

Wei LI, Bowei YANG, Guanghua SONG, Xiaohong JIANG, 2021. Dynamic value iteration networks for the planning of rapidly changing UAV swarms. *Frontiers of Information Technology & Electronic Engineering*, 22(5):687-696.

<https://doi.org/10.1631/FITEE.1900712>

Dynamic value iteration networks for the planning of rapidly changing UAV swarms

Key words: Dynamic value iteration networks; Episodic Q-learning; Unmanned aerial vehicle (UAV) ad-hoc network; Non-dominated sorting genetic algorithm II (NSGA-II); Path planning

Corresponding author: Bowei YANG

E-mail: boweiy@zju.edu.cn

 ORCID: <https://orcid.org/0000-0001-8581-3817>

Motivation

1. Following the rapid deployment of UAV ad-hoc networks (UANETs), the adaptation of artificial intelligence (AI) solutions to mobile network environments has become the research focus in today's information network era.
2. An advanced UANET communication architecture is the basis for efficient communication among UAV swarms. The successful completion of a UAV mission depends on a high-quality and efficient communication network.
3. Sparse and rapidly mobile UAVs/nodes can dynamically change the topology of UANETs, which may critically affect UAV communication performance.

Main idea

1. Value iteration networks (VINs) that are based on a CNN strategy can be effectively trained to plan several unknown static scenarios and solve complex multi-objective problems.
2. To improve the communication performance among dynamic UAVs in UANET, a dynamic value iteration network (DVIN) model is proposed based on VINs.
3. Compared with the conventional multi-UAV collaborative planning methods, such as NSGA-II, the DVIN model reduces the decision-making time for UAV/node path planning with a high average success rate.

Method

The architecture of DVIN

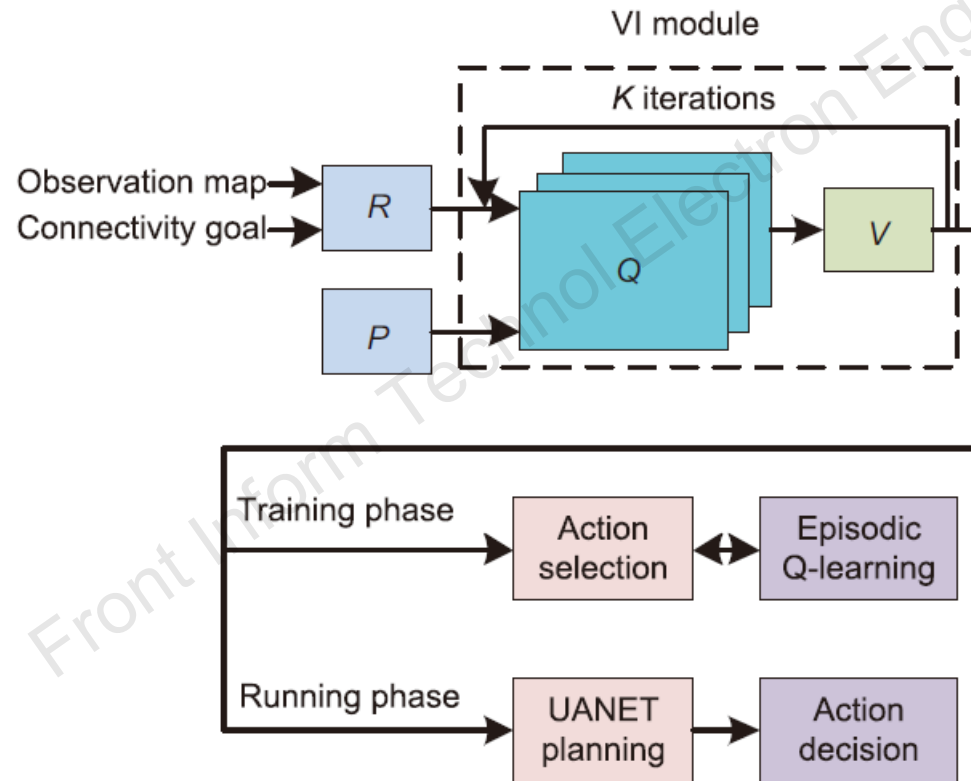


Fig. 2 Architecture of a dynamic value iteration network (DVIN)

Method (Cont'd)

- First, we concatenate the observation map data with a connectivity goal map. The observation map is an observation of the UANET scenario, including UAVs' positions and obstacle positions. The connectivity goal map is generated from the observation map, which is a one-hot representation of the aim ("1" represents a non-test UAV/node that can communicate with the test UAV/node, and "0" otherwise).
- Second, we map the concatenated data into a reward map R through two convolution operations.

Method (Cont'd)

- Third, we input the reward map R and the state transition probability P into the VI module and obtain the state value map V . In the action selection (AS) module, the state value map V is considered as an instructive state value function, which is based on the ε -greedy policy to select an action.
- Finally, in continuously moving scenarios, any selected action is trained using a trial-and-error method and backpropagation by episodic Q-learning to obtain the training parameters θ of the DVIN model.
- In the testing phase, parameters θ are used for node path planning and final action decision-making for the UANET.

