect sensors and solved the problem using a classic framework of continuous postprobability updating with the searched result. Zhu [8] also pointed out the difference of searching planning between maximizing the probability of discovery (a.k.a. detection search) and maximizing the belief of target's location (a.k.a. whereabouts search). Based on single agent search theory, Zhu et al. [9] proposed multiple agent searching planning methods using clustering and minimum spanning tree to coordinate team to minimize the cost of changing search position. It should be noted that because the sensor is modeled as a "found-or-not-found" binary model, search effort is needed only under the condition of "not found," so the strategy obtained is open loop, which is destined to end when the target discovered. Nevertheless, it is still implying that the search behavior is changed in real time according to the detection result. Whereabouts, search includes not only the optimal strategy when the search succeeds in finding the target but also the optimal strategy for guessing the target's location when the search is unsuccessful. In the field of robotics and multiagent, the reward, rate of probability return, of a search plan is often replaced by information metric using information theory. These problems are cataloged as active information gathering [10] problem, whose goal is maximizing the quality of

information collected along the route planned in real time. The research of sensors network (sensors network), which has emerged in recent years [11, 12], also covers the scope of multiple agent information gathering. In the scenario where the target distribution probability is Gaussian and the measurement noise is also Gaussian, the information quality metric can be the covariance of the target distribution probability and the measurement of sensors. In the more general scenario, the metric is often information entropy, mutual information [13], and Kullback–Leibler divergence [10]. Mutual information between measurement and target state indicates that how much one can learn about target's state from the measurement. The larger the mutual information, the more the one learn about target state based on measurements.

Although information gathering problem differs from search theory in a more generalized sensor model, both in essence lie within Bayesian inference framework, information gathering problem models the sensor as imperfect measurement of target's state with noise interference, and the target state distribution is updated every time the measurement arrives. For linear Gaussian systems, classical Kalman filtering weights the mean and covariance of inferred distribution by using the covariance of measurement and prior distribution. But real physical systems are often nonlinear and non-Gaussian [14]. For nonlinear systems, linearization methods such as extended Kalman filtering are useful. However, in highly nonlinear systems, the accuracy of extended Kalman filtering is limited, and it cannot process non-Gaussian noise. With the improvement of airborne computing resources recently, the particle filtering method based on sequential Monte Carlo (SMC) sampling is widely used because of its capacity of effectively processing nonlinear [13] and non-Gaussian noise.

Another critical issue is that distribution swarm decision is a must because it is particularly relevant to scalability and applicability. Comprising multiple loitering munitions, the swarm is expected to behave collaborative and cooperative without any centric agent or fully connected network. Some literatures often assume that there exists a centric agent or fully connected communication without delay in which information from all agents in swarm can be aggregated and dispatched. Thus, there is no information difference among all agents, and the coordination problem degenerates into a single-state decision problem with a higher state dimension. In real-world scenario, this assumption is often difficult to meet. On the other extreme, all agents cannot communicate, and each agent makes independent decisions based solely on local information, which leads to limited performance. Our goal of this work is to devise a distributed approach with scalability in the number of agents and meanwhile to still share information and decision among each other and plan cooperative route.

Zavlanos et al. [15] utilize the "distributed auction" algorithm to optimize task allocation without the need for a centric agent or fully connected communication. However, this method requires multiple bids and rebid communication during the swarm cooperative decision process. In the loitering munitions patrolling scenario where measurement

signals arrive continuously at high frequencies, the auction-based swarm coordinate method meets difficulties in practical due to the limitation of concurrent communication capacity. The coordinate descent method used in [16, 17] provides a suboptimal solution to reduce the decision space of multiagent decision-making problems from exponential growth on the number of agents to linear growth. It is proved that this method can get more than 50% return of the optimal solution and requires a smaller amount of communication [17]. Due to hierarchy structure of decision where the agent with higher rank makes the decision with higher priority, Chen [18] called this method sequential allocation and used this method to solve the problem of multi-UAVs reconnaissance and surveillance planning.

In this paper, the loitering munitions swarm online autonomous patrolling route planning problem is modeled as an information collection problem. Loitering munitions in the swarm share the measurements collected by its airborne radar seeker and their own status via wireless communication and update the radar target's position [19] distribution in real time based on the prior position distribution and the observation sequence. Mutual information is selected as the optimization goal, and a nonmyopic iterative optimization strategy in the model predictive control (MPC) scheme is adopted to conduct route planning online. This paper assumes that all loitering munitions in the swarm can communicate in ad hoc mode, but communication capacity is limited, and a fully connected communication is not afforded. It is assumed that the airborne radar seeker (signal receiver) has limited field of view, limited sensitivity, and non-Gaussian errors. Particle filtering [19] is used to infer target position information to calculate mutual information between measurements and target state. The problems of collision avoidance, obstacle avoidance, and patrolling-striking switch in real operations can also be solved within the model predictive control framework proposed in this paper through hierarchical decision-making, inner and outer loop control, and multitarget weighting. It is noted that a shorter conference version of this paper is submitted to 5th International Conference on Automation, Control and Robotics Engineering (CACRE 2020). The initial conference paper addressed the patrolling route planning problem without considering the problem caused by unknown parameters like radar transmitting power and the error introduced by quantization [20, 21] of measurements acquired by airborne radar seeker. This manuscript addresses those issues and provides a more complicated simulation experiment. The content of this article is organized as follows: Section 2 introduces the modeling of information collection problems based on the model predictive control framework. Section 3 introduces the collaborative descent method used in swarm distributed decision-making under communication constraints. Section 4 deals with the calculation of mutual information using particle filtering-based methods. Section 5 illustrates the effectiveness of the method proposed by simulation experiments using Octave [22]. Section 6 summarizes and prospects future work.

