

## ІНФОРМАЦІЙНІ ТЕХНОЛОГІЇ

UDC 004.8

DOI <https://doi.org/10.35546/kntu2078-4481.2023.4.25>

Y. O. ALBREKHT

Postgraduate Student at the Department of Information Systems and Technologies  
National Technical University of Ukraine  
“Igor Sikorsky Kyiv Polytechnic Institute”  
ORCID: 0000-0003-0093-6397

A. V. PYSARENKO

PhD, Associate Professor,  
Associate Professor at the Department of Information Systems and Technologies  
National Technical University of Ukraine  
“Igor Sikorsky Kyiv Polytechnic Institute”  
ORCID: 0000-0001-7947-218X

## DECISION-MAKING HETEROGENEOUS UAV SWARM SYSTEM WITH NEURAL NETWORK-ENHANCED REINFORCEMENT LEARNING

*This article explores how artificial intelligence and automation are significantly impacting unmanned aerial vehicles (UAVs), moving from traditional roles to versatile applications.*

*The paper addresses the problem of optimizing the composition of a UAV swarm for efficient task execution by proposing an expert decision-making system that integrates neural networks and reinforcement learning. This system dynamically selects the optimal configuration for heterogeneous UAV swarms, in particular, for searching for objects in unfamiliar terrain. In the experimental phase, an advanced level of system was implemented by combining neural networks and reinforcement learning, based on role-based and MADDPG algorithms for heterogeneous UAV swarms. Decentralized information fusion-based swarm decision making algorithm (IFDSDA) is presented to overcome communication obstacles.*

*The experiment presents a concept for improving heterogeneous UAV swarms using a neural decision network based on reinforcement learning. The environment is represented by a three-dimensional space with objects to be searched in random locations. The neural network evolves its decision-making strategy during training episodes, having an architecture with an input layer that processes information about the UAV's state, hidden layers, and an output layer that influences the swarm's behavior. The paper describes the process of direct propagation, reward-based weight adjustment, and the role of the output layer in determining collective actions.*

*The results demonstrate the effective distribution of UAV types by the swarm based on a neural network, reducing redundancy and resource waste, thereby increasing overall efficiency. The article highlights the optimal solution obtained during the experiment, accompanied by a visual representation of the reward results.*

**Key words:** swarm of UAVs, heterogeneous swarm, reinforcement learning, decision-making system, heterogeneous swarm of UAVs.

Й. О. АЛЬБРЕХТ

аспірант кафедри інформаційних систем та технологій  
Національний технічний університет України  
«Київський політехнічний інститут імені Ігоря Сікорського»  
ORCID: 0000-0003-0093-6397

А. В. ПИСАРЕНКО

кандидат технічних наук, доцент,  
доцент кафедри інформаційних систем та технологій  
Національний технічний університет України  
«Київський політехнічний інститут імені Ігоря Сікорського»  
ORCID: 0000-0001-7947-218X

## СИСТЕМА ПРИЙНЯТТЯ РІШЕНЬ З ВИКОРИСТАННЯМ НАВЧАННЯ З ПІДКРІПЛЕННЯМ ДЛЯ КЕРУВАННЯ ГЕТЕРОГЕННИМИ РОЯМИ БПЛА

*У статті досліджується, як штучний інтелект і автоматизація суттєво впливають на безпілотні літальні апарати (БПЛА), переходячи від традиційних ролей до універсальних застосувань.*

*У статті вирішується проблема оптимізації складу рою БПЛА для ефективного виконання завдань, пропонуючи систему прийняття рішень, яка інтегрує нейронні мережі та навчання з підкріпленням. Ця система*

динамічно обирає конфігурацію гетерогенних роїв БПЛА, зокрема, для пошуку об'єктів на незнайомій місцевості. На експериментальній фазі було впроваджено вдосконалений рівень системи шляхом поєднання нейронних мереж і навчання з підкріпленням, на основі рольових алгоритмів та алгоритму MADDPG. Для подолання перешкод зв'язку представлений децентралізований алгоритм прийняття рішень роєм на основі інформаційного злиття (IFDSDA).

