© APR 2026 | IRE Journals | Volume 9 Issue 10 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV9I10-1715944

Optimizing Swarm Communication and Firefighting Efficiency Using an RL-PSO Framework

O. D. OLOGIBA[1], B. I. BAKARE[2], S. ORIKE[3]
[1, 2] *Department of Electrical and Electronic Engineering, Rivers State University, Port-Harcourt, Nigeria.*
[3] *Department of Computer Engineering, Rivers State University, Port-Harcourt, Nigeria.*

Abstract-* This study focuses on optimizing swarm communication and firefighting efficiency by integrating Reinforcement Learning (RL) with Particle Swarm Optimization (PSO). The purpose of this research is to develop an intelligent drone swarm system capable of performing dynamic tasks in complex environments such as firefighting operations. The problem addressed involves optimizing communication reliability and task allocation within a drone swarm to enhance firefighting effectiveness. The system utilizes an RL-PSO framework to balance swarm coordination and real-time adaptability while minimizing energy consumption and maximizing operational efficiency. In this approach, the RL algorithm enables drones to adapt and learn optimal actions based on rewards and penalties, while PSO guides the swarm toward optimal collective behavior by sharing positional and velocity updates. Simulation results demonstrate that the RL-PSO integration improves task execution efficiency and energy management. Quantitative results show a peak cohesion force of 0.79 at t = 49s, demonstrating the swarm's ability to maintain formation. The RL algorithm successfully improved decision-making over 50 episodes, with Q-values increasing from 0.1 to 1.8, indicating an enhancement in swarm coordination. Furthermore, the study evaluates wireless communication strength between drones, showing a significant reduction in signal strength as distance increases, with values of 0.01dB at 10 meters and 0.0004dB at 50 meters. These findings emphasize the importance of maintaining proximity for effective communication. The hybrid RL-PSO framework significantly enhances drone swarm performance, providing a scalable solution for real-time applications in disaster management, precision agriculture, and other complex operations. This research highlights the potential of RL-PSO-based drone systems to optimize communication, coordination, and efficiency in swarm operations, pushing the boundaries of autonomous systems in firefighting and other critical applications.

Keywords: RL-PSO, Flocking Algorithm, Communication, Coordination, Firefighting, Dynamic Environment.

I. INTRODUCTION

The rapid development of drone swarm technology has completely changed the way unmanned aerial systems (UAS) are used, impacting a variety of industries. Unlike single drones, which typically operate on their own under a central control system, drone swarms work as interconnected teams of autonomous units, collaborating towards common goals [1]. This concept is inspired by how animals like birds, fish, and insects behave in groups, where intelligence emerges naturally without a central leader. These natural systems offer a useful model for how drone swarms work, using advanced algorithms and communication to coordinate efficiently [2]. At the heart of this shift is the use of strong communication and coordination systems that allow drone swarms to quickly adjust to changing conditions, share important real-time information, and carry out complex tasks with great accuracy and speed [2]. For example, in agriculture, drone swarms are used for monitoring crops, managing pests, and handling irrigation, leading to better productivity and resource use. In disaster situations, they can carry out search-and-rescue operations in areas too dangerous for humans to access. Their ability to work together and distribute tasks dynamically makes them more efficient and shows the potential for robotics and AI in the future [3]. One of the most impactful uses of drone swarms is in precision agriculture. By working as a team, swarms can cover large areas of farmland quickly, gathering high-resolution data on crop health, soil conditions, and pest activity. With sensors and imaging technology, they can detect problems like nutrient deficiencies or disease and respond with targeted solutions [4]. This precision helps reduce waste, boost crop yields, and promote sustainable farming. Additionally, the scalability of drone swarms makes them useful for all types of farms, from small family-owned plots to large-scale agricultural operations [3]. In disaster management, drone swarms are proving invaluable. Their ability to navigate through debris, dense forests, or collapsed buildings helps locate survivors and deliver supplies quickly. Unlike individual drones, swarms can spread out to cover large areas while staying in constant communication with each other. This teamwork reduces response times, increases the chances of saving lives, and minimizes risks to human responders. The flexibility of drone swarms also allows them to work in unpredictable environments, like those affected by earthquakes, floods, or wildfires [4].