2. Problem Formulation

Active information gathering problem refers to the optimal planning of the executor's path or action under the constraints of certain resources which aims to maximize information revenue along the route planned. Integrating information perception, data fusion, and inference, this problem is an emerging hot topic in robotics, computer science, and military operations research. To plan the route of the loitering munitions swarm using information gathering framework, it is necessary to comprehensively consider the swarm dynamics model, radar target model, and the airborne radar seeker sensing model. Table 1 lists all the notations used in the following and their meaning.

2.1. Dynamics Model of the Loitering Munitions Swarm. Consider a team of n_a loitering munitions (referred to as agents in the remaining paper) obeying the following motion dynamics:

$$\mathbf{s}_{k+1}^i = f^i(\mathbf{s}_k^i, \mathbf{u}_k^i), \quad i \in A = \{1, \dots, n_a\}, \quad (1)$$

where $\mathbf{s}_k^i \in S^i \cong \mathbb{R}^{n_s}$ is the n_s dimensional state of loitering munition i at time k and $\mathbf{u}_k^i \in U^i$ is the control action applied to the loitering munition i at time k and is selected among all possible decisions U^i . In this paper, it is assumed that during the patrolling phase, the agents are flying in the combat air zone in a fixed speed and height and the action candidates are finite discrete yaw angle, that is,

$$\begin{aligned} \mathbf{s}_{k+1}^i &= \begin{bmatrix} x_{k+1}^i & y_{k+1}^i & \varphi_{k+1}^i & \gamma_{k+1}^i \end{bmatrix}^T, \\ &= \begin{bmatrix} x_k^i + v \cos^i(\varphi_{k+1}^i) \Delta t \\ y_k^i + v \sin^i(\varphi_{k+1}^i) \Delta t \\ \varphi_k^i \\ \gamma_k^i - c \cdot \Delta t \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ \Delta t \\ d \cdot \Delta t \end{bmatrix} \mathbf{u}_k^i, \end{aligned} \quad (2)$$

where x and y are the coordinates of agent in an orthogonal plane whose origin lies in the combat air zone, v is a constant representing the unchanged velocity, φ is the yaw angle, γ is a real number representing the fuel remainder, and $c > 0$ and $d < 0$ are the fuel consumption ratios, respectively, representing the fuel consumption for flying and steering. The decision set is defined as $U^i = \{0, -\omega^i, \omega^i\}$ is the control action, i.e., flight straight, maximum left turn, and maximum right turn. It is noted that the dynamic model formulated in (2) is a special form of (1) and is formulated for engineering purpose. The method is applicable to other dynamic models formulated in (1).

2.2. Dynamics Model of the Radar Target. The mission of the swarm is to suppress the stationary radar target T at an unknown location within the combat air zone. The state of the target $r = [x_T \ y_T \ \alpha]$ is unknown, where $|x_T| \leq x_{\text{field}}$ and $|y_T| \leq y_{\text{field}}$ are the coordinates of the target in the same orthogonal plane with loitering munitions, and

TABLE 1: Nomenclature.

Nomenclature	
\mathbf{s}, S	State and state space of loitering munition
k	Time index
\mathbf{u}, \mathcal{U}	Control action and action space of loitering munition
v	Constant value of loitering speed
x_T, y_t, α	Component of radar target state
rss, bear	Radio strength and bear of radar emission
φ_{fov}	Constant value of sensor field of view
i, n_a	Index and number of loitering munitions
x, y, φ, γ	Components of loitering munition state
c, d	Cost ratio of straight flight and turning
\mathbf{r}, R	State and state space of radar target
z, Z, \mathcal{Z}	Measurement, measurement space, and measurement union set of airborne radar seeker
τ, ω	Noise of range and bear
\mathcal{T}	Finite time step of model predictive control

$\alpha > 0$ is a unknown constant relating to the transmit power of the radar.

2.3. Model of the Radar Seeker. The model of the radar seeker is described by the measurement. Agent i measures the radar target:

$${}^i z_k = h(\mathbf{s}_k^i, \mathbf{r}, v_k(\mathbf{s}_k^i)), \quad (3)$$

where ${}^i z_k \in \mathbb{R}^{n_s}$ is the measurement of agent i at time k , which is determined by the state of agent \mathbf{s}_k^i and the unknown state of radar target \mathbf{r} . $v_k(\mathbf{s}_k^i)$ is an agent-state-dependent noise, whose values are independent at any pair of times and across the agents. Unlike the study [10] based on the extended Kalman filtering, $v_k(\mathbf{s}_k^i)$ does not have to be Gaussian noise nor $h(\mathbf{s}_k^i, \mathbf{r})$ does not have to be linear mapping. In our work, we consider a generalized range bearing sensor with limited sensitivity and field of view, influenced by a mixture of Gaussian and non-Gaussian noise.