Експеримент демонструє покращення роботи системи гетерогенних роїв БПЛА з нейронною мережею для прийняття рішень на основі навчання з підкріпленням. Навколишнє середовище представлено тривимірним простором де знаходяться об'єкти для пошуку у випадкових місцях. Нейронна мережа розвиває свою стратегію прийняття рішень протягом навчальних епізодів, маючи архітектуру з вхідним шаром, що обробляє інформацію про стан БПЛА, прихованими шарами, і вихідним шаром, що впливає на поведінку рою. У статті описано процес прямого поширення, коригування ваг на основі винагороди та роль вихідного шару у визначенні колективних дій.

Результати демонструють ефективний розподіл роєм типів БПЛА на основі нейронної мережі, зменшення надмірності та втрат ресурсів, тим самим підвищуючи загальну ефективність. У статті висвітлено рішення, отримане в ході експериментів, що супроводжується візуальним представленням результатів.

**Ключові слова:** рій БПЛА, гетерогенний рій, навчання з підкріпленням, система прийняття рішень, гетерогенний рій БПЛА.

### Introduction

In the field of robotics, the convergence of AI and automation has catalyzed remarkable transformations in the capabilities of UAVs. Originally confined to roles such as remote sensing and surveillance, UAVs have evolved into multifaceted platforms with applications spanning diverse industries [1]. This evolution has been characterized by breakthroughs that have shattered traditional limitations and paved the way for innovation.

Among these advancements, the emergence of UAV swarms has been particularly intriguing. A departure from single-UAV approaches, swarm technology harnesses the collective power of multiple drones operating harmoniously to accomplish tasks previously deemed infeasible. Drawing inspiration from nature's collective behaviors, such as bird flocks and insect colonies, UAV swarms exemplify the strength in unity [2]. But the question arises if the system would work better with some decision-making tool to select the best setup of the consisting elements of the heterogeneous UAV swarm.

### Description of the problem

In the usage of heterogeneous UAV swarms, a common issue appears to be solved. There is a need to find the optimal number of each type of available UAV to perform the task the most efficient way. Solving this problem can be difficult and making calculations every time can be time-consuming. The solution would be to create a decision-making expert system over the neural network that is controlling the UAVs that would be able to find the optimal set of UAVs for a given task. In this article there was an attempt of finding a proof of concept of having this expert system on a simple but common task on searching objects in unknown area by heterogeneous UAV swarms.

### Elevating Swarm Efficiency via Neural Networks and Reinforcement Learning

In a bid to optimize the potential of UAV swarms, the experiment in the article uses an additional layer of sophistication. This method has been already used in other experiments [3], but not yet in UAV swarms. Building upon the foundation of heterogeneous UAV swarms, an advanced layer is injected into the equation: the fusion of neural networks and reinforcement learning. The goal is clear – to construct a system capable of intelligently distributing distinct UAV types within the swarm, orchestrating their collaboration to achieve optimal performance for specific tasks.

### Heterogeneous UAV swarms

There are already researches on the topic of heterogeneous UAV swarms [4]. The proposed role-based MADDPG algorithm is the base of this research. It not only enables the tracking of multiple targets but also fosters exploration for undiscovered targets via a Voronoi-based rewarding policy. The algorithm's effectiveness is demonstrated through comprehensive implementation, testing, and validation in a simulation environment. Following this, the approach is assessed using a real-world multi-robot system featuring micro drones.

For the experiment the Python gym library was used. This library was created by OpenAI Company. And it has one advancement in comparison between the other methods – it has the possibility to add custom environments [5].

### Decentralized decision-making algorithm

The study [6] introduces the Information-Fusion based Decentralized Swarm Decision Algorithm (IFDSDA) for coordinating UAV swarms in situations of communication interference or failure. Each UAV uses a monocular camera to perceive the area ahead, and the IFDSDA employs an information fusion strategy to integrate communication and visual perception data. This enables UAVs to effectively utilize different information in the absence of communication. The decentralized swarm decision module, controlling each UAV, generates heading orientation based on the fused information and basic action rules. Weight parameters for the combination are optimized using a heuristic genetic algorithm offline. Simulations demonstrate the proposed method's effectiveness, scalability, and robustness compared to the ISOA method and its variant. The study [6] aims to reduce swarm dependence on network communication and enhance adaptability in complex battlefield environments.