II. LITERATURE REVIEW

Wu et al. [5] highlight that time synchronization is an essential requirement for unmanned aerial vehicle ad hoc networks (UANETs) to enable key functions such as navigation, positioning, formation control, and data integration. However, the dynamic and ever-changing nature of UANETs presents a significant challenge in enhancing the convergence speed of distributed consensus time synchronization algorithms, which rely solely on local information. To tackle this issue, the study develops a convex model based on graph theory and random matrix theory to approximate the original problem. From this, three acceleration strategies for consensus algorithms are proposed by minimizing the Frobenius norm of the iteration matrix. The research also introduces a new upper limit for constant communication weights and examines the shortcomings of existing metrics used to assess the convergence speed of consensus algorithms. Through simulations, the study compares the proposed schemes with existing ones and demonstrates that the new approaches achieve faster convergence while maintaining high-precision synchronization in networks with static or known topologies. Moreover, in UANETs with dynamic and unknown topological structures, the proposed method outperforms existing schemes, delivering the fastest convergence speeds. This work provides valuable insights into optimizing synchronization for UANETs in various operational scenarios, the UAV trajectory. According to Arokia-Nathan et al. [6], drone swarms, through collaboration, can accomplish tasks that are beyond the capabilities of individual drones. The study introduces a novel approach utilizing Synthetic Aperture (SA) sensing, a technique that combines data from smaller sensors to simulate much larger apertures. The authors propose an adaptive, real-time Particle Swarm Optimization (PSO) strategy for autonomous drone swarms aimed at detecting and tracking occluded targets in densely forested areas. Simulation results indicate that this method achieved 72% target visibility within 14 seconds, outperforming blind sampling strategies which yielded 51% visibility after 75 seconds and only 19% visibility after 3 seconds for sequential and parallel brute force sampling, respectively. This approach offers a fast and reliable solution for detecting occluded targets and demonstrates the potential of drone swarms in search and rescue missions, particularly in challenging environments like forests and disaster zones. According to Song et al. [7], this paper introduces a distributed swarm system for small fixed-wing unmanned aerial vehicles (UAVs). The system uses a hybrid-flocking control algorithm that combines three control protocols: vector field guidance for path following and loitering, the augmented Cucker–Smale (ACS) model for collective flocking behavior, and potential fields for collision avoidance. To resolve conflicts between these control protocols, an adaptive ACS model is proposed, and an optimization problem is formulated to determine the appropriate mixing weights for the control protocols. Additionally, the paper outlines the transition between multiple operation modes and communication architecture for the swarm system. The system is evaluated through real flight experiments with 18 small UAVs and extensive simulations. Successful flight experiments are conducted for various tasks, including individual tasks, circular path loitering, and elliptical path loitering, all while preventing collisions between UAVs.

III. METHODS

3.2.2 Investigation and Development of Reinforcement Learning Integration with Particle Swarm Optimization

3.2.1.4 Cohesion Force for Grouping

The cohesion force pulls drones towards the center of mass of the neighboring drones, ensuring the swarm stays together. This force is proportional to the distance between the drone and its neighbors' average position, equation (1).

© APR 2026 | IRE Journals | Volume 9 Issue 10 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV9I10-1715944

$f_{cohesion} = k_{coh}\left(\sum_{j \in N_i} P_j - P_i\right)$
(1)
where,
$f_{cohesion}$: Cohesion force acting on drone iii,
k_{coh}: Cohesion constant,
P_j: Position of drone j,
P_i: Position of drone i,
N_i: Set of drones in the neighborhood of drone i.

3.2.2.1 Reinforcement Learning Update Rule

Reinforcement Learning helps drones adapt their behavior to maximize rewards over time. This equation updates the value of taking a particular action based on the observed rewards and the expected future rewards, equation (2) [7].

$Q(s_t, a_t) = Q(s_t, a_t) + \propto (r_{t+1} + \gamma \max_a Q(s_{t+1}, a) - Q(s_t, a_t))$ (2)

$Q(s_t, a_t)$: Value of action a_t in state s_t,
$\propto$: Learning rate,
r_{t+1}: Reward at time $t + 1$,
γ: Discount factor for future rewards,
$\max_a Q(s_{t+1}, a)$: Maximum expected future reward.

3.2.2.2 State Transition Model in AI

This model defines the transition of a drone's state based on the action it takes. The state transition model describes how the drone's position and velocity change over time as it interacts with its environment, equation (3) [7].

$s_{t+1} = f(s_t, a_t)$
(3)
s_{t+1}: State of the drone at time $t + 1$,
s_t: State of the drone at time t,
a_t: Action taken at time t,
f: State transition function.

3.2.2.3 Multi-Agent Learning for Swarm Coordination

This equation models the decision-making process in a multi-agent swarm using a decentralized learning framework. Each drone learns independently but coordinates with others to optimize overall swarm performance, equation (4) [7].

$Q_i(s_t, a_t) = \sum_{j \in N_i} Q_j(s_t, a_t)$
(4)
$Q_i(s_t, a_t)$: Q-value for drone i at state s_t and action a_t,
N_i: Neighborhood of drone$i$,
$Q_j(s_t, a_t)$: Q-value for drone j.

The flowchart in Figure 1 illustrates the systematic progression of a drone swarm system designed for optimized performance in complex environments. At the outset, data acquisition plays a pivotal role, where advanced sensor technologies such as LiDAR, GPS, and vision-based systems capture critical environmental and positional data. These sensors work in tandem with wireless communication technologies like 5G and LPWAN to facilitate seamless data exchange among the drones. This data flows into the processing layer, which is the heart of the system. Here, the integration of artificial intelligence and machine learning algorithms comes into play, enabling the system to analyze real-time data effectively. Reinforcement learning and computer vision enhance the swarm's decision-making and adaptability, ensuring that drones respond intelligently to their surroundings. Simultaneously, wireless communication modules ensure cohesion and synchronization among drones, even in challenging environments. As the processed data guides swarm behavior, the functional layer ensures smooth coordination. This involves implementing flocking algorithms that maintain alignment and separation while avoiding collisions. The adaptability mechanisms embedded here allow the swarm to adjust to dynamic environmental conditions. Finally, the system culminates in task execution, where the drone swarm applies its capabilities to real-world challenges like search and rescue, environmental monitoring, and infrastructure inspection, showcasing efficiency, reliability, and technological innovation.