Signal strength measurement is modeled as

$$\text{rss} = \mu \log_{10} \left(1 + \frac{\text{range}}{\alpha} \right) + \tau, \quad (4)$$

where $\mu > 0$ is a constant related to the performance of signal receiver and τ is the Gaussian distributed noise introduced by the environmental noise and the thermal noise of the receiver. range is defined as

$$\text{range} = \min \left(\text{range}_{\max} - \sqrt{(x_k^i - x_T)^2 + (y_k^i - y_T)^2}, 0 \right), \quad (5)$$

where $\text{range}_{\max} > 0$ is a constant relating to the sensitivity of the signal receiver.

Bearing measurement is modeled as

$$\text{bear} = \tanh \left(\frac{y_T - y_k^i}{x_T - x_k^i} \right) + w, \quad (6)$$

where w is a noise introduced by the random distortion of the target radar waveform. The probability density function

of w is characterized by a weighted mixture of Gaussian distribution and a ‘‘long-trailing’’ noise distribution (in this paper, a student- t distribution):

$$w = \xi g_t(\psi) + (1 - \xi) g_G(\psi), \quad (7)$$

where g_t is the probability density function of the t distribution, g_G is the probability density function of the Gaussian distribution, and $\xi \in [0, 1]$ is the weight representing the degree of non-Gaussian. Thus, the model of the radar seeker can be written as

$${}^i z_k = \begin{cases} \begin{bmatrix} \text{rss} \\ \text{bear} \end{bmatrix}, & \text{if } \text{rss} \geq 0 \quad |\text{bear} - \varphi_k^i| \leq \varphi_{fov} \\ \emptyset & \text{else,} \end{cases} \quad (8)$$

where φ_{fov} is a constant representing the maximum angle of field of view respect to yaw. It is noted that the sensor model formulated by using (8) is a special form of (3) and the proposed method is also applicable to other linear and nonlinear sensor models in the form of (3).

2.4. Online Patrolling Route Planning as Active Information Gathering. Based on a model predictive control scheme, we only solve the optimization problem of information gathering in a finite time horizon and apply the first optimal control action to agent. This scheme not only reduces the size of the optimization problem that needs to be solved but also adapts to the dynamic environment better.

Given a state vector of loitering munitions swarm $\{\mathbf{s}_k^1, \mathbf{s}_k^2, \dots, \mathbf{s}_k^{n_a}\}$, an unknown but fixed target state $[x_T, y_T, \alpha]$, and a constant of planning horizon T , let R be the notation of the set of all possible r , the information gathering utility is formulated as

$$\begin{aligned} U_T^* &= \arg \max_{U_T} \eta_1 I(R; \mathcal{Z}_T(U_T)) + \eta_2 J(U_T), \\ \text{s.t.} & \quad (2), (8), \\ & \quad \mathbf{u}_t^i \in \mathcal{U}^i, \forall t \in \mathcal{T}, \forall i \in A, \end{aligned} \quad (9)$$

where $I(\cdot, \cdot)$ is the information metric (mutual information in this paper); $\eta_1 \in \mathbb{R}^1$ and $\eta_2 \in \mathbb{R}^{n_j}$ are weights, respectively,

for information quality and other operational objective; $J(U_T) \in \mathbb{R}^{n_j}$ represents the costs of action related to other operational objectives (for the sake of concise and focusing on main theme, let $\eta_2 = [0]_{n_j}$ in the remaining part of this paper), $U_T = \{u_T^1, \dots, u_T^{n_a}\}$ is the union set of the sequential action of each agent in time horizon \mathcal{T} , and $\mathcal{Z}_T = \{Z_T^1, \dots, Z_T^{n_a}\}$ is the union set of measurements of each agent, which is further decomposed as ${}^i Z_T = \{Z_T^i, \dots, Z_T^{n_a}\}$ of which elements contain ${}^i Z_T = \{Z_{k+1}^i, Z_{k+2}^i, \dots, Z_{k+T}^i\}$. Equation (9) can be solved and is guaranteed to have optimal solution because of the finite decision candidates. However, the complexity of computation growth exponential and the communication capacity for completely sharing of information is not practical due to communication capacity limitation. Therefore, a coordinate descent scheme of swarm decision is utilized.

3. Distributed Collaborative Framework

Due to the interference of environmental noise and communication delay, the capacity of the airborne communication is limited, which cannot afford a fully connected two-way communication channel between all agents. To tackle the complexity problem and communication capacity limitation of swarm decision, Atanasov et al. [17] proposed a coordinate descent method. We assume that all agents in the swarm communicate in a packet broadcasting manner similar to the UDP protocol.

Suppose agent 1 plans its route without considering the other robots. In other words, it solely solves a single agent version of problem (9):

$$\begin{aligned} {}^1 U_T^* &= \arg \max_{U_T} I(R; {}^1 \mathcal{Z}_T(U_T)), \\ \text{s.t.} \quad (2), (8), \\ \mathbf{u}_t^1 &\in \mathcal{U}^1. \end{aligned} \quad (10)$$

After solving (10), agent 1 broadcasts its chosen plan and measurement to the other agent in the swarm. Then agent 2 solves a 2-agent version of information gathering problem with the constraint of fixed policy for agent 1:

$$\begin{aligned} U_T^* &= \arg \max_{U_T} I(R; {}^{1:2} \mathcal{Z}_T(U_T)), \\ \text{s.t.} \quad (2), (8), \\ \mathbf{u}_t^1 &\in {}^1 U_T^*, \\ \mathbf{u}_t^2 &\in \mathcal{U}^2. \end{aligned} \quad (11)$$

