



LEEDS
BECKETT
UNIVERSITY

Citation:

Seerangan, K and Nandagopal, M and Govindaraju, T and Manogaran, N and Balusamy, B and Selvarajan, S (2024) A novel energy-efficiency framework for UAV-assisted networks using adaptive deep reinforcement learning. *Scientific Reports*, 14. pp. 1-23. ISSN 2045-2322 DOI: <https://doi.org/10.1038/s41598-024-71621-x>

Link to Leeds Beckett Repository record:

<https://eprints.leedsbeckett.ac.uk/id/eprint/11415/>

Document Version:

Article (Published Version)

Creative Commons: Attribution-Noncommercial-No Derivative Works 4.0

© The Author(s) 2024

The aim of the Leeds Beckett Repository is to provide open access to our research, as required by funder policies and permitted by publishers and copyright law.

The Leeds Beckett repository holds a wide range of publications, each of which has been checked for copyright and the relevant embargo period has been applied by the Research Services team.

We operate on a standard take-down policy. If you are the author or publisher of an output and you would like it removed from the repository, please [contact us](#) and we will investigate on a case-by-case basis.

Each thesis in the repository has been cleared where necessary by the author for third party copyright. If you would like a thesis to be removed from the repository or believe there is an issue with copyright, please contact us on openaccess@leedsbeckett.ac.uk and we will investigate on a case-by-case basis.



OPEN

A novel energy-efficiency framework for UAV-assisted networks using adaptive deep reinforcement learning

Koteeswaran Seerangan¹, Malarvizhi Nandagopal², Tamilmani Govindaraju³, Nalini Manogaran⁴, Balamurugan Balusamy⁵ & Shitharth Selvarajan^{6,7}✉

In the air-to-ground transmissions, the lifespan of the network is based on the “unmanned aerial vehicle’s (UAV)” life span because of the limited battery capacity. Thus, the enhancement of energy efficiency and the outage of the ground candidate’s minimization are significant factors of the network functionality. UAV-aided transmission can highly enhance the spectrum efficacy and coverage. Because of their flexible deployment and the high maneuverability, the UAVs can be the best alternative for the situations where the “Internet of Things (IoT)” systems utilize more energy to attain the essential information rate, when they are far away from the terrestrial base station. Therefore, it is significant to win over the few troubles in the conventional UAV-aided efficiency approaches. Thus, this proposed work is aimed to design an innovative energy efficiency framework in the UAV-assisted network using a reinforcement learning mechanism. The energy efficiency optimization in the UAV offers better wireless coverage to the static and mobile ground user. Presently, reinforcement learning techniques effectively optimize the energy efficiency rate of the system by employing the 2D trajectory mechanism, which effectively removes the interference rate attained in the nearby UAV cells. The main objective of the recommended framework is to maximize the energy efficiency rate of the UAV network by performing the joint optimization using UAV 3D trajectory, with the energy utilized during interference accounting, and connected user counts. Hence, an efficient Adaptive Deep Reinforcement Learning with Novel Loss Function (ADRL-NLF) framework is designed to provide a better energy efficiency rate to the UAV network. Moreover, the parameter of ADRL is tuned using the Hybrid Energy Valley and Hermit Crab (HEVHC) algorithm. Various experimental observations are performed to observe the effectualness rate of the recommended energy efficiency model for UAV-based networks over the classical energy efficiency framework in UAV Networks.

Keywords Unmanned aerial vehicles, Energy efficiency, Deep reinforcement learning, Novel loss function, Hybrid energy valley and hermit crab

The motivation for the proposed work is mainly due to the lack of comprehensive approaches in the existing systems, which were not multi-dimensional, were either focusing on agent optimization, distance optimization, energy optimization, focusing on node detection, distance minimization, or error control. There was no such method that dealt with both distance optimization, performance enhancement and energy optimization. The proposed method was inspired from¹⁻³, which were dealing all these factors separately. The proposed method uses HEVHC and ADRL-NLF integrated frameworks which reduces the complexity of both the HEV and HCO besides, improving the efficiency of the model, enhancing the reward, and reducing the penalty. The innovation