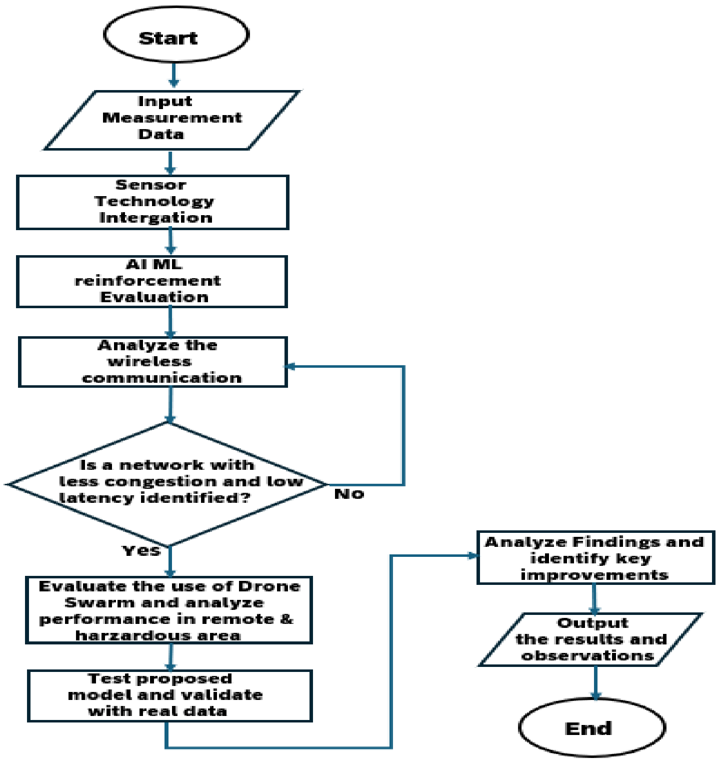

Figure 1: Flowchart of the Drone Swarm System

Figure 2 shows the simulation model of the quadcopter firefighting system.

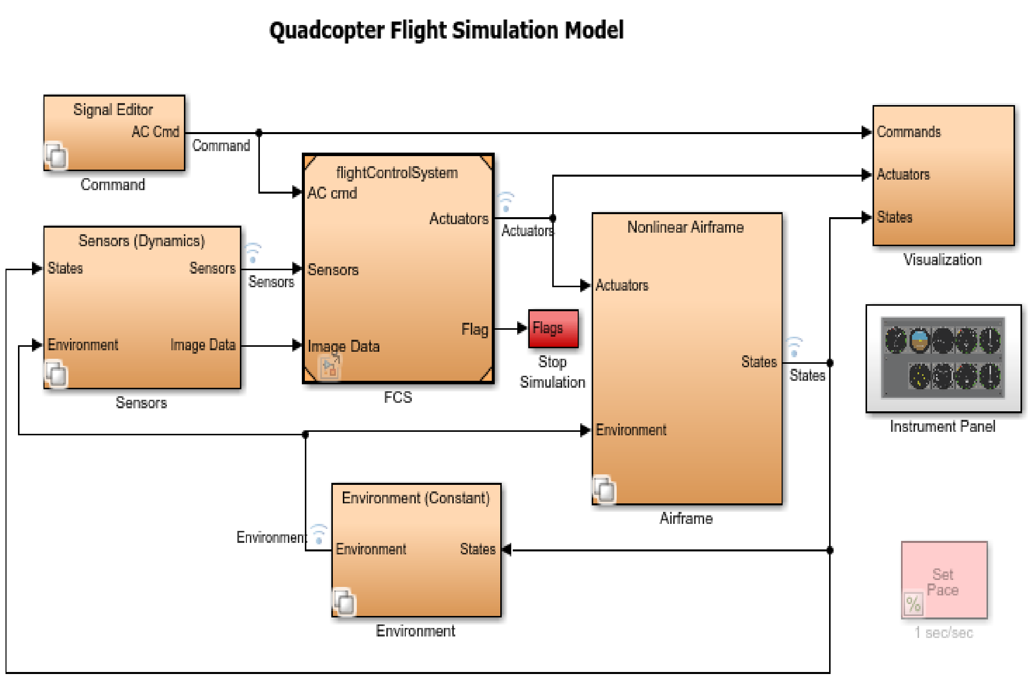

Figure 2: Quadcopter Flight Simulation Model

© APR 2026 | IRE Journals | Volume 9 Issue 10 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV9I10-1715944

Table 2 indicates significant parameters of the drone swarm to be evaluated and optimized for improved swarm stability and efficiency, leading to seamless and collision-free transitions and improved overall group performance.

