2024 23rd ACM/IEEE International Conference on Information Processing in Sensor Networks (IPSN)

Demo Abstract: CARL: Collaborative Altitude-Adaptive Reinforcement Learning for Active Search with UAV Swarms

Chen-Chun Hsia*
Shenzhen International Graduate School (SIGS), Tsinghua University
Shenzhen, China
xcj21@mails.tsinghua.edu.cn

Yanggang Xu*
Shenzhen International Graduate School (SIGS), Tsinghua University
Shenzhen, China
xyg22@mails.tsinghua.edu.cn

Jiyuan Ren*
Shenzhen International Graduate School (SIGS), Tsinghua University
Shenzhen, China
rjy22@mails.tsinghua.edu.cn

Xinlei Chen†
Shenzhen International Graduate School, Tsinghua University
Pengcheng Laboratory
RISC-V International Open Source Laboratory
Shenzhen, China
chen.xinlei@sz.tsinghua.edu.cn

ABSTRACT

Sensing noise and complex decision-making pose critical challenges to active search for lost persons amid disasters, impeding efficient rescue efforts. We introduce CARL, a collaborative altitude-adaptive reinforcement learning framework for UAV swarms. CARL integrates confidence-informed assessment with Sparse Bayesian Learning to diminish the noise impact on sensor performance, and an altitude-adaptive planner for collaborative active search strategy. Simulation experiments with up to 50 targets and 10 UAVs demonstrate CARL's superior performance compared to baseline methods in lost person active search scenarios.

KEYWORDS

Reinforcement learning, Bayesian learning, Collaborative UAV swarms, Active search framework

1 INTRODUCTION

The urgency of efficient search and rescue operations following natural and man-made disasters is underscored by the World Disaster Report [3], which reveals that nearly 80% of urban disaster survivors are surface victims, visibly ensnared amidst rubble. This highlights the critical role of Unmanned Aerial Vehicles (UAVs) in performing active searches to rapidly locate and rescue these victims. Active target search, a key component of search-and-rescue missions, involves identifying an unknown number of targets in varied environments, a task for which UAVs are uniquely suited due to their agility and aerial perspective.

The challenges facing multi-UAV active search in large-scale, three-dimensional environments[5, 6] with sparse targets are considerable. First, the sparse target distribution complicates decision-making, often leading to myopic methods that fail to find global optima. Second, UAVs' partial observations hinder their understanding of the search area and ability to efficiently locate survivors. Third, the explore-exploit dilemma presents a significant obstacle: UAVs must choose between broader coverage with increased noise at higher altitudes or improved precision with narrower views at lower altitudes. These challenges critically impact the UAVs' ability to accurately locate surface victims, determining the mission's success.

Our research introduces a novel approach that significantly enhances the target recovery rate by integrating an altitude-adaptive planner into a multi-UAV collaborative search framework. By incorporating the impact of altitude on sensing quality, our method enables UAVs to dynamically adjust their sensing strategies based on their flying height. We utilize Sparse Bayesian Learning (SBL)to effectively mitigate the noise impact on the sensing model. Concurrently, a confidence-informed assessment mechanism, empowered by Long Short-Term Memory (LSTM) networks for capturing temporal dependencies and Proximal Policy Optimization (PPO) for policy learning[1, 4], enhances the UAVs' decision-making in coordinating their search efforts. By addressing altitude-dependent sensing noise and complex multi-UAV coordination, our approach substantially improves target search efficiency.

2 SYSTEM MODEL

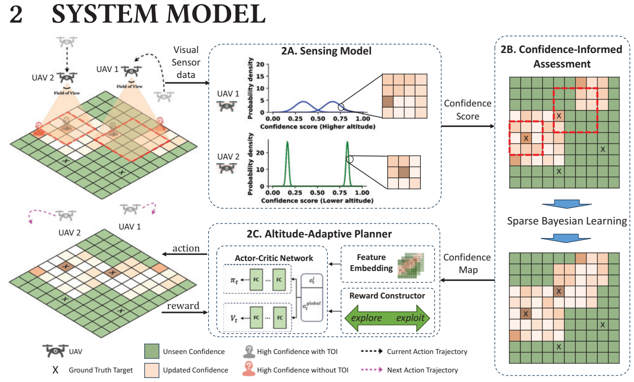


Figure 1: Scenario with CARL framework.

