

Hybrid AI Algorithm for Optimizing Radio Coverage in UAV-based Wireless Mesh Networks*

Eryk DŁUGOSZ and Krzysztof MAŚŁANKA

Faculty of Electronics, Military University of Technology, Warsaw, Poland

Correspondence should be addressed to: Eryk DŁUGOSZ, eryk.dlugosz@wat.edu.pl

*Presented at the 47th IBIMA International Conference, 29-30 June 2026, Madrid, Spain

Abstract

This article introduces a hybrid GA–BO–DRL optimization framework designed to enhance ground node coverage in UAV enabled wireless mesh networks. The proposed methodology integrates a global search mechanism based on a genetic algorithm, a Bayesian optimization module for local refinement, and a deep reinforcement learning agent enabling real time adaptation to node mobility. The combined approach maximizes coverage, minimizes the number of active UAV platforms, and maintains network robustness under dynamic conditions. Simulation results demonstrate significant improvements in coverage efficiency and energy conservation, underscoring the method's applicability to time critical scenarios such as disaster relief operations and military deployments..

Keywords : Unmanned aerial vehicles; Wireless mesh networks; Genetic algorithm; Bayesian optimisation

Introduction

Advancements in unmanned aerial vehicle (UAV) technologies have facilitated their increasing deployment across both civilian and defense oriented communication systems. A particularly promising application involves the establishment of UAV based wireless mesh networks (WMNs), wherein drone platforms cooperatively form an adaptable communication infrastructure. In modern military and crisis response environments, where conventional infrastructure may be absent or damaged, rapid provisioning of reliable communication links is of paramount importance. UAV based WMNs have been widely investigated as a means of providing resilient, rapidly deployable communication backbones (Wei et al. (2018), Zhou J. et al. (2025), Davoli et al. (2021), Esrafilian et al. (2020), Aljaber A.R. (2019)).

A fundamental challenge in such networks concerns the optimal placement and mobility control of UAVs to ensure maximal coverage of ground nodes. The problem exhibits substantial complexity due to the combinatorial nature of possible UAV configurations and the dynamic spatial distribution of ground users. Consequently, the optimal configuration is not static, and real time re optimization is required as ground nodes reposition themselves. Classical approaches, such as predefined UAV placements or heuristic coverage strategies generally fail to achieve adequate performance in highly dynamic environments.

Cite this Article as: Eryk DŁUGOSZ and Krzysztof MAŚŁANKA Vol. 2026 (12) "Hybrid AI Algorithm for Optimizing Radio Coverage in UAV-based Wireless Mesh Networks " Communications of International Proceedings, Vol. 2026 (12), Article ID 4718626, <https://doi.org/10.5171/2026.4718626>

This paper presents a hybrid algorithm combining genetic algorithm (GA), Bayesian optimization (BO), and deep reinforcement learning (DRL) to solve the UAV deployment problem. The hybrid GA-BO-DRL algorithm operates in stages: first, the GA performs a global search for good UAV deployment configurations. Afterwards, the BO fine-tunes the best solution and finally, the DRL agent continuously adjusts the UAV positions in response to the movements of ground nodes during operation. By combining these methods, this approach aims to achieve high initial coverage (thanks to GA and BO) and maintain coverage over time despite node mobility (thanks to DRL). This paper focuses on optimizing ground node coverage using this hybrid algorithm. The following sections detail the current state of the problem, the characteristics of the proposed solution, simulation results, an assessment of the solution's applicability, possible further development, and conclusions.

Optimization in UAV-based wireless mesh networks

Numerous studies have examined UAV placement and mobility optimization using heuristic, evolutionary, and machine learning methods. Traditional evolutionary algorithms—particularly genetic algorithms and particle swarm optimization—offer a robust mechanism for exploring large configuration spaces without requiring gradient information (Goldberg and Holland (1988) as well as Senel and Aschenbruck (2018)). However, evolutionary strategies alone tend to converge slowly and usually generate static deployment plans unsuitable for dynamic environments.