The algorithm continues in this fashion for each agent in the swarm. The mutual information in optimization objective (9) can be obtained by using the chain rule of mutual information:

$$I(r; \mathcal{Z}_T) = \sum_{i=1}^{n_a} I(R; {}^{1:i} \mathcal{Z}_T). \quad (12)$$

Let ${}^l \Psi_K$ be the notation of the data package broadcasted by agent l and received by agent i :

$${}^i \Psi_K = \begin{cases} \{Z_k^l, \mathbf{s}_k^l, \mathbf{u}_T^l\} & \text{if } l \leq i \\ \emptyset & \text{else.} \end{cases} \quad (13)$$

Given a set of data package $\{{}^i \Psi_K, \dots, {}^{n_a} \Psi_K\}$ and finite planning horizon T , using coordinate descent, the swarm decision can be decomposed as

$$\begin{aligned} \mathbf{u}_T^* &= \arg \max_{\mathbf{u}_T} I(R; {}^{1:i} \mathcal{Z}_T(\mathbf{u}_T^i | \mathbf{u}_T^{\setminus i})), \\ \text{s.t.} \quad (2), (8), (13), \\ \mathbf{s}_{t+1}^i &= f^i(\mathbf{s}_t^i, \mathbf{u}_t^i), \mathbf{s}_{t+1}^l = f^l(\mathbf{s}_t^l, \mathbf{u}_t^l), \\ \mathbf{s}_{t+1}^m &= \mathbf{s}_t^m, \\ \mathbf{u}_t^i &\in \mathcal{U}^i, \forall t \in T, \\ \forall l &\in \{1, \dots, i-1\}, \\ \forall m &\in \{i+1, \dots, n_a\}. \end{aligned} \quad (14)$$

(14) can be solved exactly for that all solutions are finite and enumerable. Comparing to (9), the computation for each agent in swarm grows linearly with the number of agents in the swarm, i.e., $O(n_a |u|^T)$. Thus, it is scalable to large swarm.

4. Mutual Information Computation Using Particle Filtering

To calculate mutual information, the probability distribution of all possible target states and the probability of union distribution of measurement and target state are needed, which is infeasible in practice. In the following part, we compute mutual information by the differential between prior entropy and posterior entropy. Both prior entropy and posterior entropy are estimated by the particle filtering-based method.

4.1. Mutual Information as Entropy Differential. In the objective function of (14), using information theory,

$$\begin{aligned} I(R; {}^{1:i} \mathcal{Z}_T) &= \sum_{t=k+1}^{k+T} I(R; {}^{1:i} Z_{k+1:t}, {}^{1:i} Z_{k+1:t-1}), \\ &= \sum_{t=k+1}^{k+T} H(R | {}^{1:i} Z_{k+1:t-1}) - H(R | {}^{1:i} Z_{k+1:t}, {}^{1:i} Z_{k+1:t-1}), \end{aligned} \quad (15)$$

where $H(\cdot)$ denotes a conditional entropy defined as

$$H(X | Y) = - \int_{y \in Y} \int_{x \in X} \mathbb{P}(x, y) \log \mathbb{P}(x | y) dx dy. \quad (16)$$

$H(R | {}^{1:i} Z_{1:k})$ and $H(R | {}^{1:i} Z_{1:k+1})$ are the prior entropy and posterior entropy respectively, which can be derived by prior probability $p(r | {}^{1:i} z_{1:k})$ and posterior probability $p(r | {}^{1:i} z_{1:k+1})$.

4.2. Target State Distribution Inference. Given target state distribution $p(\mathbf{r} | {}^{1:i}z_{1:k})$ at time k , since the target is stationary, consider the measurement arrived at time $k+1$, the posterior distribution can be derived from Bayes theorem:

$$p(\mathbf{r} | {}^{1:i}z_{1:k+1}) = p(\mathbf{r} | {}^{1:i}z_{1:k}) \frac{\prod_{j=1}^i P_k({}^jz_{k+1} | \mathbf{r})}{\int_{\mathbf{r} \in \mathbb{R}} P_k(\mathbf{r}, {}^{1:i}z_{1:k+1}) d\mathbf{r}} \quad (17)$$

In the finite-state hidden Markov model, the integral part of (17) can be a summation of finite terms. In the linear system and the Gaussian distribution, the posterior distribution is still Gaussian distribution. Here for the non-Gaussian nonlinear of (8), probability density function or the integral part generally does not have close form solutions. In this paper, the particle filtering method is used to approximate the posterior probability density function.

4.3. Sequential Importance Resampling-Based Particle Filtering. The basic idea of particle filtering is to use the Monte Carlo sampling method to obtain conditionally distributed weighted random samples (particles) to approximate the posterior probability density, including the prediction and update steps. The randomly sampled particles in the prediction step update the state according to the dynamics followed by the target. Because it is assumed that the target is stationary in this paper, the prediction step can be omitted. Particles of filter are initially sampled from a distribution of target location given by prior knowledge. In the update step, the particles' weight is updated using the measurement and radar seeker model can be formulated as follows. Let $\langle {}^i\hat{\mathbf{r}}_{p,k}, {}^i\omega_{p,k} \rangle$ be the notation of a group that has n_p particles and its corresponding weights known by agent i at time k , where ${}^i\omega_{p,k} \in \mathbb{R}_+$ and $\sum_{p=1}^{n_p} {}^i\omega_{p,k} = 1$. Equation (17) can be approximated as

$$p(\mathbf{r} | {}^{1:i}z_{1:k}) \approx \sum_{p=1}^{n_p} {}^i\omega_{p,k} \delta(\mathbf{r} - \hat{\mathbf{r}}_{p,k}). \quad (18)$$