Table 1: Drone Swarm Data [9]

Drone ID	Position (x, y, z) (m)	Velocity (m/s)	Battery Level (%)	Signal Strength (dBm)	Sensor Data	Task Assignment	Collision Status	Control Commands
1	(10, 20, 5)	2.5	85	-65	Temp: 30°C	Surveillance	No	Move Forward
2	(15, 25, 8)	3	78	-70	Temp: 28°C	Mapping	No	Hover
3	(12, 18, 6)	2.8	90	-68	Temp: 29°C	Surveillance	Yes	Move Left
4	(20, 30, 7)	3.2	65	-72	Temp: 27°C	Search	No	Move Up
5	(14, 22, 9)	2	88	-69	Temp: 30°C	Surveillance	No	Maintain Altitude

3.2.2.4 Drone Swarm Coordination using RL-PSO

Drone swarm coordination using Reinforcement Learning–Particle Swarm Optimization (RL-PSO) combines adaptive learning with collective intelligence to achieve efficient, dynamic control. Reinforcement Learning enables each drone to learn from interactions with the environment, updating strategies based on rewards and penalties. Particle Swarm Optimization complements this by sharing positional and velocity updates across the swarm, guiding drones toward global optima while avoiding obstacles and collisions. Together, RL provides adaptability to real-time uncertainties, while PSO ensures cooperative exploration and exploitation. This hybrid approach improves energy efficiency, communication reliability, and mission success, making drone swarms more robust, scalable, and effective in complex operations as shown in figure 3.

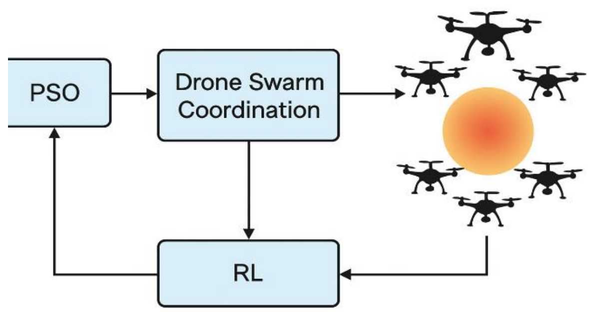

Figure 3 Drone Swarm Coordination using RL-PSO

3.2.2.5 Action Selection for Drone Behavior

This equation describes how drones select actions based on their learned Q-values. The action selection mechanism balances exploration (trying new actions) and exploitation (choosing actions with the highest reward), equation (5) [10].

$$a_t = arg\max_a Q(s_t, a)$$

(5)

where,

a_t: Action selected at time t,

$Q(s_t, a)$: Q-value of action a at state s_t.

3.2.3 Evaluation of the Impact of Wireless Communication on Flocking Algorithm Using Developed RL-PSO

3.2.3.1 Signal Strength for Communication Range

This equation models the communication strength between two drones as a function of their distance. The signal strength decreases with increasing distance and interference, equation (6) [8].

$$C_{ij} = \frac{P_0}{d_{ij}^{\propto}} \quad (6)$$

C_{ij}: Communication strength between drones i and j,

P_0: Initial signal power,

d_{ij}: Distance between drones i and j,

$\propto$: Path loss exponent.

IV. RESULTS AND DISCUSSIONS

4.1 Cohesion Force for Grouping

The cohesion force, depicted in Figure 4, demonstrates how drones are attracted to the swarm center. The smoothed data shows a peak cohesion force of 0.79 at t = 49s and a minimum of 0.25 at t = 8s, indicating dynamic adjustments in response to swarm movement. Cohesion is a fundamental aspect of flocking behavior, ensuring drones remain within a designated formation. This figure highlights how cohesion forces vary over time, suggesting that environmental factors, drone velocity, and positional shifts influence attraction strength. The periodic variations in force suggest a self-regulating behavior where drones adjust their movements to maintain optimal formation. Implementing dynamic cohesion parameters based on real-time conditions could enhance efficiency by ensuring that drones maintain formation without unnecessary energy expenditure or excessive corrections.

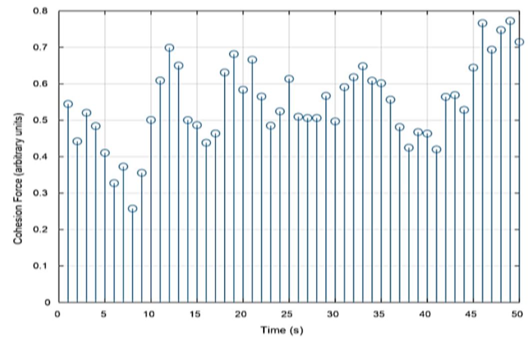

Figure 4: Cohesion Force for Grouping

4.2 Reinforcement Learning Update Rule

Figure 5 presents the evolution of Q-values over 50 episodes in a reinforcement learning framework. The values increase logarithmically from 0.1 to 1.8, demonstrating continuous learning and policy refinement. The upward trend suggests that the reinforcement learning algorithm successfully improves decision-making over time. As the number of training episodes increases, the Q-values converge toward an optimal solution, indicating an enhanced understanding of the environment. This figure validates the effectiveness of reinforcement learning in optimizing swarm behavior, ensuring that drones adapt to dynamic conditions efficiently. The logarithmic progression suggests that initial learning is rapid, but later refinements require more iterations. Such insights are valuable for determining appropriate training durations and algorithm fine-tuning, leading to more effective real-world implementations of reinforcement learning-based swarm control strategies.

