ISSN 2522-9818 (print)
ISSN 2524-2296 (online) Innovative technologies and scientific solutions for industries. 2025. No. 4 (34)

UDC 629.735.3 DOI: https://doi.org/10.30837/2522-9818.2025.4.032

M. YENA, O. POHUDINA

INTEGRATED SIMULATION MODEL OF Swarm CONTROL AND ADAPTIVE ROUTEING OF UAVS IN A CHANGING AIR ENVIRONMENT

Subject matter: the processes of swarm control and adaptive routing of unmanned aerial vehicles (UAVs) in complex and dynamically changing air conditions using adaptive algorithms. **Goal**: to develop an integrated simulation model that combines swarm control methods, adaptive PID control and adaptive routing algorithms to ensure the safety, optimality and efficiency of UAV fleet movement in conditions of a changing air environment. **Tasks**: to analyze existing approaches to swarm control and adaptive routing of UAVs; to develop a mathematical model of an integrated system that takes into account the specifics of interaction between UAVs, collision avoidance and dynamic changes in the air environment; to create a swarm control algorithm based on adaptive PID regulation of UAV movement parameters; to develop and implement an adaptive routing algorithm that responds to changes in traffic, weather conditions and other airspace factors; to implement the integrated model in a simulation environment and test its effectiveness; to conduct a comparative analysis of the efficiency of UAV operation with and without the developed algorithms. **Methods**: use of adaptive PID control methods for dynamic regulation of UAV movement trajectories and ensuring flight accuracy and stability; application of swarm control algorithms (boids-type methods) for synchronization of movement and collision avoidance in UAV groups; nonlinear optimization of routes taking into account dynamically changing conditions, which allows minimizing collision risks, energy consumption and flight time; construction of a graph-theoretic model of airspace for effective route planning and situation forecasting; creation of digital twins of the air environment for conducting simulation experiments. **Results**: an integrated simulation model of swarm control and adaptive routing of UAVs was developed, which takes into account air environment variables; adaptive PID control and swarm control algorithms ensured a reduction in the average positioning error and collision avoidance of UAVs; According to the results of simulation experiments, an increase in the reward of agents by ≈50%, an increase in the successful completion of episodes by ≈50%, and a reduction in agent errors on the way to the goal by ≈10%. **Conclusions**: created integrated model allows for effective management of UAV flotillas in conditions of a changing air environment, significantly increasing the safety and optimality of routes; the use of adaptive algorithms and graph-theoretic models provides high forecasting accuracy and risk minimization; the results of the study confirm the prospects for implementing the developed algorithms for UAV control in urban and regional conditions.

Keywords: unmanned aerial vehicles; swarm control; adaptive routing; PID control; real-time data processing; route optimization.

Introduction

The rapid development of unmanned aerial vehicles (UAVs) and the growth in their use in civil and specialized areas necessitate the creation of effective air traffic management methods. This issue is particularly relevant in urban and complex dynamic environments, where the number of UAVs is rapidly increasing and airspace is characterized by high density, changing weather conditions, and limited resources. Effective management of UAV fleets requires new approaches that take into account group behavior, dynamic response to changes in the environment, and the ability to self-learn [1] (hereinafter, the term "agent" refers to a single unmanned aerial vehicle in the model; these terms are used synonymously in this paper).

There are currently numerous approaches to optimizing UAV routes and coordinating their movement: the use of theoretical graph models, nonlinear optimization methods, evolutionary algorithms, collision avoidance strategies, etc. [2]. However, most of them are either focused on individual trajectory planning or do not take into account the peculiarities of collective fleet dynamics and adaptation to changes in airspace in real time. This significantly limits their application for automated control systems of a large number of UAVs in real operating conditions.

To solve the outlined problems, the paper proposes an integrated simulation model that combines swarm intelligence, adaptive PID control, and dynamic routing algorithms. The developed model makes it possible to simulate the complex collective behavior of a group of UAVs, ensure coordinated avoidance of obstacles and collisions, and dynamically update routes in response to changes in air and weather conditions. The proposed approach is universal for various types of air missions,

ISSN 2522-9818 (print)
Сучасний стан наукових досліджень та технологій в промисловості. 2025. № 4 (34) ISSN 2524-2296 (online)

from monitoring and search and rescue operations to logistics in urban environments.

Research objectives and tasks

The aim of the work is to develop an integrated simulation model of swarm control and adaptive routing for a group of unmanned aerial vehicles, which ensures safe, optimal, and efficient movement of fleets in a changing air environment.

The main criteria for the effectiveness of the model are:

– minimization of the average distance to the target for each agent (UAV);

– avoidance of collisions;

– ability to dynamically adapt to changes in airspace;

– stability and control even in complex and turbulent conditions.

To achieve the goal, the following tasks must be performed:

– analyze modern approaches to swarm control and UAV routing;

– formalize a mathematical model of group movement, taking into account the interaction of agents (UAVs) and dynamic changes in the environment;

– develop and investigate algorithms for collective behavior and adaptive trajectory control;

– implement and test the model in a simulation environment, compare its effectiveness with basic approaches.

Analysis of recent research and publications

Over the past decade, the issue of UAV Traffic Management (UTM) has been attracting increasing attention from the scientific community due to the prospects for the widespread use of unmanned aerial vehicles in urban and regional airspace. Researchers are focusing on developing models for decentralized fleet management [3], route optimization, collision avoidance, and the integration of artificial intelligence into automated systems [4].

Among the most cited works in recent years are those on modeling collective behavior (*swarm intelligence*) for UAVs, which implement the *Boids* approach and its modifications [5, 6]. Such models make it possible to implement coordinated control of the movement of a group of devices, adapt to dynamic obstacles, ensure self-coordination, and avoid collisions.

Special attention is paid to the use of adaptive controllers in UAV trajectory control tasks. Modern research actively implements PID and LQR controllers [7], which are capable of automatically adapting their parameters to changes in the dynamics of the environment [8]. However, most of the work focuses on controlling single UAVs or small groups, while often ignoring the effects of mass interaction and a large number of agents in real airspace.

In the field of dynamic routing, graph models of airspace have become widespread, allowing for effective trajectory planning based on space topology, traffic, and weather conditions [9]. The use of nonlinear optimization methods, evolutionary algorithms (in particular, genetic algorithms, particle swarm algorithms [10], etc.) and hybrid approaches helps to find optimal routes in multifactorial scenarios.