When agent i receives the measurements ${}^{1:i}z_{k+1}$, the particle weights are updated according to each measurements ${}^jz_{k+1}$, $j \in 1, \dots, i$ using the radar seeker measurement likelihood probability model and sequential importance sampling/resampling (SISR) method, in which weight of particle is updated using

$$\begin{aligned} & {}^i\omega_{p,k+1}^{(j)} \alpha^i \omega_k^{(j)} p({}^jz_{k+1} | \hat{\mathbf{r}}_{p,k}), \\ & {}^i\omega_{p,k+1} = \prod_j {}^i\omega_{p,k}^{(j)}. \end{aligned} \quad (19)$$

In order to mitigate the particle degeneracy problem, a standard importance resampling step is imposed after the weight of particles is updated. Before the measurements are sent to neighbor agents via wireless communication, the

measurements have to be quantized. The received measurements at each agent is given by

$$\hat{z} = \begin{cases} 0 & q_0 < z < q_1 \\ 1 & q_1 < z < q_2 \\ \dots & \dots \\ L-1 & q_{L-1} < z < q_L, \end{cases} \quad (20)$$

where \hat{z} is the quantized component of measurement z (one of rss or bear) and q_0, q_1, \dots, q_L are the thresholds for a $K = \log_2 L$ bit quantizer. Note that $q_0 = -\infty$ and $q_L = \infty$. The corresponding likelihood function of quantized component measure can be derived from the probability of the quantized measurement value conditioned on the target state:

$$p(\hat{z} | \mathbf{r}) = \begin{cases} Q(q_0 < \hat{z}) - Q(\hat{z} < tq_1) & \hat{z} = 0 \\ Q(q_0 < \hat{z}) - Q(\hat{z} < tq_1) & \hat{z} = 1 \\ \dots & \dots \\ Q(q_{L-1} < \hat{z}) - Q(\hat{z} < tq_L) & \hat{z} = L-1, \end{cases} \quad (21)$$

where $Q(\cdot)$ is the complementary distribution function of the distribution of measurement component, i.e., rss or bear, which can be derived from (5) or (6). It should be noted that the quantized measurement $p(\hat{z} | \mathbf{r})$ is a function of the agent state and radar target state. Thus the quantized likelihood can be used in particle filter weight update step.

5. Simulation Results and Analysis

To evaluate the effectiveness of the method proposed in this paper, the Octave platform is used to conduct a 100-time Monte Carlo simulation experiment in which loitering munitions swarm patrol and search for an unknown static radar target in a predefined combat air zone. We compared the proposed method in this paper with baselines to show the improvement.

5.1. Experiment Conditions. The simulated combat air zone is a plane area of $[-1000, 1000] \times [-1000, 1000]$ (100m per unit). Target radar coordinates are (300, 300), and a constant related to radar transmit power $\alpha = 22.3$. The loitering munition swarm comprises 4 identical agents, and their initial states are, respectively, velocity $v = 30$ (1.75 m/s per unit, about 190 km/h), starting yaw uniformly distributed random number $\varphi \sim U(0, 2\pi)$, fuel remainder $\gamma = 200$, straight flight fuel consumption $c = 1$, steering fuel consumption $d = -1$, maximum steering angle $\omega^i = 15^\circ$, and starting positions are $(-500, -480)$, $(-600, -600)$, $(-500, -720)$, and $(-400, -600)$. The airborne radar seeker measurement model is set to be with a sensitivity constant $\text{range}_{\max} = 600$, $\mu = -66.8$, non-Gaussian noise $\tau \sim N(0, 3^2)$, and $\omega \sim 0.3T(5) + 0.7N(0, \pi^2/12)$. The simulation runs with the following parameters: model predictive control time horizon $T = 3$ and particle filter with particle

TABLE 2: Mean and standard deviation with respect to RMSE and entropy at the end of simulation from 100 simulation runs.

Approaches	Mean and standard deviation with respect to entropy ^a and RMSE ^a at the end of simulation			
	Entropy mean (nat)	Entropy Std dev (nat)	RMSE mean (m)	RMSE Std dev (m)
Random	14.7725	0.1237	557.5756	56.7275
Coverage of field offline planning	13.4732	1.0593	292.5080	181.2276
Gaussian noise-based IEKF	14.7810	0.1383	555.8947	54.0263
Proposed method	12.4540	0.4812	145.4161	73.3486

^aThe lower the better.