More recently, reinforcement learning (RL) and deep reinforcement learning (DRL) methods have gained prominence in UAV networking research. DRL enables autonomous agents to learn mobility strategies by interacting with the environment, thus allowing UAV platforms to adaptively respond to variations in ground node distribution, environmental obstacles, and communication quality constraints (Sutton and Barto (2018) and Sewak (2019)). Nevertheless, purely DRL based systems require long training times, exhibit sensitivity to reward function design, and may experience instability during learning, especially in multi UAV contexts.

Hybrid strategies, which combine classical optimization techniques with machine learning components, have emerged as a promising direction for improving convergence and operational robustness (Liang et al. (2021)). The approach presented in this paper contributes to this research stream by tightly integrating GA based global search, BO driven local refinement, and DRL based adaptive control into a unified optimization pipeline.

Hybrid AI algorithm for optimizing radio coverage in WMNs

The proposed solution is a hybrid algorithm that sequentially integrates a genetic algorithm, Bayesian optimization, and deep reinforcement learning to optimize the coverage of ground-based UAV nodes. The process begins with a GA that generates an initial population of UAV deployment solutions. Each solution (individual) encodes the positions (and active status) of all UAVs in the network. A problem-specific fitness function evaluates each individual by considering the number of covered ground-based nodes (to be maximized) and penalizing undesirable attributes (such as using too many UAVs, overlapping coverage areas, or placing UAVs outside the allowed area). The GA uses selection, crossover, and mutation operations to evolve the population over multiple generations, gradually improving the coverage. This yields a set of candidate distributions, and the best solution found by the GA serves as a strong starting point.

The best solution found by the GA is then passed to the Bayesian optimization stage. BO treats the placement optimization as a function optimization problem (Snoek et al. (2012)). BO selects further candidate configurations for evaluation (either through simulation or using a model) that may improve coverage or other metrics. Because BO can intelligently explore small adjustments, it fine-tunes the BSP positions more efficiently than random search. In practice, BO may slightly adjust the BSP height or coordinates to cover an additional node or reduce overlap, thereby improving the GA solution based on algorithm published in Gong et al. (2022).

Finally, a DRL agent was used to handle the movements of ground nodes over time. The agent observes the network state (e.g., current UAV positions, which ground nodes are covered or uncovered, etc.) at each time step and can issue actions to move some UAVs to new positions. The agent's policy is trained in simulation to maximize cumulative rewards, where the reward can be defined based on instantaneous range. In this approach, the DRL agent starts with an optimized static deployment (from GA + BO) and learns to make incremental adjustments as needed

when some ground nodes move. This allows the system to consistently maintain high coverage without having to restart the GA from scratch for each change. The DRL component effectively provides real-time adaptability by adjusting UAV positions in response to environmental changes. A simplified DRL model, limited to a moderate neural network size and planning horizon, was used to ensure the solution could operate in real time. The overall optimization objective is defined as:

$$f_1(\Omega) = \frac{Coverage^{\alpha_1}}{\alpha_2 * UAVno} - P(\Omega) \quad (1)$$

where:

- Ω – UAV configuration (positions and activation states),
- α_1, α_2 – weighting factors balancing coverage and resource cost,
- UAVno – number of active UAVs,
- Coverage – number of ground nodes covered by at least one UAV,
- $P(\Omega)$ – penalty term for overlap or boundary violation:

$$P(\Omega) = \alpha_{ov} \cdot P_{ov}(\Omega) + \alpha_{bd} \cdot P_{bd}(\Omega) \quad (2)$$

- $P_{ov}(\Omega)$ – penalty for overlapping UAV range,
- $P_{bd}(\Omega)$ – penalty for crossing the border of the area of operation,
- α_{ov} – UAV range overlap penalty factor,
- α_{bd} – penalty factor for exceeding the boundary of the area of operation.