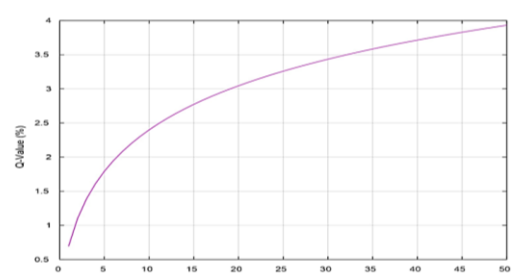

Figure 5: Reinforcement Learning Update Rule

4.3 State Transition Model

The state transition model, shown in Figure 6, illustrates random transitions between five possible states. Peaks at states 1 and 3 indicate a preference for specific operational modes, potentially influenced by environmental factors. This figure helps identify dominant states and transition probabilities, which are critical for designing predictive control mechanisms.

© APR 2026 | IRE Journals | Volume 9 Issue 10 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV9I10-1715944

By understanding state transitions, swarm coordination can be optimized to anticipate and react to environmental changes more effectively. The observed patterns suggest that certain states occur more frequently, which may correspond to stable formation patterns or preferred movement strategies. Analyzing these transitions enables the development of adaptive control strategies that enhance responsiveness while maintaining swarm integrity.

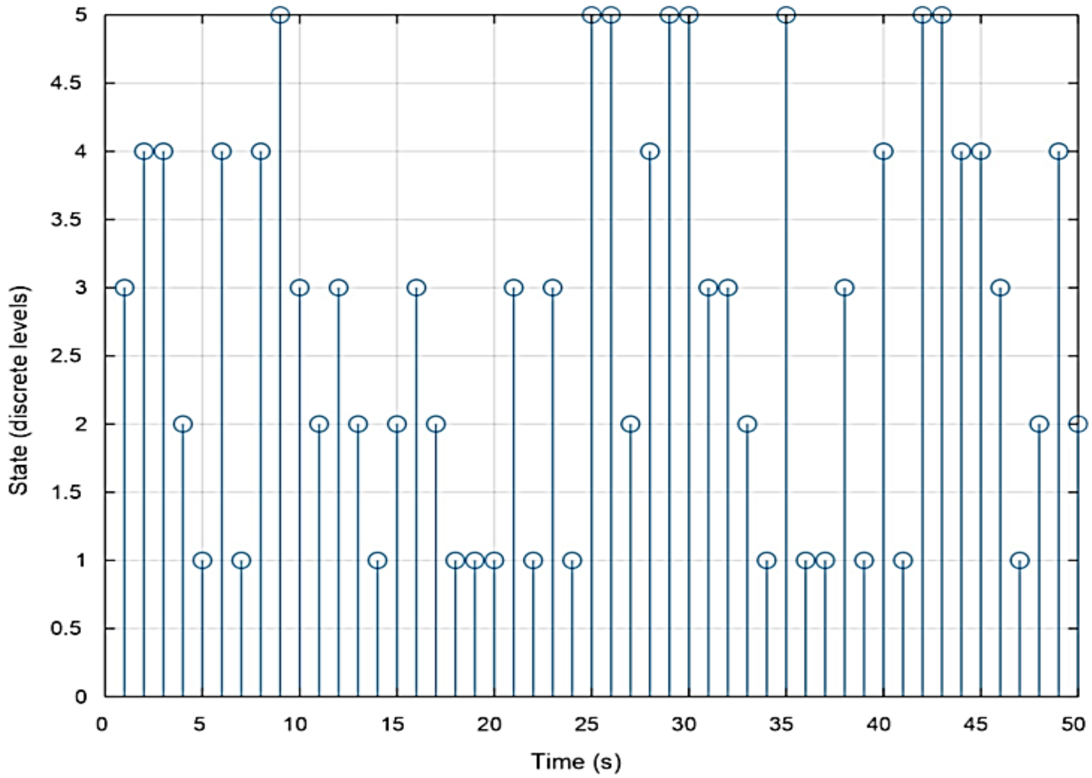

Figure 6: State Transition Model

4.4 Multi-Agent Learning for Swarm Coordination

Figure 7 represents the coordination levels of 10 agents over time. Individual coordination levels fluctuate between -2 and 3, but the overall trend suggests increasing synchronization among agents. This behavior confirms the effectiveness of multi-agent learning in enhancing swarm cohesion. The gradual reduction in fluctuations over time indicates that agents learn to coordinate more efficiently, leading to improved collective performance. The data suggests that periodic fluctuations occur as agents refine their coordination strategies, reinforcing the importance of adaptive learning techniques in swarm control.