Recently, there have been studies on the use of reinforcement learning (RL) for routing and distributed control of UAVs [11, 12]. RL approaches allow agents to be trained to interact effectively with an unknown environment, but the issues of stability and interpretability of such systems remain open.

A review of the literature shows that, despite significant progress in the development of theory and tools for UAV traffic control, there is still a lack of comprehensive models that combine adaptive routing, swarm control, and the ability to integrate real-time environmental variables. Accordingly, the development and research of integrated simulation models capable of combining these components is a relevant scientific task.

Mathematical model of swarm control and adaptive routing systems

Formalization of the task of controlling groups of UAVs

This section presents a formalization of the problem of collective control of a group of UAVs in complex airspace. The space in which the devices move is modeled as a directed graph $G=(V,E)$, where the set of nodes V corresponds to key geographical or control points (e.g., launch zones, target positions, or obstacle avoidance points), and the edges E define the permissible trajectories of movement between these points [13]. This approach makes it possible to naturally take into account the structure of the environment, the

ISSN 2522-9818 (print)
ISSN 2524-2296 (online) Innovative technologies and scientific solutions for industries. 2025. No. 4 (34)

presence of restrictions, and variable movement conditions. Each agent in this model, understood as a separate UAV, is defined by three basic parameters: spatial position, velocity vector, and a set of adaptive PID controller coefficients. Formally, the state of the *i*-th agent at time step *t* is defined as:

– $x_i(t)$ – current position in the corresponding coordinate system;

– $v_i(t)$ – speed of movement;

– $p_i(t)=[K_{p,i},K_{i,i},K_{d,i}]$ – vector of PID parameters, which are adaptively adjusted for each UAV taking into account changes in the environment.

The goal of control is to minimize the average distance to target points for all agents (UAVs) while avoiding collisions and ensuring efficient routing in complex and dynamic conditions. This approach allows not only to coordinate group movement, but also to take into account the individual characteristics of the trajectories and behavior of each device.

The dynamics of agent movement are determined by a system of differential equations that describe changes in position and velocity, taking into account the actions of control signals and random disturbances (e.g., noise or turbulence). This makes it possible to model the behavior of UAVs in realistic conditions:

$$x_i(t+1)=x_i(t)+u_{i,x}(t)dt+\eta_{i,x}(t),$$
$$y_i(t+1)=y_i(t)+u_{i,y}(t)dt+\eta_{i,y}(t), \quad (1)$$

where $u_{i,x}$, $u_{i,y}$ – resulting control actions (determined by the control system); dt – time discretization step; $\eta_{i,*}$ – random component (noise, turbulence).

Swarm Control

Collective behavior is achieved using *swarm* control, which is based on a modified *Boids* algorithm. It takes into account three main components:

– ***cohesion*** **–** inclination towards the center of mass of neighbors;

– ***separation*** – avoidance of collisions when approaching other agents;

– ***alignment*** – aligning the direction of movement with the nearest neighbors.

The summary control action of the *swarm* component is determined as follows:

$$u_{swarm}(t)=\alpha alignment_i\times\beta cohesion_i\times\gamma separation_i, \quad (2)$$

where α,β,γ – are selected or optimized during simulations, ensuring a balance between group cohesion and individual safety.

Adaptive PID control

To accurately follow the specified trajectories, each UAV is equipped with an adaptive [14] PID controller, which takes into account not only the current error, but also its accumulation and rate of change. The control action of the PID is described by the equation:

$$u(t)=K_p(t)e(t)+K_i(t)\int e(t)dt+K_d(t)\frac{de(t)}{dt}, \quad (3)$$

wher $e(t)$ – the current distance between the agent and the specified target, and the derivative $de(t)/dt$ reproduces the change in this distance over time; $K_p(t)$ – proportional coefficient of the PID controller at time *t*, responsible for responding to the current error; $K_i(t)$ – integral coefficient, compensates for systematic error (accumulated error over time); $K_d(t)$ – differential coefficient, responsible for anticipating and responding to the rate of error change.

These parameters can be adapted, in particular, using *reinforcement learning*, which helps to increase resistance to external influences, turbulence, or changes in the environment.

The proposed formalization makes it possible to model the behavior of a group of UAVs in both calm and changing airspace, ensuring the flexibility, scalability, and adaptability of the control system for solving complex tasks in modern aeronautics.

Routing model and route updates

In a dynamic airspace with variable factors such as traffic, weather conditions, and the emergence of danger zones, it is critical for the system to be able to quickly update UAV routes. To do this, a graph model is used, in which the vertices represent key points in the airspace and the edges represent possible trajectories. At each time step, the optimal route to the target is determined for each agent (UAV) based on current information about the state of the environment. Route optimization consists of finding a path for which the total "cost" of travel is minimal:

$$\pi_i^* \arg\min_{\pi\in P}(\sum e\in\pi\omega_e(t)), \quad (4)$$

ISSN 2522-9818 (print)
Сучасний стан наукових досліджень та технологій в промисловості. 2025. № 4 (34) ISSN 2524-2296 (online)

where $w_e(t)$ – current "cost" (delay, risk) of the edge e; π – possible route in the set of all permissible routes P.

The "cost" of an edge is determined comprehensively: traffic levels in the area, weather threats, flight restrictions, or *no-fly* zones are all taken into account. With each change in the environment (e.g., the appearance of a new obstacle or a change in weather conditions), the routing model recalculates the optimal path for each agent in real time.

This approach ensures the flexibility of the system, allowing agents to quickly adapt to changing conditions and minimize risks during missions.

Reward function and performance criteria

One of the key elements of the control system is a properly designed reward function that guides agent learning and influences their behavior strategy [15, 16]. The main goal of the reward function is to encourage the agent to reduce the average positional error, avoid collisions [17], effectively achieve goals, and minimize resource expenditure on the route.

Within the current approach, the reward at each step of the simulation is determined by the following formula:

$$r_t = w_\delta(\varepsilon_{t-1} - \varepsilon_t) - w_{step} + r_{success} \times I[\varepsilon_t - \varepsilon_{th}] + r_{collision} I[condition], \quad (5)$$

where ε_t – mean average distance to target (MAE) per step *t*; w_δ, w_{step} – weight coefficients for progress and penalties per step; $r_{success}$ – additional reward for achieving the goal; $I[condition]$ – indicator equal to 1 if the condition is met, 0 otherwise.