The GA, BO, and DRL components are integrated into a unified framework: GA and BO operate before a mission to determine the initial, near-optimal UAV deployment, and a trained DRL agent then operates during the mission to adapt to changes. This hybrid GA-BO-DRL approach leverages the strengths of each method: GA finds a good global configuration, BO efficiently fine-tunes it, and DRL continuously refines the solution on the fly. The result is an optimized and adaptive coverage solution.

Simulation Results

The hybrid algorithm was implemented and tested in a simulation environment with 20 mobile ground nodes in a 100×100 area and up to 10 UAVs. Parameters can be adjusted to fit different scenarios. Table 1 summarizes model parameters, and Fig. 1 shows an example output.

Table 1. Simulation model parameters

Name		Value	Description
MAX UAV		10	maximum number of UAV
GROUND NODES		20	number of ground nodes
DIM PER DRONE		4	number of UAV dimensions [flag, x, y, z]
DIMENSION		3	number of dimensions of space (x, y, z)
DRONE RANGE		10.0	basic UAV range at minimum altitude
MAX_ALTITUDE		22.0	height to which the range increases linearly (above which constant)
BOUND_LOW, BOUND_UP		0.0, 100.0	area boundaries (x, y coordinates)
MIN_ALTITUDE		0.0	minimum drone height
MAX_ALTITUDE		50.0	maximum height of the drone

Key performance metrics included the number of covered ground nodes (network coverage), the number of actively

deployed UAVs, and the overall cost-to-benefit ratio (CBR), which combines coverage (benefit) and resource utilization (cost). CBR is defined as the ratio of covered nodes to active UAVs. The algorithm was evaluated in several scenarios reflecting different optimization priorities.

Scenario 1 – maximizing coverage

In scenario 1, the target weights were set to prioritize coverage of as many ground nodes as possible, with less emphasis on minimizing the number of UAVs. The GA-BO stages tended to activate the majority of available UAVs to cover almost all ground nodes. Results showed that, on average, ~17 of the 20 ground nodes (85%) were covered using 4 UAVs (with a fifth UAV sometimes deployed if necessary). The CBR in this scenario was high (~1.0), meaning approximately one UAV for every 4 ground nodes. As the ground nodes moved, the DRL agent made frequent adjustments to reposition the UAVs and successfully maintained coverage above 80% at all times. Fig. 2 shows the simulation results.