Large-scale UAV swarms have diverse applications, including express logistics, agricultural plant protection, emergency relief, and reconnaissance. Collaborative decision-making is vital for autonomous UAV swarms, often relying on centralized or decentralized control with wireless communication. However, these approaches face limitations due to interference or unreliable data links. Decentralized swarm decision models assume ideal communication, but they become inefficient during communication outages. The study [6] addresses this issue by proposing the IFDSDA, focusing on decentralized swarm decision-making under communication interference.

The research landscape involves improving communication network invulnerability, exploring novel swarm decision-making mechanisms, and compensating for UAV perception in failure cases. The IFDSDA contributes by presenting:

- a decentralized algorithm composed of information fusion and decision-making processes, enhancing swarm scalability without global information dependence;
- an information fusion strategy integrating communication and visual perception data, increasing swarm reliability and adaptability;
- macro-level swarm behaviors generated through basic action rules, allowing flexibility and optimization through a genetic algorithm;
- simulation validation of IFDSDA's effectiveness, scalability, and adaptability in collision/obstacle avoidance and area search missions.

### Experiment Setup

This experiment introduces an elevated setup, augmenting the conventional heterogeneous UAV swarm with a decision-making neural network with reinforcement learning. This neural network acts as the conductor, directing the swarm's synergy and evolution based on acquired knowledge.

#### *Environment Formation.*

An environment is designed to challenge the swarm's abilities in a search for object groups within an unknown area. This controlled environment mirrors real-world complexities, providing a testing ground for the enhanced swarm. During the experiment, the environment consists of a number of objects the agents have to first find, and then continue following until the time of the experiment ends.

The environment consists of a 3 dimensional field in 100x100x100 area and N objects in random placement, where N – the number of UAVs used in the experiment.

#### *Heterogeneous UAV Compositions.*

This diversity empowers the swarm with a versatile skill set that can tackle multifaceted challenges. The system consists of two different types of agents with one of them being slower (slower max speed) but with the ability to detect objects faster (from a bigger radius) and the second with these parameters reversed.

#### *Neural Network Architecture.*

To steer the swarm's actions intelligently, a neural network was added. Guided by reinforcement learning algorithms, the network takes into account each UAV's state and steers the swarm towards actions that maximize rewards within the environment. Over time, the network evolves, adapting its decision-making strategy to optimize outcomes. The decision-making strategy is also built as another layer of a neural network to determine which setup of UAV agents would be the best solution for the given scenario.

### Algorithm Structure

The algorithm governing the training of the heterogeneous UAV swarm, empowered by neural networks and reinforcement learning, can be outlined as follows:

```

begin
  init_state = initialize the UAV environment
  params = define reinforcement learning parameters
  decision_params = define reinforcement learning parameters of the decision making layer
  model = define neural network model, including the values that state the type of
objects in the swarm
  metrics = define metrics
  N = number of episodes
  T = number of timesteps
  S = number of UAVs in the swarm
  for training_episode = 0 to N do:
    environment = init_state
    reward = 0
    for time_step = 0 to T do:
      for i = 0 to S do:
        observe_current_state(UAV[i])
        set_next_decision(UAV[i])

```

```

endfor
environment = next_step(model, environment)
reward = calculate_reward(environment)
update_network(model, params, reward)
endfor
record_metrics_for_episode(reward, training_episode, metrics)
update_decision_layer(model, decision_params, reward)
endfor
create_diagram(metrics)
end

```

### Neural Network Architecture

The neural network that drives the decision-making process within enhanced UAV swarm is structured to extract and process relevant information from each UAV's state. This information is used to determine the most suitable action for the swarm as a collective entity.

Let's look at the neural network architecture in more detail.