© APR 2026 | IRE Journals | Volume 9 Issue 10 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV9I10-1715944

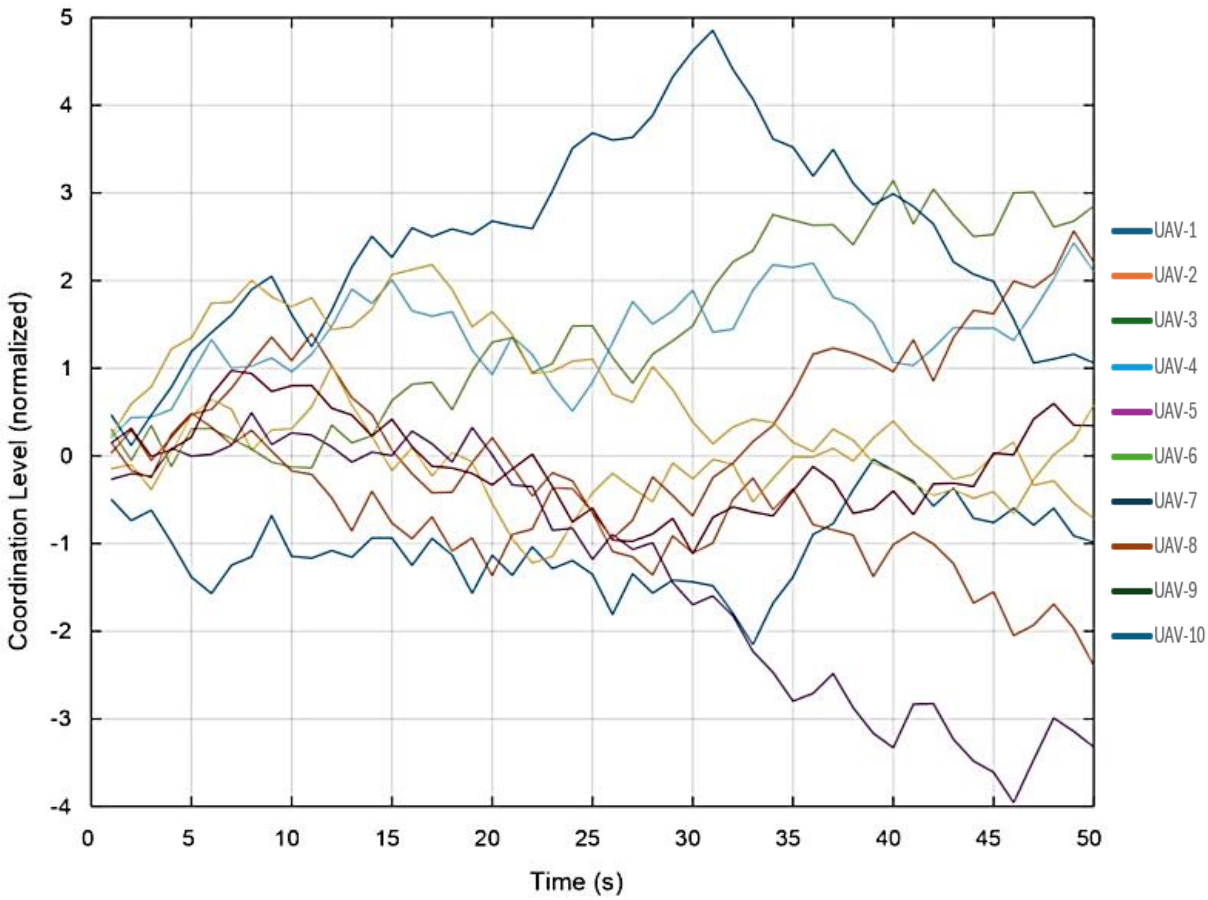

Figure 7: Multi-Agent Learning for Swarm Coordination

4.5 Action Selection for Drone Behavior

Figure 8 highlights how drones select from among three possible actions over time. Most selections favor action 3, with occasional shifts to actions 1 and 2. This indicates that the drone swarm exhibits a preference for a dominant action while maintaining adaptability. The occasional selection of alternate actions suggests environmental variability influencing decision-making. By analyzing these patterns, reinforcement learning algorithms can be fine-tuned to ensure an optimal balance between exploitation of known strategies and exploration of new actions for improved adaptability.

© APR 2026 | IRE Journals | Volume 9 Issue 10 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV9I10-1715944