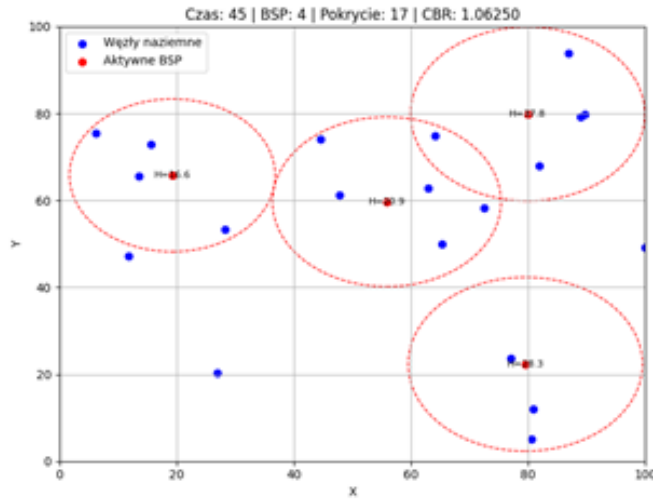


Fig. 1. Example result of the simulation

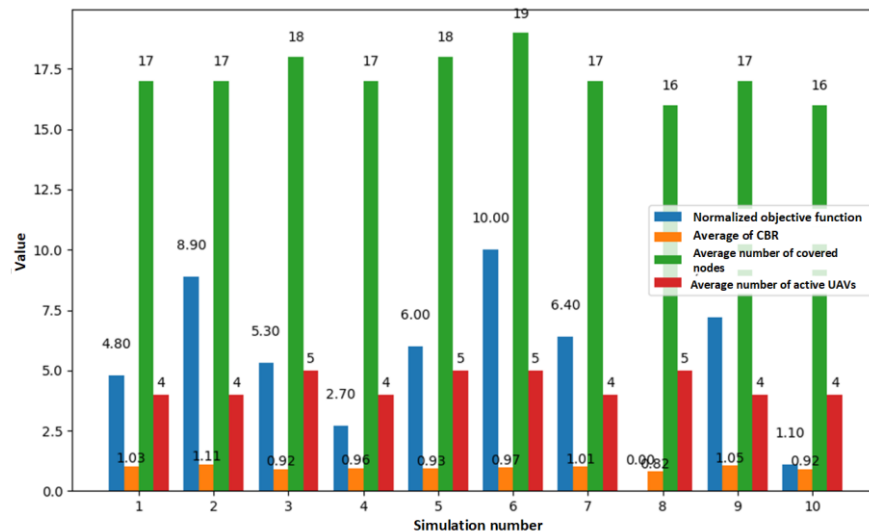


Fig. 2. Simulation results for scenario 1

Scenario 2 – minimizing the number of active UAVs

In scenario 2, the penalty for using additional UAVs was increased, so the algorithm favored solutions with fewer UAVs, even though the coverage decreased slightly. In the GA-BO phases, deployments using only 2-3 UAVs were found. As expected, the coverage was lower, with the network covering approximately 12 ground nodes on average (~60%). However, the CBR was improved (~1.2) because each UAV covered more nodes on average. This scenario is suitable when saving resources (such as available UAVs) is a priority over maximizing coverage. Figure 3 shows the simulation results.

Scenario 3 – Balanced scenario

In scenario 3, balanced weighting was used to avoid extreme behavior, aiming for good coverage with moderate resource utilization. The GA-BO result typically activated 3-4 UAVs covering approximately 15-16 nodes (~75-80% coverage). The CBR was around 1.0-1.1, indicating balanced performance. In the simulation, this scenario appeared to be a compromise: the DRL agent occasionally adjusted UAV positions to cover newly discovered nodes, but also occasionally activated an additional UAV. Overall, coverage remained high (~75% or higher) with higher coverage than in scenario 2. Figure 4 shows the simulation results.