*Input Layer:*

The input layer of the neural network accepts the state information of each UAV as input. Let  $s_i$  denote the state of UAV  $i$ , which includes attributes such as the UAV's position ( $p_i$ ), sensor data ( $sd_i$ ), communication status ( $cs_i$ ), and task-specific cues ( $tc_i$ ). Mathematically, the input to the neural network's input layer can be represented as:

$$Input_i = [p_i, sd_i, cs_i, tc_i].$$

*Hidden Layers.*

The neural network consists of multiple hidden layers that process the input data, extracting features and patterns that influence decision-making [7]. Each hidden layer contains neurons interconnected through weighted connections. The activation of a neuron  $j$  in layer  $l$  ( $a_j^{(l)}$ ) is computed using the weighted sum of inputs ( $z_j^{(l)}$ ) and an activation function ( $\sigma$ ). For a given neuron  $j$  in hidden layer  $l$ , the calculation is:

$$z_j^{(l)} = \sum_{i=1}^{n^{(l-1)}} \omega_{ji}^{(l)} * a_i^{(l-1)} + b_j^{(l)},$$

where  $a_j^{(l)} = \sigma(z_j^{(l)})$ ,

$\omega_{ji}^{(l)}$  – weight parameter of a given layer,

$b_j^{(l)}$  – bias parameter for the given hidden layer [8].

*Forward Propagation.*

Moving forward the system adjusts the weights of the network's connections based on the collected rewards ( $r_i$ ) and experiences, facilitating learning. The adjustment process, guided by reinforcement learning algorithms, aims to maximize cumulative rewards over time. The reward is based on the number of UAVs that are near the searchable object and is decremented based on the UAVs that are not near any object. The weight update rule for a connection between neuron  $j$  in layer  $l$  and neuron  $k$  in layer  $l+1$  can be expressed as:

$$\Delta\omega_{jk}^{(l)} = \alpha * a_j^{(l)} * \delta_k^{(l+1)},$$

where  $\delta_k^{(l+1)} = \frac{\partial L}{\partial z_k^{(l+1)}}$ ,

$L$  – loss function,

$z_k^{(l+1)}$  – intermediate neuron value, on neuron  $k$  in layer  $(l+1)$ . It is calculated from the weight, bias and neuron value of the previous layer.

The weight and bias parameters are updated by subtracting the partial derivation of the loss function with respect to those parameters [9].

*Output Layer:*

The final layer of the neural network produces the output action for the swarm. This action guides the collaborative behavior of the UAVs, directing them toward the most beneficial actions within the environment. Let  $a_i$  represent the output action for UAV  $i$ , which influences the swarm's collective behavior.

The neural network evolves through training episodes as the swarm engages in tasks. The adaptation of its decision-making strategy occurs as the network learns from experiences and refines its approach, culminating in optimal performance based on the task's requirements.

### Insights from Results

The neural network-fueled swarm aptly distributes different UAV types based on the task's demands. This dynamic allocation enhances the swarm's effectiveness, facilitating efficient task completion in diverse scenarios.

The neural network-orchestrated swarm demonstrates heightened efficiency in task execution. Its ability to make intelligent decisions on optimal UAV selection curtails redundancy and resource wastage, thus amplifying overall efficiency.

The results of the experiment showed that the best solution for the given scenario was the one with 2 slow types of agents and 3 faster ones. To prove the concept a new experiment was conducted, where all the different possible scenarios were tested with the combination of 5 drones consisting of different types. The results of the reward are in the picture below (Fig. 1) proving the best solution with 2 slower agents (called planes in the picture).