The parameters of the reward function were selected experimentally to achieve a balance between speed and safety. In particular, an increase of w_δ stimulates rapid convergence with the target, but excessively high values can increase the risk of collisions. The penalty w_{step} limits the time to complete the task, while the bonus $r_{success}$ motivates the agent to complete the mission rather than move in endless loops near the target. Typical values of the coefficients used in the experiments: $w_\delta = 400.0$; $w_{step} = 0.2$; $r_{success} = 500.0$.

Analysis of the simulation results showed that this reward function setting not only effectively reduces MAE, but also ensures a high *success rate* (percentage of goals achieved) at an acceptable level of safety.

Next, we will look how the described models and criteria are integrated into a single management system and what the architecture of the simulation model looks like.

Adaptive *Swarm*+PID controller: algorithm

To ensure both coordinated collective behavior and individual adaptation of each UAV in a variable airspace, the proposed model uses a combined controller. It combines the advantages of *swarm* approaches with classic PID control and also includes reinforcement learning mechanisms.

At each step of the simulation, a total control signal is formed for each agent *i*:

$$u_i(t) = u_{swarm,i}(t) + u_{pid,i}(t), \quad (6)$$

where $u_{swarm,i}(t)$ – collective behaviour vector (*alignment, cohesion, separation*);

$u_{pid,i}(t) = K_P e_i(t) + K_i \sum_{r=0}^{t} e_i(r) + K_d \Delta e_i(t)$ – PID-regulation with dynamic adaptation of coefficients.

The controller's adaptability is ensured by dynamic adjustment of PID parameters based on feedback, which allows the system to respond flexibly to changes in turbulence, the appearance of new obstacles, and other external influences.

The logic of the combined controller (algorithm)

The sequence of actions for each simulation step involves several stages.

1. **Obtaining the state.** The RL agent analyzes the current state of the environment (*St*), which contains the positions, speeds, and locations of targets, as well as information about obstacles.

2. **Generation of strategic action.** The RL agent determines the optimal actions for the swarm module, i.e., generates signals for the collective behavior of agents.

3. **Collective coordination.** The swarm module coordinates the interaction between agents, ensures collision avoidance, and maintains group integrity.

4. **Individual adjustment.** For each agent, a PID correction is calculated based on the current position error, after which a total action is formed.

ISSN 2522-9818 (print)
ISSN 2524-2296 (online) Innovative technologies and scientific solutions for industries. 2025. No. 4 (34)

5. **Status and reward updates.** The system updates the position of agents in space, calculates the reward function, and proceeds to the next step of the simulation.

This architecture ensures high flexibility and adaptability of the control system, allowing the UAV to operate effectively even in complex and dynamic conditions.

The next important stage of the research was the design and implementation of the simulation model architecture, which combines all the components considered into a single system.

Simulation model architecture

To implement and test the proposed control algorithms, a modular simulation architecture was created that combines all key components of the system into a single integrated environment. This approach provides flexibility in conducting research, facilitates scaling, and allows for in-depth analysis of the interaction of various control elements in a variety of experimental scenarios.

The simulation model contains several specialized modules, each of which plays a specific role in ensuring the effective operation of the entire system. In particular, the **swarm control module** (*Swarm Controller*) is responsible for coordinating the collective behavior of a group of UAVs. Based on a modified *Boids* algorithm [18], it dynamically adjusts the speed and trajectories of agents, promoting coordinated movement and avoiding collisions.

In addition, an adaptive PID controller allows each UAV to individually change control parameters according to the current positioning error, environmental changes, and accumulated experience. This mechanism increases the stability and adaptability of the system to various influences, including turbulence or the appearance of new obstacles.

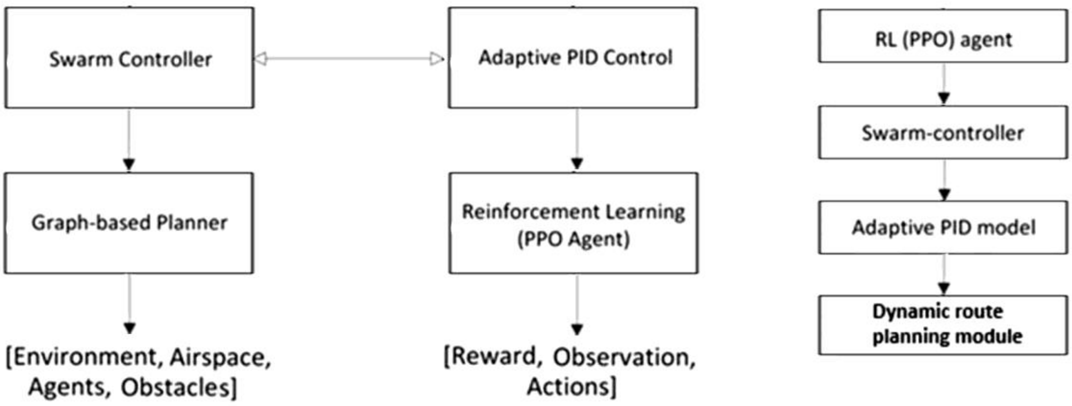
The graph router dynamically updates the optimal routes for each agent, taking into account the current configuration of the airspace, traffic, weather conditions, and possible movement restrictions. This helps the system to respond quickly to changes and optimize trajectories in real time.

An important component of the architecture is the *reinforcement learning agent* (RLA), which acts as the "brain" of the UAV group. This agent analyzes the global state of the system and generates strategic actions for the entire group, constantly improving the policy based on feedback and collective experience.

Interaction between modules is implemented through a main control loop, in which the RL agent generates strategic signals that are sent to the swarm controller. The latter coordinates local actions between agents, after which each UAV individually adjusts its trajectory using a PID controller. At each step, information about the current state of the environment, such as positions, obstacles, turbulence, or other changes in conditions, is sent to all modules. This ensures their synchronization, adaptation, and the integrity of the entire system's functioning.

This modular architecture not only improves the efficiency of group control, but also makes it easy to investigate the impact of individual components or settings on the overall behavior of the fleet in various simulation scenarios.

The model structure is shown in Fig. 1.



A B

Fig. 1. (A) – basic module architecture;
(B) – diagram of interaction between the RL agent, swarm controller, and PID controller in the task of routing a group of UAVs

ISSN 2522-9818 (print)
Сучасний стан наукових досліджень та технологій в промисловості. 2025. № 4 (34) ISSN 2524-2296 (online)

This approach allows for the effective integration of modern artificial intelligence algorithms with classical control methods, which significantly expands the possibilities for modeling and practical application of the system.