In this paper, we employ n UAVs, equipped with visual sensors and sensing models, to collaborative active search for m Target of Interest (TOIs).

*These authors contributed equally to this research.
†Xinlei Chen is the corresponding author.

979-8-3503-6201-5/24/$31.00 ©2024 IEEE
DOI 10.1109/IPSN61024.2024.00025

Sensing model: Ideally, the output of the visual sensing models, denoted as $\Phi_p \in \{0, 1\}$, indicates whether a TOI exists at location p ($\Phi_p = 1$) or not ($\Phi_p = 0$). But practically, there is a significant correlation between the sensing error and the UAV's altitude h [3]. Therefore, the TOI confidence score of the sensing algorithm is modeled as $\phi_p = clip(\varphi_p, 0, 1),\ \varphi_p \sim \mathcal{N}(\mu(\Phi_p, h), \sigma^2(h))$.

Confidence informed assessment: A confidence map $\mathcal{B} \in \mathbb{R}^{L_1 \times L_2}$ is employed to characterize the entire environment. Each value $P(\psi_p)$ in $\mathcal{B}$ is initialized to 0.5 at $t = 0$, and $P(\psi_p)$ represents the probability of the existence of a TOI at position p initially. And based on SBL, we define the confidence update rule as

$$\mathcal{B}_t^i = \frac{\beta_t^i \cdot \mathcal{B}_t^{i-1}}{(1-\beta_t^i)\cdot(1-\mathcal{B}_t^{i-1}) + \beta_t^i \cdot \mathcal{B}_t^{i-1}}, \mathcal{B}_t^0 = \mathcal{B}_t,\ i \in (1, ..n)$$

, where β_t^i contains confidence scores ϕ measured by UAV i. Additionally, the length F_l and width F_w of the field of view of UAVs exhibit a linear relationship with h, i.e. $F_l = k_l h$ and $F_w = k_w h$.

Thus, the proposed algorithm needs to plan the action trajectory $\tau = \{a_1, a_2, \cdots\}$ of UAV swarms to maximize the objective, i.e. $\tau* = \arg\max_\tau \sum_{p,t} I(\psi_p; \tau)/T$. And we have

$$I(\psi_p; \tau) = \begin{cases} w_e log \frac{P(\psi_p|\tau)}{P(\psi_p)}, & \text{if } \Phi_p = 1 \\ w_f log \frac{1-P(\psi_p|\tau)}{1-P(\psi_p)}, & \text{if } \Phi_p = 0 \end{cases}, \tag{1}$$

where w_e and w_f are importance weights, and $P(\psi_p|\tau)$ denotes the TOI existence probability after UAVs adopt trajectory τ.

Adaptive planner: Our adaptive planner synthesizes confidence map, UAV altitude, and observations into a state encoding $S_t = \mathcal{B}(x, y), A(x, y), O(x, y)$, capturing essential information for dynamic decision-making. The reward function, comprising entropy reduction: $r_1 = \mathcal{I}_{t+1} - \mathcal{I}_t$ and target-space discrimination:$r_2 = w_h \sum_{p_h} \Delta\phi_t^{p_h} + w_e \sum_{p_e} \Delta\phi_t^{p_e}$, refined by a temporal discount: $(w_1 \cdot r_1 + w_2 \cdot r_2) \cdot T_{max} - t/T_{max}$, to balance efficient exploration with rapid target detection.

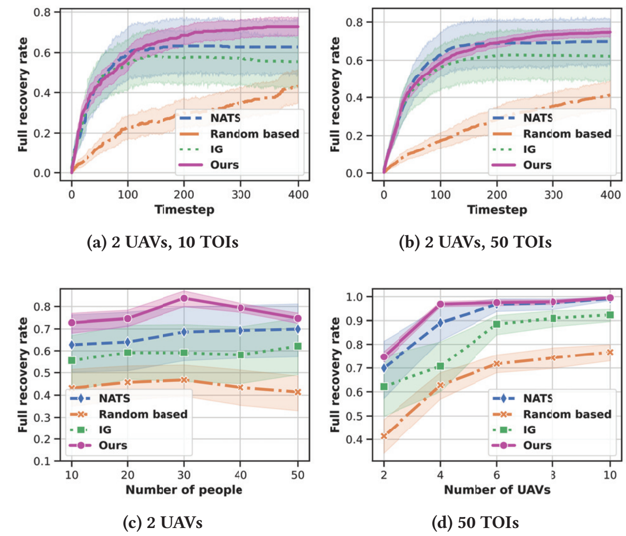
We standardize observations for input consistency and leverage a CNN-LSTM architecture for spatial-temporal feature extraction, vital for navigating changing environments. To ensure synchronized UAV teamwork in complex searches, the planner employs the PPO strategy, facilitating effective, coordinated actions.