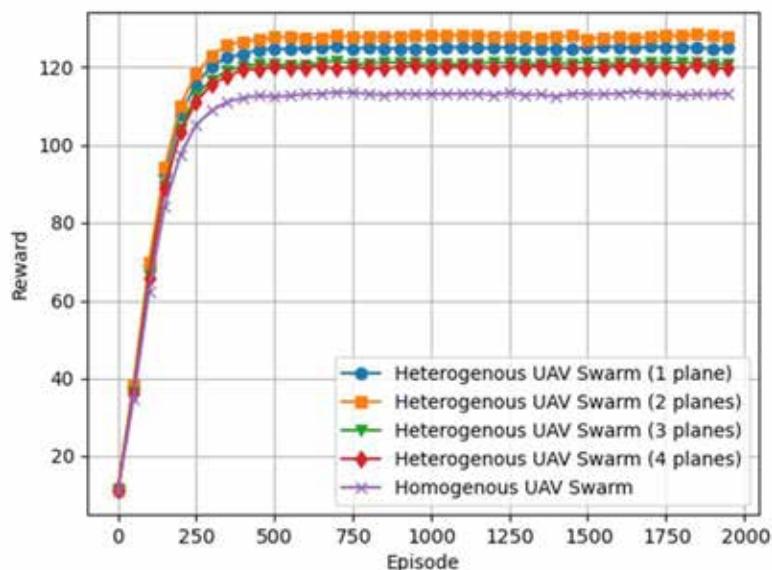


Fig. 1. Rewards of the system for different setup of the swarm

### Conclusion

The dynamic environment of UAV swarms evolves further as neural networks and reinforcement learning converge. The outcomes of the experiment reverberate with practical implications, as it gives the proof of concept of a system that would be able to choose the best approach of existing UAVs to maintain the fastest and most efficient result. The fusion of neural networks and reinforcement learning to optimize heterogeneous UAV swarms underscores the essential role of AI techniques in fully exploiting the potential of collaborative aerial systems. As industries embrace swarm-based technologies, the ability to dynamically allocate resources within the swarm broadens horizons for efficiency gains and refined task accomplishment.

Upcoming research could explore scenarios such as dynamic environments, varied task distributions, and real-time decision-making complexities. Exploring mechanisms to integrate external data sources into the decision-making process could further augment the swarm's capabilities. Adding the more advanced physics of the flight with some dynamic changes of the environment would be the way to better imitate the real world, getting the experiment to the new level. Other way of advancing the experiment could be creation of an algorithm of adding new types of UAVs to the list during the experiment and making possible to add additional UAVs to the swarm in the middle of the process.

### References

1. Yongkun Zhou, Bin Rao, Wei Wang. (2020) UAV Swarm Intelligence: Recent Advances and Future Trends. *School of Electronics and Communication Engineering*. DOI: 10.1109/ACCESS.2020.3028865
2. Hanno Hildmann, Ernő Kovacs, Fabrice Saffre, A. F. Isakovic. (2019) Nature-Inspired Drone Swarming for Real-Time Aerial Data-Collection Under Dynamic Operational Constraints. *Drones*.
3. Xiaofeng Hong, Yonghui Zhao, Nasreen Kausar, Ardashir Mohammadzadeh, Dragan Pamucar, Nasr Al Din Idecor. (2022) A New Decision-Making GMDH Neural Network: Effective for Limited and Fuzzy Data. *Computational Intelligence and Neuroscience*.

4. Maryam Kouzeghar, Youngbin Song, Malika Meghjani, Roland Bouffanais. (2023) Multi-Target Pursuit by a Decentralized Heterogeneous UAV Swarm using Deep Multi-Agent Reinforcement Learning. *ICRA*. DOI:10.1109/ICRA48891.2023.10160919
5. Ashish Rana. (2018) Reinforcement Learning with OpenAI Gym. *Towards Data Science*.
6. Ziquan Wang, Juan Li, Jie Li, Chang Liu. (2023) A decentralized decision-making algorithm of UAV swarm with information fusion strategy. *Expert Systems with Applications*.
7. Jide Nosakare Ogunbo, Olufemi Adigun Alagbe, Michael Ilesanmi Oladapo, Changsoo Shin. (2020) N-hidden layer artificial neural network architecture computer code: geophysical application example. *Computer Science*.
8. Yesmina Jaafraa, Jean Luc Laurent, Aline Deruyvera, Mohamed Saber Naceur. (2019) Reinforcement Learning for Neural Architecture Search: A Review. *Elsevier*.
9. Bhavika (2019) Mathematics behind the Neural Network. *Machine Learning Model*.