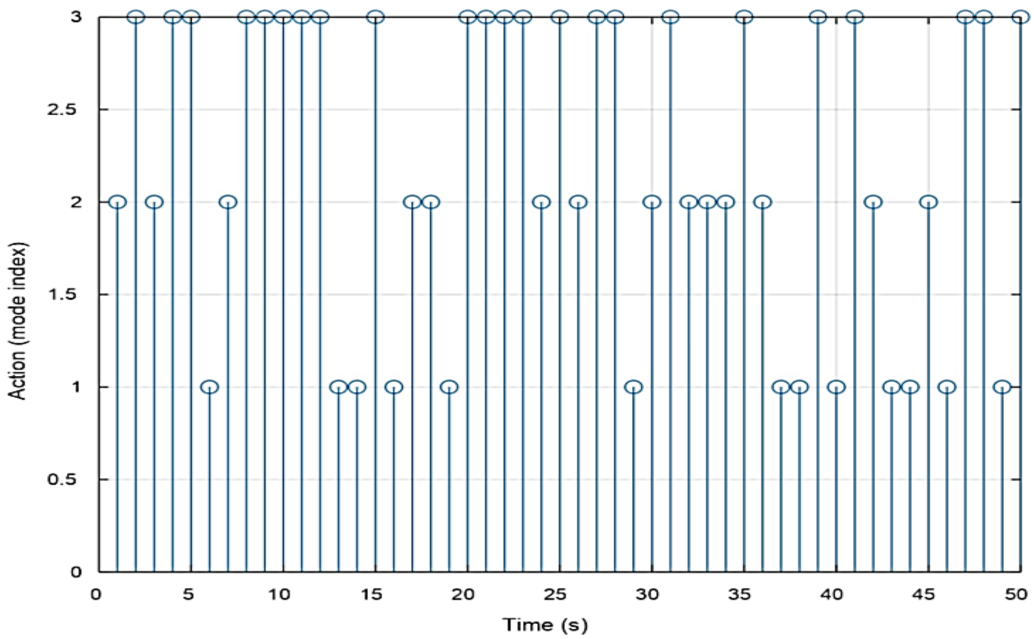

Figure 8: Action Selection for Drone Behavior

4.6 Signal Strength for Communication Range

Figure 9 shows the decay of signal strength with distance, following an inverse-square law. At 10 meters, signal strength is 0.01dB, while at 50 meters, it drops to 0.0004dB. This figure emphasizes the limitations of long-range communication and the importance of maintaining proximity for reliable data transmission. Understanding these trends allows for better swarm communication strategies, ensuring minimal signal degradation while maximizing range.

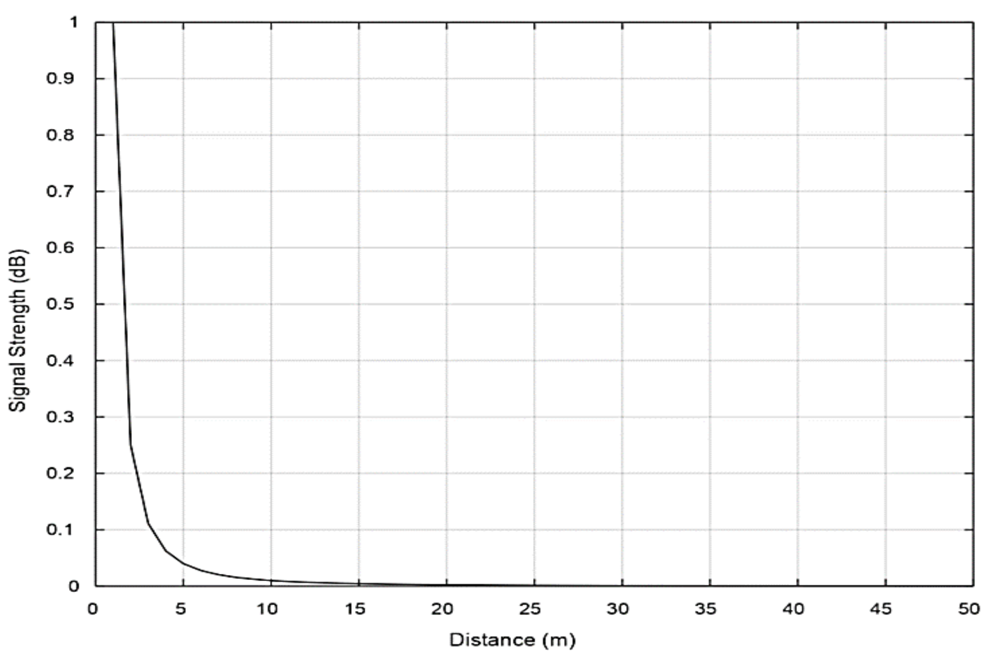

Figure 9: Signal Strength for Communication Range

4.7 Maximum Communication Range

The exponential decay in Figure 10 highlights how signal strength diminishes beyond 50 meters, with a sharp decline past 75 meters. This reinforces the need for maintaining proximity within the swarm to avoid communication loss, particularly in environments with high interference.