The model is configured by selecting hyperparameters for conducting a series of simulation experiments.

Key hyperparameters (PPO, PID, *Swarm*)

When configuring the combined *RL+Swarm+PID* system, special attention was paid to the selection of hyperparameters that determine the quality of training, noise resistance, adaptability, and the effectiveness of UAV group coordination.

In particular, for a reinforcement learning agent using the Proximal Policy Optimization (PPO) algorithm, it was important to select parameters such as **learning_rate, ent_coef, clip_range, net_arch, and total_timesteps**. For example, the selected learning rate value (***learning_rate*** **= 0.0003**) helped to find the optimal balance between the speed of policy updates and its stability: increasing this parameter led to instability, while decreasing it led to excessively slow convergence.

The entropy coefficient (***ent_coef*** **= 0.05**) determines the agent's activity level: increasing this parameter above 0.1 made the policy too chaotic, while decreasing it slowed down adaptation to new conditions. The parameter ***clip_range*** **= 0.25** ensured smooth policy updates, and the selected neural network architecture (***net_arch*** **= [128, 128]**) made it possible to effectively approximate complex multi-agent strategies.

The PID controller parameters were adjusted according to the need to compromise between quick response to changes and control stability in complex or noisy environments. Additional tests showed that overly aggressive parameters negatively affect the result in turbulent conditions.

The *Swarm* module was modeled using weight coefficients (***alpha, beta, gamma***) that determined the strength of alignment, attraction to the center of the group, and collision avoidance. The introduction of Gaussian noise (***sigma_turbulence*** **= 0.05**) made it possible to evaluate the robustness of the controller in turbulent conditions.

All these parameters were selected empirically based on numerical experiments with cross-validation. The detailed influence of each parameter on the simulation results is summarized in Table 1, which facilitates a reasonable choice of configuration for different tasks. This allows the system configuration to be selected depending on the tasks and experiment scenarios. It was this combination of hyperparameters that provided the best results in terms of MAE, *success rate*, and convergence speed for different experimental scenarios.

Table 1. *Hyperparameters and their impact*

Parameter	Mean	Impact
learning_rate	0.0003	Speed of agent adaptation
ent_coef	0.05	Level of exploration
clip_range	0.25	Stability of learning
net_arch	[128, 128]	Quality of policy approximation
Kp, Ki, Kd	1.0/0.1/0.01	Stability and speed of PID control
sigma_turb	0.05	Realism of the model (noise testing)
total_timesteps	100_000	Duration of learning for each series of experiments

Training was performed using a train-test scheme with random initial positions to avoid overfitting. Each model was tested on 30 independent episodes. Stability was checked using cross-validation on different seeds.

The choice of parameters was justified by a series of tests that showed that increasing *ent_coef* to 0.1 reduces stability, and smaller networks ([64, 64]) worsen the convergence rate.

The entropy coefficient allows the agent to explore trajectories more actively in complex situations. Increasing the number of neural network layers contributes to better policy approximation in multi-agent cases.

Success criterion: an episode is considered successful if MAE < 0.7 in 40 steps. The MAE threshold value was chosen as a compromise between positioning accuracy and time/resource constraints and was confirmed by experiments.

The following experiments are based on the above-selected system hyperparameters.

Research sequence

The effectiveness of the proposed system was evaluated through a series of numerical experiments implemented in a specialized simulation environment developed on the basis of the ***Gymnasium*** platform. The research was conducted according to a carefully designed scenario that included several interrelated stages.

The first stage involved **preparing the environment**, during which airspace was formed

ISSN 2522-9818 (print)
ISSN 2524-2296 (online) Innovative technologies and scientific solutions for industries. 2025. No. 4 (34)

in the form of an oriented graph with specified nodes (start, targets, obstacle zones) and edges denoting possible movement routes. To increase the realism of the simulation, random obstacles, high-risk areas, and so-called no-fly zones were added. This configuration made it possible to test the system's ability to adapt to complex and unpredictable scenarios.

Next, **the agents were initialized**: a group of UAVs was placed in random or predetermined starting positions. Each agent received individual controller parameters, as well as initial speed and orientation values. This ensured diversity of behavior within the group and made it possible to evaluate the model's ability to coordinate flexibly.

The simulation consisted of sequential episodes, in each of which the agents moved according to a combined *RL+Swarm+PID* algorithm. At each step of the trajectory, the trajectories were dynamically updated to account for changes in the environment, interactions between agents, and the presence of turbulence or obstacles. Particular attention was paid to testing the system's behavior under stress conditions, such as the appearance of new obstacles or increased noise.

During the simulation process, **key metrics were collected**: mean absolute error (MAE), success rate, number of collisions, average number of steps to reach the goal, and other performance indicators. Analysis of these indicators provided an objective picture of the system's performance.

To substantiate the advantages of the proposed architecture, the results were **compared with basic approaches**, in particular with classic PID control without swarm control and reinforcement learning elements. This comparative analysis helped to identify specific advantages of the new model in terms of robustness, flexibility, and convergence speed.

A separate stage was **repeating the experiments for different scenarios**: the number of agents (from 1 to 10) was changed, the noise level, the number and topology of obstacles were varied. To avoid overfitting, a *train-test* scheme with random initial conditions and cross-validation was used, which ensured the reliability of the results.

This comprehensive approach made it possible to comprehensively assess the stability and adaptability of the developed control system in various operating conditions, as well as to determine the influence of individual parameters on the overall efficiency of the model's functioning.

At each stage of the experiments, the influence of the key hyperparameters of the RL agent, PID controller, and *swarm* module on the main performance metrics was analyzed separately: average absolute positioning error, frequency of successful episodes, and system robustness to external disturbances. The results of comparing different configurations made it possible to identify the optimal settings for achieving stable model operation in various airspace scenarios.

Experiment results

To evaluate the effectiveness of the proposed approach, numerical simulations were performed in a self-developed environment based on the ***Gymnasium*** platform. The experiment scenarios included both single UAV flights and group interactions of up to 10 agents in environments with varying amounts of obstacles and noise levels. All models were trained for 100,000 steps using pre-selected hyperparameters (see Table 1), which ensured balanced learning dynamics and resistance to noisy conditions.

The criterion for the success of the experiment was to achieve a mean absolute error (MAE) of less than 0.7 over 40 steps. This threshold value was chosen based on previous studies as one that guarantees approach to the target with minimal collision risk and meets realistic requirements for positioning accuracy in multi-component control systems.