3 EVALUATION

In the experiment, we validate the proposed method in simulation by comparing it with the NATS [2], information gain (IG) based and random based methods. And full recovery rate $\kappa = \sum_p \mathbb{1}_{\{P(\psi_p)>0.8, \Phi_p=1\}}/m$ is used to evaluate the model. In figure 2, (a) and (b) depict the results when using 2 UAVs to search for 10 and 50 TOIs. (c) and (d) illustrate the results when 2 UAVs are employed to search for varying numbers of TOIs and when different numbers of UAVs are used to search for 50 TOIs. The results indicate that the proposed model outperforms other methods.

4 CONCLUSION

In this paper, we propose CARL for TOI active search. The approach leverages reinforcement learning and a confidence-informed assessment based on SBL to achieve superior performance. Simulations



(a) 2 UAVs, 10 TOIs

(b) 2 UAVs, 50 TOIs

(c) 2 UAVs

(d) 50 TOIs

Figure 2: Experiment validation.

validate CARL's effectiveness, outperforming state-of-the-art methods. Future work will focus on enhancing CARL's architecture to better capture spatial-temporal dependencies and address real-world challenges in search and rescue missions.

ACKNOWLEDGMENTS

This paper was supported by the National Key R&D program of China No. 2022YFC3300703, the Natural Science Foundation of China under Grant No. 62371269. Guangdong Innovative and Entrepreneurial Research Team Program No. 2021ZT09L197, Shenzhen 2022 Stabilization Support Program No. WDZC20220811103500001, and Tsinghua Shenzhen International Graduate School Cross disciplinary Research and Innovation Fund Research Plan No. JC20220011.

REFERENCES

[1] Xuecheng Chen, Haoyang Wang, Zuxin Li, Wenbo Ding, Fan Dang, Chenye Wu, and Xinlei Chen. 2022. DeliverSense: Efficient Delivery Drone Scheduling for Crowdsensing with Deep Reinforcement Learning. https://doi.org/10.1145/3544793.3560412

[2] Ramina Ghods, William J. Durkin, and Jeff Schneider. 2021. Multi-Agent Active Search using Realistic Depth-Aware Noise Model. In *2021 IEEE International Conference on Robotics and Automation (ICRA)*. 9101–9108. https://doi.org/10.1109/ICRA48506.2021.9561598

[3] Ajith Anil Meera, Marija Popović, Alexander Millane, and Roland Siegwart. 2019. Obstacle-aware Adaptive Informative Path Planning for UAV-based Target Search. In *2019 International Conference on Robotics and Automation (ICRA)*. 718–724. https://doi.org/10.1109/ICRA.2019.8794345

[4] Jiyuan Ren, Yanggang Xu, Zuxin Li, Chaopeng Hong, Xiao-Ping Zhang, and Xinlei Chen. 2023. Scheduling UAV Swarm with Attention-based Graph Reinforcement Learning for Ground-to-air Heterogeneous Data Communication. 670–675. https://doi.org/10.1145/3594739.3612905

[5] Haoyang Wang, Xuecheng Chen, Yuhan Cheng, Chenye Wu, Fan Dang, and Xinlei Chen. 2022. H-SwarmLoc: Efficient Scheduling for Localization of Heterogeneous MAV Swarm with Deep Reinforcement Learning. In *Proceedings of the 20th ACM Sensys*. 1148–1154.

[6] Haoyang Wang, Jingao Xu, Chenyu Zhao, Zihong Lu, Chen Cheng, Yuhan, Xuecheng Chen, Xiao-Ping Zhang, Yunhao Liu, and Xinlei Chen. 2024. Transform-Loc: Transforming MAVs into Mobile Localization Infrastructures in Heterogeneous Swarms. In *Proceedings of the IEEE INFOCOM*.