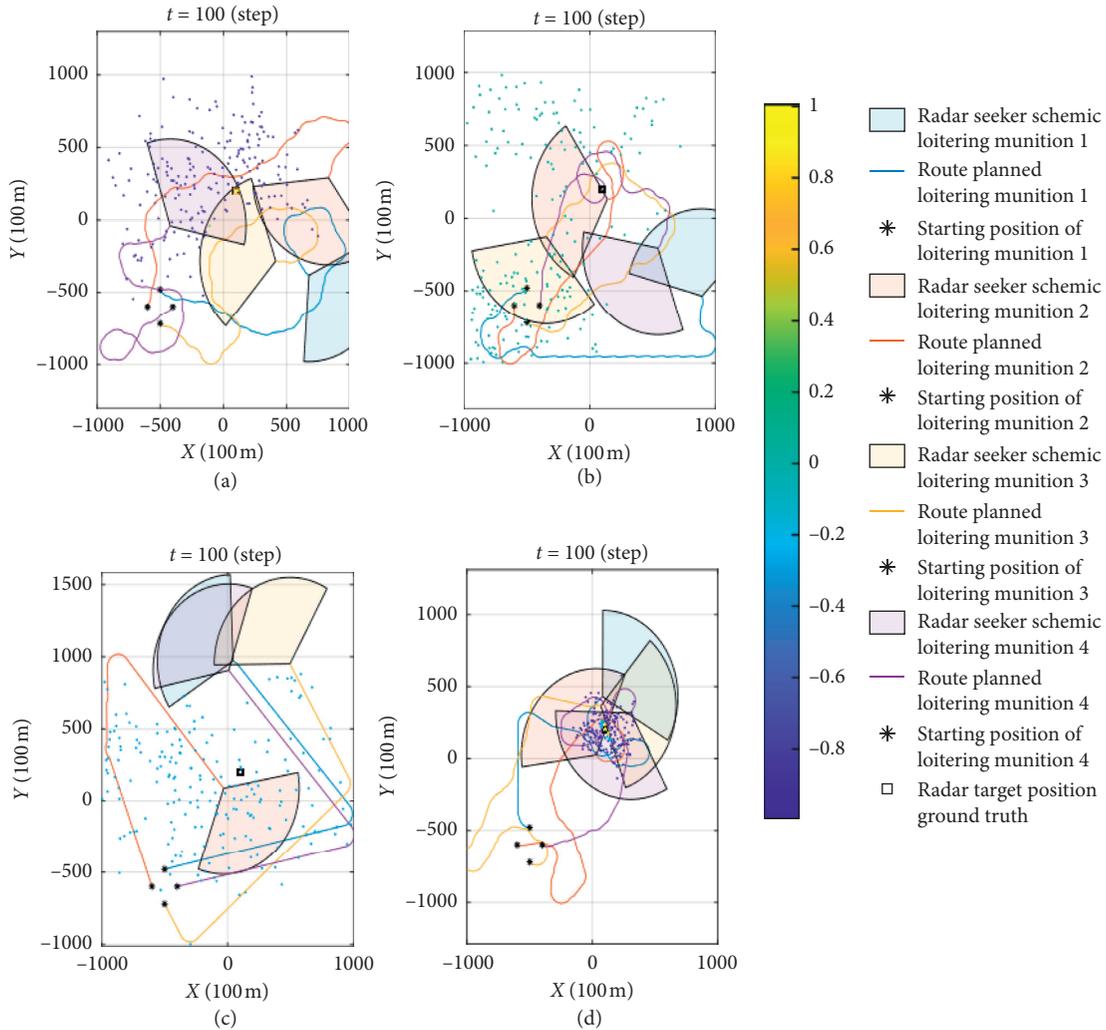


FIGURE 2: Route planned in simulation 1, particles presenting estimation are plotted with scattered color relating to its weight. (a) Random strategy, (b) coverage of field, (c) Gaussian noise IEKF, and (d) the proposed method.

number of $n_p = 200$. The total number of simulation time steps is set to $n_{\text{step}} = 100$.

5.2. Baselines

5.2.1. Random. In this paper, random strategy is selected as the baseline to show the effect of planning. In the random strategy, each loitering munition randomly selects an action in the action space or steering to avoid getting out of combat air zone.

5.2.2. Coverage of Field Offline Planning. This planning method is proposed in [1]. The starting position of route is predefined as the starting of each agent. Because the actual target position is unknown in this paper, the estimated position of radar is used, and the end of route is set to offset the estimated radar target position by (100, 100). Parameters for the algorithm are, respectively, $\text{popsize} = 200$, $\text{pc} = 0.9$, $\text{pm} = 0.05$, and $\text{MAXGEN} = 50$ (for the mean of these parameter, please refer to [1]).