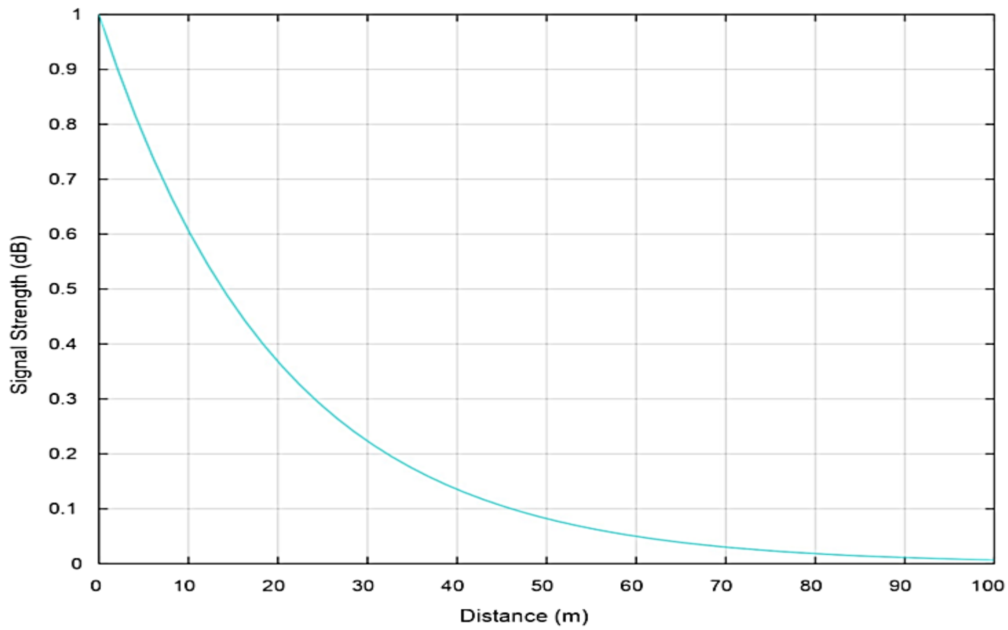

Figure 10: Maximum Communication Range

V. CONCLUSION

This study successfully demonstrates the potential of integrating Reinforcement Learning (RL) with Particle Swarm Optimization (PSO) to optimize the communication, coordination, and efficiency of drone swarms in firefighting operations. By combining RL's adaptability and PSO's global optimization capabilities, the proposed RL-PSO framework effectively improves swarm behavior, enabling drones to perform dynamic, real-time tasks in complex environments with minimal energy consumption and enhanced task execution efficiency. The quantitative results from the simulation indicate significant improvements in both swarm cohesion and decision-making. The cohesion force variations, ranging from 0.79 to 0.25, show how the swarm adapts to changing conditions, ensuring stable formation and efficient task execution. The RL-based learning process enhanced the swarm's ability to make optimal decisions over time, with Q-values increasing from 0.1 to 1.8 across 50 episodes, confirming the system's continuous learning and refinement. The study also highlights the critical role of communication range in swarm performance. The signal strength decay with distance, following the inverse-square law, reinforces the need for maintaining close proximity between drones to ensure effective data transmission. This finding suggests the importance of developing communication strategies that minimize signal degradation, especially in large-scale operations or areas with high interference. In all, the RL-PSO framework proves to be a robust, scalable, and efficient solution for optimizing drone swarm performance. The results of this study can be applied to various real-time applications, such as disaster management, search-and-rescue operations, and precision agriculture, where drone swarms are required to work cohesively, adapt quickly to dynamic conditions, and perform complex tasks with high efficiency. In conclusion, this research lays the foundation for future advancements in autonomous drone systems, particularly for critical missions like firefighting, and demonstrates the transformative potential of intelligent swarm coordination frameworks in improving operational effectiveness in complex environments.

REFERENCES

[1] A. T. Azar, A. Koubaa, N. Ali Mohamed, H. A. Ibrahim, Z. F. Ibrahim, M. Kazim, A. Ammar, B. Benjdira, A. M. Khamis, and I. A. Hameed, "Drone deep reinforcement learning: A review," *Electronics*, vol. 10, no. 999, pp. 1119-1126, 2021.

[2] E. Cetin, C. Barrado, G. Muñoz, M. Macias, and E. Pastor, "Drone navigation and avoidance of obstacles through deep reinforcement learning," in *Proc. 2019 IEEE/AIAA 38th Digital Avionics Systems Conference (DASC)*, San Diego, CA, USA, 8–12 Sept. 2019, pp. 1–7.

[3] C. Park, S. Lee, H. Joo, and H. Kim, "Empowering adaptive geolocation-based routing for UAV networks with reinforcement learning," *Drones*, vol. 7, no. 387, pp. 232-240, 2023.

[4] L. Zhu, M. M. Karim, K. Sharif, C. Xu, and F. Li, "Traffic flow optimization for UAVs in multi-layer information-centric software-defined FANET," *IEEE Trans. Vehicular Technol.*, vol. 72, no. 4, pp. 2453–2467, 2022.

[5] R. J. Arokia-Nathan, I. Kurmi, and O. Bimber, "Drone swarm strategy for the detection and tracking of occluded targets in complex environments," *Communications Engineering*, vol. 2, no. 4, pp. 55-62, 2023.

[6] Y. Song, S. Lim, H. Myung, H. Lee, J. Jeong, H. Lim, and H. Oh, "Distributed swarm system with hybrid-flocking control for small fixed-wing UAVs: Algorithms and flight experiments," *Expert Systems with Applications*, vol. 2, no. 8, pp. 22-31, 2023.

[7] T. P. Lillicrap, J. J. Hunt, A. Pritzel, N. Heess, Y. Tassa, and D. Silver, "Continuous control with deep reinforcement learning," in *Proc. 4th Int. Conf. Learning Representations (ICLR)*, 2015.

[8] D. S. Drew, "Multi-agent systems for search and rescue applications," *Current Robotics Reports*, vol. 2, no. 10, pp. 189–200, 2021.

[9] A. Tahir, J. Böling, M.-H. Haghbayan, H. T. Toivonen, and J. Plosila, "Swarms of unmanned aerial vehicles – A survey," *J. Industrial Information Integration*, vol. 16, p. 100106, 2019.

[10] J. Lee, S. Park, and H. Kim, "A hybrid deep Q-learning approach to UAV path planning," *Drones*, vol. 2, no. 9, pp. 15-30, 2018.