Each model underwent a series of 30 independent episodes, which made it possible to evaluate the statistical stability, robustness, and reproducibility of the results. Two approaches were tested to compare their effectiveness: RL+Swarm (a combination of deep reinforcement learning Proximal Policy Optimization with swarm control) and a classic PID controller with fixed coefficients. The evaluation was based on the average MAE, the proportion of successful episodes, the convergence rate, and adaptability to changing environmental conditions.

Key evaluation metrics

A comparison of the total rewards of *RL+Swarm* and *Baseline PID* is shown in Fig. 2.

As can be seen from the graph, *RL+Swarm* demonstrates significantly higher rewards for most episodes.

The MAE values for each episode are shown in Fig. 3.

Сучасний стан наукових досліджень та технологій в промисловості. 2025. № 4 (34)
ISSN 2522-9818 (print)
ISSN 2524-2296 (online)

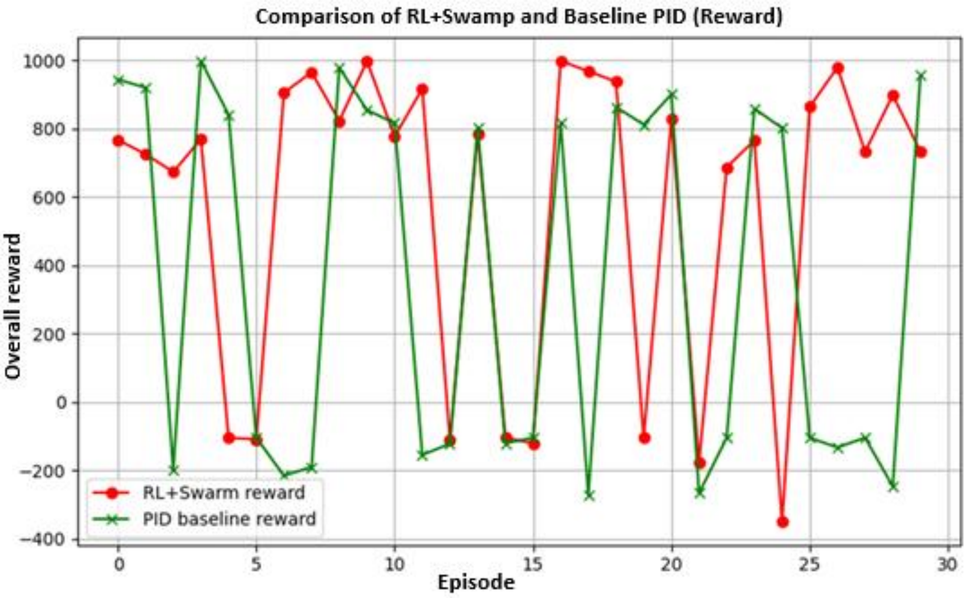


Fig. 2. Reward comparison between algorithms

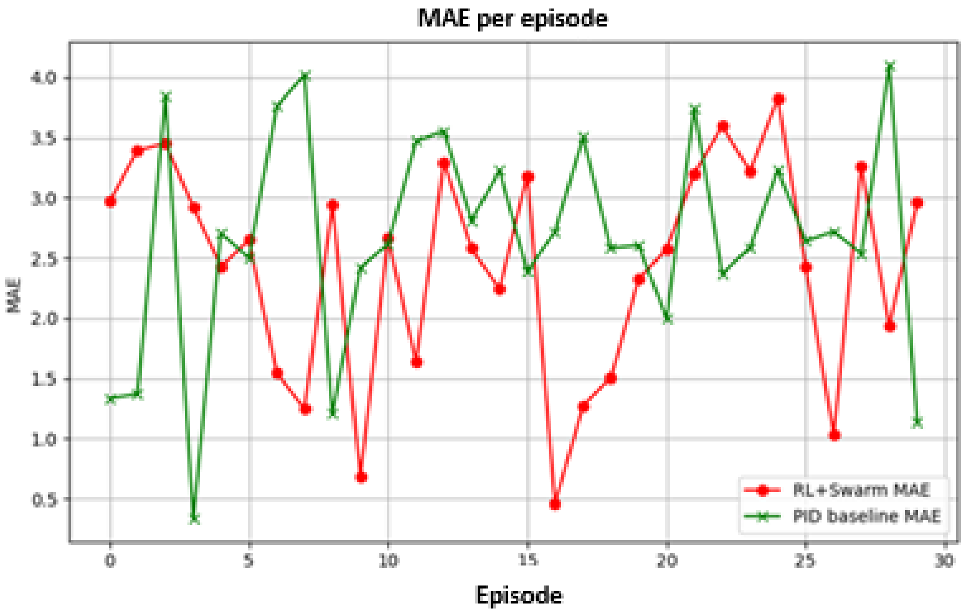


Fig. 3. Comparison of MAE distribution per episode for both algorithms

The value of *Mean RL MAE* is 2.4, and *Mean PID MAE* is 2.67. This means that the number of errors made by agents on the way to the target for the RL algorithm has decreased.

The percentage of successful episodes is shown in Fig. 4.

The goal achievement rate for RL is 73.3%, and for **PID** = 50%. This indicates that the proposed algorithm achieves the goal approximately 50% more often.

The number of steps to the goal is shown in Fig. 5.

The analysis shows that *RL+Swarm* demonstrates better adaptation to dynamic changes and learns more effectively in scenarios with a changing environment. At the same time, *PID baseline* reaches the goal faster in terms of the number of steps, but has a lower success rate.

Fig. 6 shows a comparison of the movement of an agent controlled by *RL+Swarm* (red line, Fig. 6, a) and a classic PID controller (green line, Fig. 6, b) during the first episode of the experiment, where the target is marked as *Goal*.

As can be seen in Fig. 6, *RL+Swarm* demonstrates a more straightforward and efficient path to the target, while the PID controller deviates more often from the shortest trajectory.