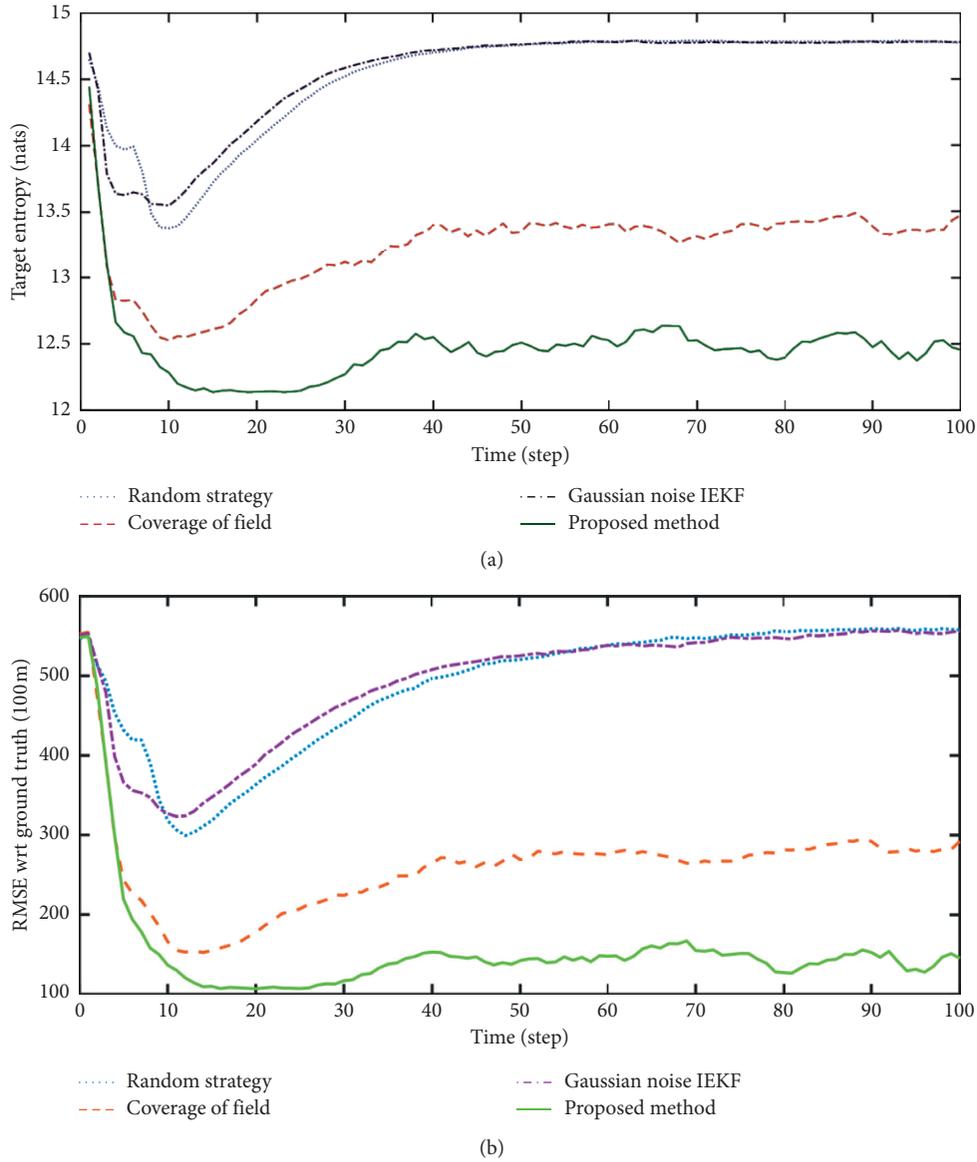


FIGURE 3: Comparison of average target entropy (a) and average RSME (b) along the simulation between baselines and the proposed method.

5.2.3. Gaussian Noise-Based IEKF. To evaluate whether the particle filter is necessary or not, the particle filtering part in this paper is replaced by the iteration extended Kalman filtering in [10], i.e., a first-order linear model is used to approximate the nonlinear seeker model. The non-Gaussian noise is replaced with a Gaussian noise with higher covariance in planning, that is, $\omega \sim N(0, \pi^2/4)$. The noise in simulation is still non-Gaussian and other parts of simulation are unchanged.

5.3. Simulation Results. Simulation results are shown in Table 2 and Figures 2 and 3. We can obtain the following conclusions from the results: (1) As the particle distribution and mean square error value at the end of the simulation shown in Table 2 and Figure 2, the results obtained from the random decision and Gaussian noise-based IEKF are poor.

The method proposed in this paper is better than the coverage of field offline planning in which comparing the former to the latter, the square error value is reduced by more than 30%. (2) The estimation accuracy of the target position measured by RMSE obtained from the Gaussian noise-based IEKF is lower than the method proposed in this paper, indicating that for a highly nonlinear, non-Gaussian noise measurement system, IEKF may divergence. Thus, the u particle filtering method is necessary and effective. (3) Comparing to random strategy, the offline route planning method proposed in [1] improves the “field coverage,” which is beneficial to the estimation of the target state, but the effect is not as good as the method proposed in this paper, which is aiming at optimizing mutual information.

Note that in Figure 3, as the simulation proceeds, both the entropy and RMSE of target rise after the falling. This is

caused by the dynamic constraint of loitering munition, and the loitering munitions must take many steps to turn its heading to achieve a better cooperative measurement. The results also reveal that the performance of the proposed method is better.

6. Conclusions and Future Works

The autonomous online route planning method for loitering munitions swarm proposed in this paper uses the real-time measurement from the airborne radar seeker and model predictive control law to iteratively optimize the mutual information gathered along the route. Compared with the offline planning algorithm based on “field coverage,” our method is more effective. The communication capacity requirement is reduced, and the scalability is enabled by introducing the coordinate descent method for swarm distribution decision. The particle filter method used in this paper, can be applied to nonlinear dynamic models and sensor models with non-Gaussian initial distribution and measurement noise. The method proposed in this paper is also capable to add other objective functions related to operational requirements or extended functions such as obstacle avoidance and threat perception through a hierarchical decision framework, inner/outer loop control [23].

Although the formulation is helpful in real-world scenario, it requires a finite set of action and relatively more computation resource for longer time horizon. Future research work will focus on introducing such box particle filtering and adaptive particle filtering to reduce the computational cost of particle filtering and tackling multiple time-variant radar targets.

Data Availability

The research designed and tested a method framework for path planning without using any data, and there is no data availability need to state. The reader can reproduce our results by the initial state given in the manuscript.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was supported by the National Natural Science Foundation of China under grant no.61473263.