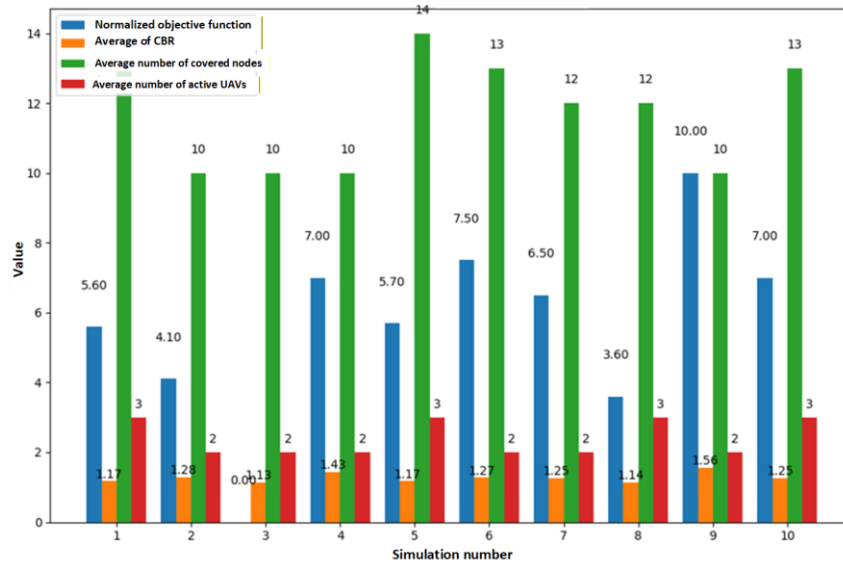


Fig. 3. Simulation results for scenario 2

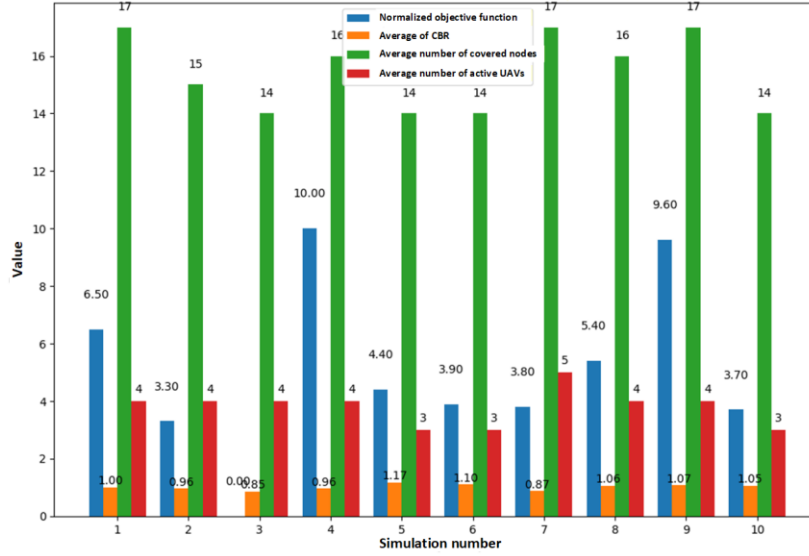


Fig. 4. Simulation results for scenario 3

Presented above results indicate that the proposed GA-BO-DRL algorithm is effective and versatile. It can be configured to favor maximum coverage or resource protection, depending on the needs, making it applicable to a range of real-world scenarios. For example, during a rescue operation, maintaining reliable connectivity for as many victims or responders as possible (Scenario 1 approach), while in a long-term military mission, minimizing the number of active UAVs may be crucial (Scenario 2 approach). A balanced strategy (Scenario 3) would be appropriate where both objectives are moderately important. Importantly, the algorithm components can operate in a distributed manner. Initial GA and BO calculations can be performed by a central command system or prior to deployment. The DRL agent could be deployed centrally or on-board the UAV (each UAV would run an instance or a distributed variant of the multi-agent agent) to make rapid decisions. The GA-BO part may be too slow to run in real time for large-scale problems, but since it is a starting point, it can be done in advance or periodically when major re-optimization is needed.

Future works

One direction for future work is to incorporate multi-agent reinforcement learning, so that each UAV can intelligently interact with others, rather than relying on a single central agent. Such techniques could enable a more scalable solution as the number of UAVs increases. The current DRL agent could also be enhanced with transfer learning, allowing knowledge acquired in one environment (e.g., one disaster area) to be transferred to another, reducing training time (Liu et al. (2024)).

Another area requiring development is the objective function. This study focused on range and basic resource utilization. Future work could include more advanced network performance metrics, such as data throughput or latency. For example, ensuring that each ground node has not only a connection but also a minimum data rate could be formulated as part of the objective function. An energy model could also be incorporated: a complete energy consumption model (including drone flight dynamics and battery limits) would make the solution more suitable for long-duration missions.

Additionally, field testing the algorithm with real UAVs would be an invaluable next step. Simulations often fail to capture all real-world effects (GPS errors, wind, unpredictable obstacles, etc.). By implementing the GA-BO-DRL algorithm on a small fleet of UAVs in a controlled environment, it would be possible to verify that the adaptive range behavior works outside of simulation. Any discrepancies could lead to adjustments in the models or training procedure (for example, adding noise to the simulation to train a more robust DRL agent).

Conclusions

This paper addresses the problem of optimizing ground node coverage in UAV-based wireless mesh networks. A hybrid GA-BO-DRL algorithm is proposed, which combines genetic algorithm, Bayesian optimization, and reinforcement learning to achieve both high initial coverage and continuous adaptation to changes. This approach is tested in simulation, and the results show that it can successfully deploy UAVs to cover a high percentage of ground nodes while intelligently managing the number of UAVs.