ISSN 2522-9818 (print)
ISSN 2524-2296 (online)
Innovative technologies and scientific solutions for industries. 2025. No. 4 (34)

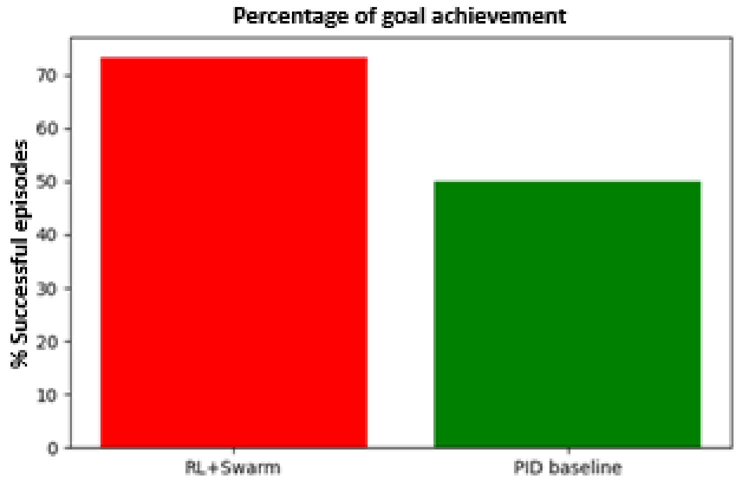


Fig. 4. Percentage of goals achieved

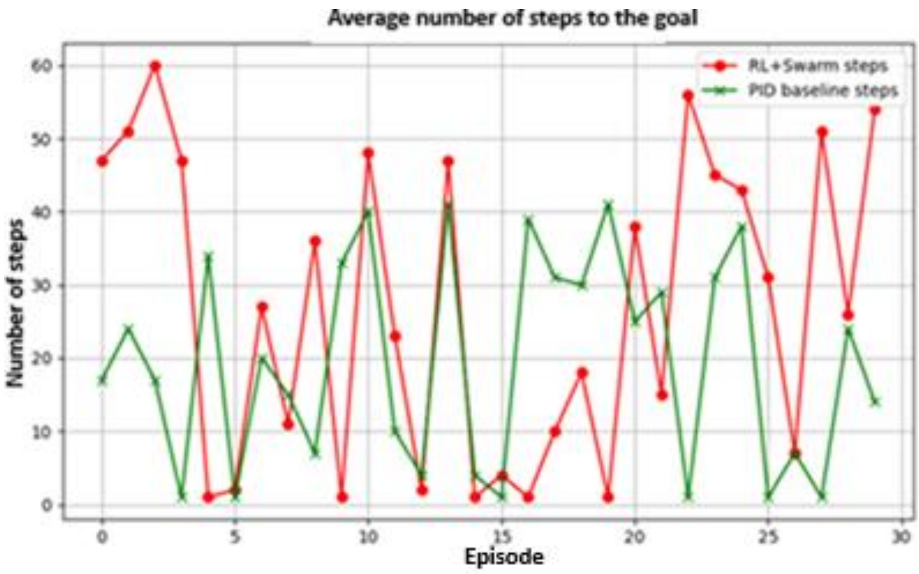


Fig. 5. Average number of steps for each UAV for two algorithms

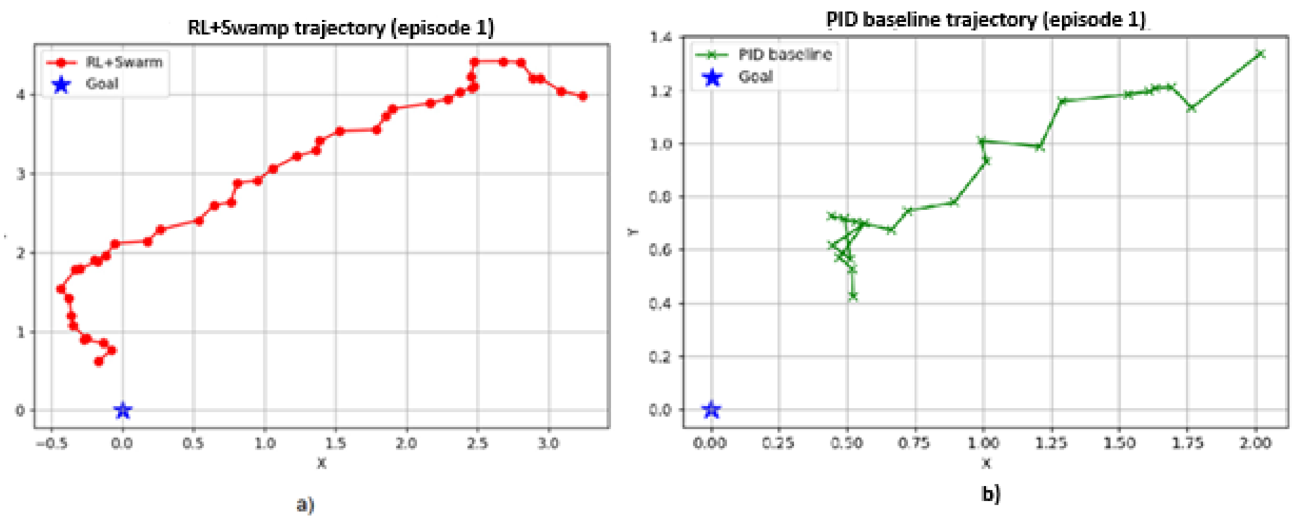


Fig. 6. Typical trajectories of a single agent in an *XY* environment

ISSN 2522-9818 (print)
Сучасний стан наукових досліджень та технологій в промисловості. 2025. № 4 (34) ISSN 2524-2296 (online)

This analysis allows to visually assess the advantage of the trained *RL+Swarm* model over the classical approach, especially in complex or noisy environments.

Research results and their discussion

The results show that *RL+Swarm* demonstrates a significantly higher total reward compared to the PID controller (average 576.6 vs. 357.1). In *MAE* metrics, *RL+Swarm* also outperforms PID: the average MAE value was 2.45 vs. 2.67. The percentage of successful episodes for the RL agent reached 73.3%, while for PID it was only 50%. At the same time, the average number of steps to the target is lower for PID, which is associated with aggressive behavior, but at the same time, there are more unsuccessful episodes (the agent "flies" past the target or hits an obstacle).

The RL agent provides a significantly more stable approach to the target under difficult conditions (obstacles, noise), while the PID controller does not adapt to dynamic changes in the environment. In general, RL+Swarm takes better account of both individual and collective movement characteristics, adapts to dynamic changes and complex topologies, although it sometimes requires more steps to reach the target.

The PID controller is easier to implement but less resistant to obstacles and gets stuck in local minima. Analysis of unsuccessful episodes showed that the shortcomings of *RL+Swarm* are related to strong noise, aggressive starting positions, or insufficient training.