References

- [1] P. B Liu and H. J. sheng, “The route planning of anti-radiation UAV search based on improved artificial potential field method,” *Journal of projectiles, Rockets, Missiles and Guidance*, vol. 1, p. 5, 2020, in Chinese.
- [2] P. B Liu and H. J. sheng, “Route planning of antiradiation UAVs for collaborative search,” *Electronics Optics & Control*, vol. 27, no. 3, pp. 65–68, 2020, in Chinese.
- [3] Z. Kang, H. Ling, T. Zhu, and H. Luo, “Coverage flight path planning for multi-rotor UAV in convex polygon area,” in *Proceedings of the 2019 Chinese Control and Decision Conference (CCDC)*, pp. 1930–1937, Nanchang, China, June 2019.
- [4] Institute of Actuaries of India, *Fire And Forget: Harpy Is An Autonomous Weapon For All Weather*, Institute of Actuaries of India, Navi Mumbai, India, 2020, <https://www.iai.co.il/p/harpy>.
- [5] X. Shao, B. Tian, and W. Yang, “Estimator-based MLP neuroadaptive dynamic surface containment control with prescribed performance for multiple quadrotors,” *Aerospace Science and Technology*, vol. 97, Article ID 105620, 2020.
- [6] B. Koopman, *The Theory of Search*, Institute for Defense Analyses, Arlington, VA, USA, 1956.
- [7] L. D. Stone, *Theory of Optimal Search*, INFORM, New York, NY, USA, 1975.
- [8] Q. ZHU, *Optimal Search Theory in Continuous and Discrete Space*, Science Press, Henderson, NV, USA, 2005, in Chinese.
- [9] T. Zhu, W. He, H. Ling, and Z. Zhang, “Cooperative search strategies of multiple UAVs based on clustering using minimum spanning tree,” in *Proceedings of the International Conference on Sensing and Imaging*, pp. 112–121, Liuzhou City, China, October 2018.
- [10] B. Schlotfeldt, D. Thakur, N. Atanasov, V. Kumar, and G. J. Pappas, “Anytime planning for decentralized multirobot active information gathering,” *IEEE Robotics and Automation Letters*, vol. 3, no. 2, pp. 1025–1032, 2018.
- [11] J. Carroll, H. Hmedi, and A. Arapostathis, “Optimal scheduling of multiple sensors which transmit measurements over a dynamic lossy network,” in *Proceedings of the 2019 IEEE 58th Conference on Decision and Control (CDC)*, pp. 684–689, Nice, France, December 2019.
- [12] B. J. Julian, M. Angermann, M. Schwager, and D. Rus, “Distributed robotic sensor networks: an information-theoretic approach,” *The International Journal of Robotics Research*, vol. 31, no. 10, pp. 1134–1154, 2012.
- [13] S. Moon and E. W. Frew, “A communication-aware mutual information measure for distributed autonomous robotic information gathering,” *IEEE Robotics and Automation Letters*, vol. 4, no. 4, pp. 3137–3144, 2019.
- [14] X. Shao, L. Wang, J. Li, and J. Liu, “High-order ESO based output feedback dynamic surface control for quadrotors under position constraints and uncertainties,” *Aerospace Science and Technology*, vol. 89, pp. 288–298, 2019.
- [15] M. M. Zavlanos, L. Spesivtsev, and G. J. Pappas, “A distributed auction algorithm for the assignment problem,” in *Proceedings of the IEEE Conference on Decision and Control*, pp. 1212–1217, Cancun, Mexico, December 2008.
- [16] N. Atanasov, J. Le Ny, K. Daniilidis, and G. J. Pappas, “Information acquisition with sensing robots: algorithms and error bounds,” in *Proceedings of the 2014 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 6447–6454, Hong Kong, China, May 2014.
- [17] N. Atanasov, J. Le Ny, K. Daniilidis, and G. J. Pappas, “Decentralized active information acquisition: theory and application to multi-robot SLAM,” in *Proceedings of the 2015 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 4775–4782, Seattle, WA, USA, May 2015.
- [18] S. Chen, *Planning for Reconnaissance and Monitoring Using UAV Swarms*, PhD Thesis, Graduate School of National University of Defense Technology, Changsha, Hunan, China, 2016, In Chinese.
- [19] Y. Xue, H. Chen, J. Chen, and J. Wang, “Pipeline Gaussian particle filter and hardware design for attitude estimation with UAV,” *Mathematical Problems in Engineering*, vol. 2020, Article ID 7092526, 6 pages, 2020.

- [20] X. Shao and Yi Shi, "Neural adaptive control for MEMS gyroscope with full-state constraints and quantized input," *IEEE Transactions on Industrial Informatics*, vol. 16, no. 10, pp. 6444–6454, 2020.
- [21] X. Shao, Si Haonan, and W. Zhang, "Fuzzy wavelet neural control with improved prescribed performance for MEMS gyroscope subject to input quantization," *Fuzzy Sets and Systems*, 2020.
- [22] J. W. Eaton, D. Bateman, S. Hauberg, and R. Wehbring, *A High-Level Interactive Language For Numerical Computations*, GNU Octave version 5.2.0 Manual, Norway, Europe, 2019.
- [23] W. Li, Z. Wang, Q. Liu, and L. Guo, "An information aware event-triggered scheme for particle filter based remote state estimation," *Automatica*, vol. 103, pp. 151–158, 2019.