Major results

One of the UANET scenarios is emulated by Mininet-WiFi

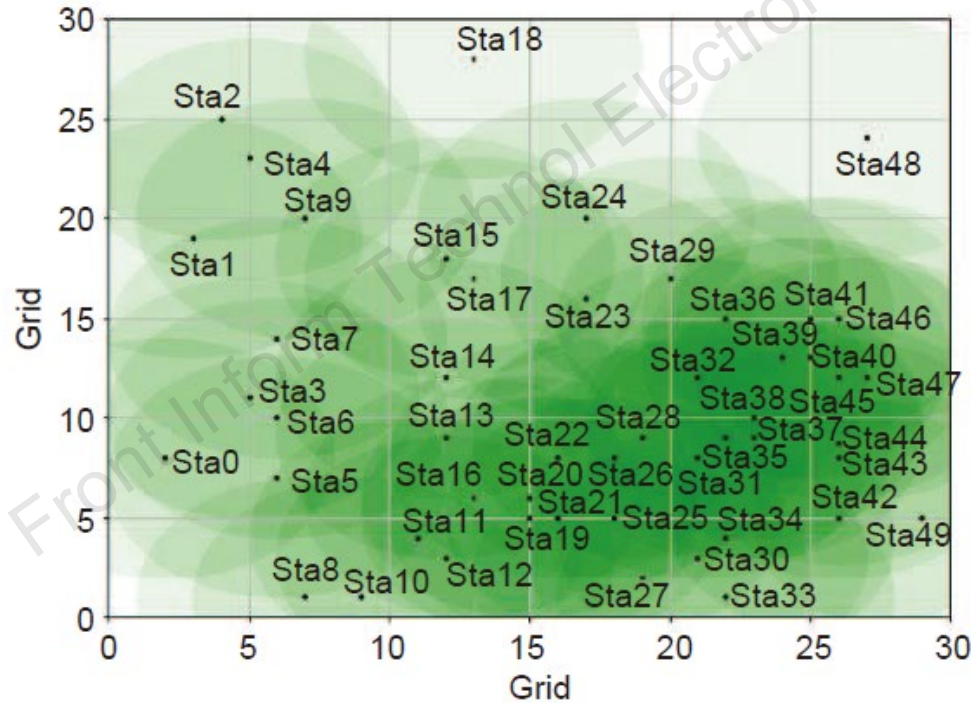


Fig. 5 A simulation result of a 2D maze for UAV swarms

Major results (Cont'd)

Training results of the DVIN model

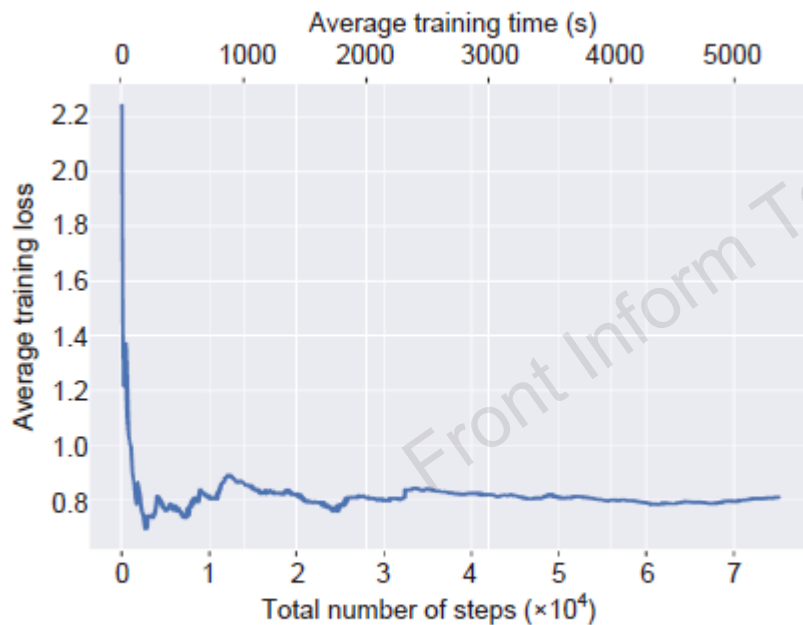


Fig. 6 Average loss of the training epochs (75 160 steps, 5364.14 s per epoch)



Fig. 7 Average reward value of each training epoch for DVIN

Major results (Cont'd)

The performance comparison results of DVIN model, NSGA-II, and the exhaustive method (EM)

Table 1 Performance comparison using different methods

Algorithm	Success rate (%)	Average time per episode (s)
DVIN	34.30	1.19
NSGA-II	30.41	43.98
EM	42.81	81.31

Conclusions

- Although the DVIN model has a greater training time, the decision-making time in the running phase is considerably less than that of NSGA-II. Thus, the DVIN model achieves higher real-time re-planning efficiency.
- In a UANET, it is difficult to achieve full network connectivity. However, using the DVIN model, we can identify an optimal trajectory in 1 or 2 s, which ensures efficient UANET communication.
- The average success rate of each episode of the DVIN model is higher than that of models using NSGA-II.



Wei LI received the BS degree in information security from North China Electric Power University, China, in 2018. He is currently pursuing the MS degree in School of Aeronautics and Astronautics, Zhejiang University, China. His research interests include artificial intelligence, UAV ad-hoc networks, software defined network, etc.



Bowei YANG received the PhD degree in computer science and technology from Zhejiang University, China, in 2011. Currently, he is an associate professor with the School of Aeronautics and Astronautics, Zhejiang University, since 2016. His main research interests include software-defined satellite networking and UAV ADHOC networks, AI-assisted wireless communications technology, AI-big data joint optimization in cellular networks and D2D communications, etc.