Conclusions

The paper proposes and investigates an integrated simulation model of swarm control and adaptive routing of a group of UAVs in a dynamic air environment. Numerous experiments confirmed the effectiveness of the proposed approach according to a number of metrics: average reward, MAE, and proportion of successful episodes.

Scientific novelty of the work

In this study, for the first time, a hybrid approach is proposed for collective control of a group of UAVs in a dynamic air environment, combining:

– adaptive PID control with parameter optimization through gradient or evolutionary mechanisms;

– *Swarm intelligence* (a collective behavior algorithm modified by Boids) for coordinating movement and avoiding collisions;

– Deep reinforcement learning (PPO) methods used for dynamic real-time adjustment of control strategies.

For the first time:

– RL agent interaction with PID and swarm controller for UAV routing tasks has been implemented;

– A reward function has been introduced that takes into account progress towards the goal, penalties for steps, collisions, and goal achievement, allowing for a balance between speed and safety;

– A comparative analysis with classic PID control was performed, and performance metrics (MAE, reward, % of successful episodes, number of steps to the goal) were presented for representative experimental scenarios.

– The influence of various RL and PID hyperparameters, as well as external factors (noise, obstacles) on the result was investigated.

Comparison with analogues (shown in Table 2) demonstrates that the proposed approach provides:

– an increase in the proportion of successful episodes by 20–35% compared to the baseline PID;

– a reduction in the mean error (MAE) and more stable agent behavior in the event of changes in external conditions.

Table 2. *Comparison of the proposed algorithm with the classic PID*

Metrixes	*RL+Swarm*	*PID Baseline*	Result
Mean reward	576.6	357.1	RL advantage: ≈ +61%
Mean MAE	2.45	2.67	RL advantage: less error on the way to the goal
Success Rate	73.3 %	50.0 %	RL advantage: almost 50% more successful episodes
Steps to goal	26.8	19.4	RL agent takes more steps on average

As can be seen from Table 2, the proposed approach demonstrates a significant increase in performance indicators.

Limitations and prospects

The main limitations of the study are the need for large computing resources for training, as well as the potential complexity of scaling to larger groups of agents in real-world conditions. Further research

ISSN 2522-9818 (print)
ISSN 2524-2296 (online) Innovative technologies and scientific solutions for industries. 2025. No. 4 (34)

is planned to focus on optimizing the reward function, hybridization with other planning algorithms, testing on more complex topologies, and conducting field experiments.

Practical significance of the results achieved

Application of the proposed architecture for monitoring systems, logistics, search and rescue operations, as well as in urban mobility scenarios (*Urban Air Mobility*).

References

1. Debnath, D., Vanegas, F., Sandino, J., Hawary, A. F., Gonzalez, F. (2024), "A review of UAV path-planning algorithms and obstacle avoidance methods for remote sensing applications". *Remote Sensing*, № 16 (21), 4019 p. DOI: 10.3390/rs16214019
2. Martins, F. G., Coelho, M. A. N. (2000), "Application of feedforward artificial neural networks to improve process control of PID-based control algorithms". *Computers & Chemical Engineering*, № 24 (2–7). P. 853-858. DOI: https://doi.org/10.1016/S0098-1354(00)00339-2
3. Liu, X., Peng, Z.R., Zhang, L.Y. (2019), "Real-time UAV rerouting for traffic monitoring with decomposition based multi-objective optimization". *Journal of Intelligent & Robotic Systems*, № 94, P. 491–501. DOI: 10.1007/s10846-018-0806-8
4. Almeida, E. N., Campos, R., Ricardo, M. (2022), "Traffic-aware UAV placement using a generalizable deep reinforcement learning methodology". *2022 IEEE Symposium on Computers and Communications (ISCC)*, P. 1–6. DOI: 10.48550/arXiv.2203.08924
5. Madani, A., Engelbrecht, A. Ombuki-Berman, B., (2023), "Cooperative coevolutionary multi-guide particle swarm optimization algorithm for large-scale multi-objective optimization problems". *Swarm and Evolutionary Computation*. № 82. 101262 p. DOI: https://doi.org/10.1016/j.swevo.2023.101262
6. Luo, J., Tian, Y., Wang, Z. (2024), "Research on unmanned aerial vehicle path planning". *Drones*, № 8(2). 51 p. DOI: 10.3390/drones8020051
7. Li, C. Lian, J., (2007), "The Application of Immune Genetic Algorithm in PID Parameter Optimization for Level Control System". *Proceedings of the 2007 IEEE International Conference on Automation and Logistics (ICAL)*, Jinan, China, P. 2670–2674. DOI: https://doi.org/10.1109/ICAL.2007.4338670
8. Yang, F., Lu, Q., Li, R., Xu, Y., Yuan, W., Wu, X. (2023), "Real-time optimal path planning and fast autonomous flight for UAV in unknown environments". *IEEE*. DOI: 10.23919/CCC58697.2023.10240971
9. Li, Q., Li, R., Ji, K., Dai, W. (2015), "Kalman filter and its application". *IEEE*. DOI: 10.1109/ICINIS.2015.35
10. Hooshyar, M., Huang, Y.M. (2023), "Meta-heuristic algorithms in UAV path planning optimization: A systematic review (2018–2022)". *Drones*, № 7(12), 687 p. DOI: 10.3390/drones7120687
11. Li, H., Zhang, Z.-y. (2012), "The application of immune genetic algorithm in main steam temperature of PID control of BP network". *Physics Procedia*, № 25, P. 80-86. DOI: https://doi.org/10.1016/j.phpro.2012.02.013
12. Zhang, M., Liu, Y., Wang, Y., Li, F., Chen, L. (2023), "Real-time path planning algorithms for autonomous UAV". *IEEE*. DOI: 10.1109/CAC57257.2022.10054770
13. Kim, D. H. (2003), "Comparison of PID controller tuning of power plant using immune and genetic algorithms". *The 3rd International Workshop on Scientific Use of Submarine Cables and Related Technologies*, Lugano, Switzerland, P. 358–363. DOI: 10.1109/CIMSA.2003.1227222
14. Zhang, G. (2023), "6G enabled UAV traffic management models using deep learning algorithms". *Wireless Networks*, № 30, P. 6709–6719. DOI: 10.1007/s11276-023-03485-4
15. Bezkorovainyi, V., Binkovska, A., Noskov, V., Gopejenko, V., Kosenko, V. (2025), "Adaptation of logistics network structures in emergency situations", *Advanced Information Systems*, Vol. 9, No. 4, P. 39–50. DOI: https://doi.org/10.20998/2522-9052.2025.4.06
16. Kuchuk, N., Kashkevich, S., Radchenko, V., Andrusenko, Y. Kuchuk, H. (2024), "Applying edge computing in the execution IoT operative transactions", *Advanced Information Systems*. Vol. 8, No. 4, P. 49–59. DOI: https://doi.org/10.20998/2522-9052.2024.4.07
17. Yena, M. (2024), "Optimizing air traffic control: Innovative approaches to collision avoidance in UAV operations". *Integrated Computer Technologies in Mechanical Engineering – 2023 (ICTM 2023)*. P. 543–553. DOI: 10.1007/978-3-031-60549-9_41
18. Saadi, A. A., Soukane, A., Meraihi, Y., Benmessaoud Gabis, A., Mirjalili, S., Ramdane-Cherif, A. (2022), "UAV path planning using optimization approaches: a survey". *Archives of Computational Methods in Engineering*. № 29. P. 4233–4284. DOI: 10.1007/s11831-022-09742-7