Compared to purely static optimization, this method ensures that coverage is not lost when the environment changes: the DRL agent quickly adapts to the mobility of ground nodes. Compared to using a purely reactive learning agent, the hybrid approach starts with a near-optimal configuration, improving performance and reducing the learning burden. The scenarios analyzed showed that by adjusting the parameters, the operator can obtain different results. This flexibility is the main advantage of the method.

In summary, the hybrid GA-BO-DRL algorithm offers a solution for maintaining reliable communication in dynamic UAV environments. It effectively bridges the gap between one-time optimization and real-time control. Future improvements and real-world tests will further validate this approach, but the current results already suggest that such intelligent hybrid optimization can significantly improve the reliability and performance of UAV-assisted networks in practical applications.

Acknowledgment

This work was financed by the Military University of Technology under Research Project no. UGB 531-000128-W300-22.

References

- Wei Z., Guo Z., Feng Z., Zhu J., Zhong J., Wu O, Wu H. (2018), 'Spectrum Sharing between UAV-based Wireless Mesh Networks and Ground Networks', Proceedings of the 10th International Conference on Wireless Communications and Signal Processing (WCSP), Hangzhou, China, 2018, pp. 1-6, doi: 10.1109/WCSP.2018.8555855
- Zhou J. et al. (2025), 'U-Mesh: Deploying UAV-Mesh Network for Automatic Powerline Inspection in Remote Areas' Proceedings of the IEEE 45th International Conference on Distributed Computing Systems (ICDCS), Glasgow, United Kingdom, pp. 1088-1098, doi: 10.1109/ICDCS63083.2025.00109.
- Davoli L, Pagliari E, Ferrari G.(2021), 'Hybrid LoRa-IEEE 802.11s Opportunistic Mesh Networking for Flexible UAV Swarming,' Drones. 5 (2), 26, <https://doi.org/10.3390/drones5020026>
- Esrafilian O., Gangula R. and D. Gesbert (2020), 'Autonomous UAV-aided Mesh Wireless Networks' Proceedings of the IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS), Toronto, ON, Canada, pp. 634-640, doi: 10.1109/INFOCOMWKSHPS50562.2020.9162753.
- Aljaber A.R. (2019), Modeling a UAV-based mesh network to analyze latency and throughput, Naval Postgraduate School, Monterey, California
- Goldberg, D.E., Holland, J.H. (1988), 'Genetic Algorithms and Machine Learning,' Machine Learning, 3, 95–99, <https://doi.org/10.1023/A:1022602019183>
- Senel, F., Aschenbruck, N. (2018), 'Optimizing the Deployment of UAV for Mesh Access Points' Proceedings of the IEEE Int. Conf. on Wireless and Mobile Computing, pp. 104-111.
- Sutton, R.S., Barto, A.G. (2018), Reinforcement Learning: An Introduction, 2nd ed., MIT Press, Cambridge, Massachusetts, USA.
- Sewak, M. (2019), Deep Q Network (DQN), Double DQN and Dueling DQN: A Step Towards General Artificial Intelligence. Deep Reinforcement Learning, Springer Singapore.
- Liang, L.; Wang, J.; Zhang, H. et al. (2021), 'GARL: Genetic-Algorithm-Augmented Reinforcement Learning to Detect Optimal Policies,' Applied Soft Computing, vol. 111, art. 107675
- Snoek, J.; Larochelle, H.; Adams, R.P. (2012) 'Practical Bayesian Optimization of Machine-Learning Algorithms,' Advances in Neural Information Processing Systems, vol. 5, pp. 2951-2959.

- Gong, S.; Wang, M.; Gu, B. et al. (2022) 'Bayesian Optimization for Energy-Efficient UAV Communications,' IEEE Internet of Things Journal, 9 (4), pp. 3127-3140.
- Liu, Y.; Li, X.; Wang, J. et al. (2024) 'Deep-RL-Based Dynamic UAV Deployment for Maritime Coverage,' IEEE Access, vol. 12, pp. 45211-45224.