Received (Надійшла) 24.06.2025
Accepted for publication (Прийнята до друку) 30.11.2025
Publication date (Дата публікації) 28.12.2025

ISSN 2522-9818 (print)
ISSN 2524-2296 (online)

Відомості про авторів / About the Authors

Yena Maksym – National Aerospace University "Kharkiv Aviation Institute", PhD Student, Department of "Information Technology Design", Kharkiv, Ukraine; e-mail: yenamaxim98@gmail.com; ORCID ID: https://orcid.org/0009-0006-0664-3244; Scopus ID: https://www.scopus.com/authid/detail.uri?authorId=59161031800

Pohudina Olha – PhD (Engineering Sciences), Associate Professor, National Aerospace University "Kharkiv Aviation Institute", Associate Professor at the Department of "Information Technology Design", Kharkiv, Ukraine; e-mail: olha.pohudina@poliba.it; ORCID ID: https://orcid.org/0000-0001-5689-2552; Scopus ID: https://www.scopus.com/authid/detail.uri?authorId=57204907264

Єна Максим Вікторович – Національний аерокосмічний університет "Харківський авіаційний інститут", аспірант кафедри інформаційних технологій проєктування, Харків, Україна.

Погудіна Ольга Костянтинівна – кандидат технічних наук, доцент, Національний аерокосмічний університет "Харківський авіаційний інститут", доцент кафедри інформаційних технологій проєктування, Харків, Україна.

ІНТЕГРОВАНА СИМУЛЯЦІЙНА МОДЕЛЬ РОЙОВОГО УПРАВЛІННЯ Й АДАПТИВНОЇ МАРШРУТИЗАЦІЇ БпЛА В УМОВАХ ЗМІННОГО ПОВІТРЯННОГО СЕРЕДОВИЩА

Предмет дослідження – процеси ройового управління й адаптивної маршрутизації безпілотних літальних апаратів (БпЛА) у складних і динамічно змінних повітряних умовах із використанням адаптивних алгоритмів. **Мета** – розробити інтегровану симуляційну модель, що поєднує методи ройового управління, адаптивного PID-контролю та алгоритмів адаптивної маршрутизації для забезпечення безпеки, оптимальності й ефективності руху флотилій БпЛА в умовах мінливого повітряного середовища. **Завдання**: проаналізувати наявні підходи до ройового управління та адаптивної маршрутизації БпЛА; розробити математичну модель інтегрованої системи, яка бере до уваги особливості взаємодії між БпЛА, уникнення зіткнень та динамічні зміни повітряного середовища; створити алгоритм ройового управління, оснований на адаптивному PID-регулюванні параметрів руху БпЛА; розробити та впровадити алгоритм адаптивної маршрутизації, що реагує на зміни трафіку, погодних умов та інших факторів повітряного простору; реалізувати інтегровану модель у симуляційному середовищі та протестувати її ефективність; проаналізувати й порівняти ефективність роботи БпЛА з розробленими алгоритмами та без них. **Методи**: упровадження методів адаптивного PID-контролю для динамічного регулювання траєкторій руху БпЛА та забезпечення точності й стабільності польотів; застосування алгоритмів ройового управління (методи типу *boids*) для синхронізації руху та уникнення зіткнень у групах БпЛА; нелінійна оптимізація маршрутів з огляду на динамічно змінні уми, що дає змогу мінімізувати ризики зіткнень, витрати енергії та час польоту; побудова теоретико-графової моделі повітряного простору для ефективного планування маршрутів і прогнозування ситуацій; створення цифрових двійників повітряного середовища для проведення симуляційних експериментів. **Результати**: розроблено інтегровану симуляційну модель ройового управління та адаптивної маршрутизації БпЛА, яка зважає на зміни повітряного середовища; алгоритми адаптивного PID-контролю та ройового управління забезпечили зменшення середньої похибки позиціювання та уникнення зіткнень БпЛА; за результатами симуляційних експериментів досягнуто збільшення нагороди агентів на $\approx 50\%$, збільшення успішного завершення епізодів на $\approx 50\%$, а також зменшення помилок агентів на шляху до цілі на $\approx 10\%$. **Висновки**: створена інтегрована модель дає змогу ефективно управляти флотиліями БпЛА в умовах мінливого повітряного середовища й водночас значно підвищити безпеку та оптимальність маршрутів; використання адаптивних алгоритмів і теоретико-графових моделей забезпечує високу точність прогнозування та мінімізацію ризиків; результати дослідження підтверджують перспективність упровадження розроблених алгоритмів для управління БпЛА в міських і регіональних умовах.

Ключові слова: безпілотні літальні апарати; ройове управління; адаптивна маршрутизація; PID-регулювання; оброблення інформації в реальному часі; оптимізація маршрутів.

Бібліографічні описи / Bibliographic descriptions

Єна М. В., Погудіна О. К. Інтегрована симуляційна модель ройового управління й адаптивної маршрутизації БпЛА в умовах змінного повітряного середовища. *Сучасний стан наукових досліджень та технологій в промисловості*. 2025. № 4 (34). С. 32–43. DOI: https://doi.org/10.30837/2522-9818.2025.4.032

Yena, M., Pohudina, O. (2025), "Integrated simulation model of swarm control and adaptive routeing of UAVS in a changing air environment", *Innovative Technologies and Scientific Solutions for Industries*, No. 4 (34), P. 32–43. DOI: https://doi.org/10.30837/2522-9818.2025.4.032