¹Department of CSE (AI&ML), S.A. Engineering College (Autonomous), Chennai, Tamil Nadu 600077, India. ²Department of CSE, School of Computing, Vel Tech Rangarajan Dr.Sagunthala R&D Institute of Science and Technology, Chennai, Tamil Nadu 600062, India. ³Department of Computational Intelligence, SRM Institute of Science and Technology, Chennai, Tamil Nadu 603203, India. ⁴Department of CSE, S.A. Engineering College (Autonomous), Chennai, Tamil Nadu 600077, India. ⁵Shiv Nadar (Institution of Eminence Deemed to Be University), Greater Noida, Uttar Pradesh 201314, India. ⁶Department of Computer Science, Kebri Dehar University, 250 Kebri Dehar, Ethiopia. ⁷School of Built Environment, Engineering and Computing, Leeds Beckett University, LS6 3QS Leeds, UK. ✉email: ShitharthS@kdu.edu.et

of the work is mainly understood in optimization of the framework with HEVHC and ADRL-NLF framework that not only increases the performance but also reduces the penalty, which was not addressed in the similar research works. The energy efficacy is the utilization of the low power to do the similar work or generate the same pattern of service. For instance, the energy-efficient constructions and homes utilize low power to the appliances⁴. The energy-efficient production facilities utilize less power to generate the products. The energy-efficient innovations can play an important part in minimizing the emissions and the consumption of the fossil fuels in the overall economic factors. For instance, the energy-effective “light emitting diode (LED)” light bulbs can generate a similar amount of light as the incandescent bulbs by utilizing less than 80% low electricity⁵. The families can save more amount of money on the bills by utilizing the energy-efficient appliances. The energy efficacy can be estimated by the section of the overall power input to the system or machines that are utilized in the essential task⁶. The energy efficacy can provide various kinds of merits, including minimizing the price on the economic range for the households, minimizing the requirement for the power imports, and minimizing the emissions of greenhouse gases⁷. The energy can be secured by deploying the “Compact Fluorescent Lamps (CFL)” lights, switching off the appliances, utilizing more daylight, handling the air leaks, and minimizing the temperature of the room when not in the utility⁸.

In the emergency communications, the UAV is a successive innovation due to its high mobility, low cost, and flexible deployment. But, because of the restricted power of the onboard battery, the UAV’s service period is highly restricted⁹. In the UAV-based networks, the energy efficacy is the rate of total efficient ability of the downlink candidates to the UAV’s power usage. The power usage of the UAV includes the power utilized by the interaction and the power utilized by the hovering¹⁰. In order to enhance the experience of the candidate, the UAVs must enhance their power efficacy by tuning their processing capacity in the restricted service period. The other difficulty in the UAV-aided networks is designing energy-based trajectories for UAVs¹¹. The UAVs can meet the trouble from the neighbourhood, when UAV access points or other cells distributing a similar frequency band. This trouble may affect the energy efficacy of the device¹². One UAV may be charged both by the charging stations and the solar energy. This can lead to sustainable interaction services, when preventing the power outage. In order to enhance the energy efficacy of the UAV-aided wireless networks, experts can concentrate on the trajectory planning of the UAV, UAV’s interaction on resource allocation, computing, UAV’s 3D hovering region decision, and caching¹³. The protocols that reduce the power usage may support resolving the issue of network lifespan. This is done by the two protocols such as Power Efficient Gathering in Sensor Information Systems (PEGASIS) and Low-Energy Adaptive Clustering Hierarchy (LEACH)⁴. The LEACH is the “Time-Division Multiple Access (TDMA)”-aided protocol that utilizes the clustering approach to enhance the “Wireless Sensor Networks (WSN)’s” lifespan. The LEACH is developed by adjusting the clustering task. The approach is enhanced by integrating the sink mobility to enhance the energy efficacy. The other protocol named PEGASIS that minimizes the amount of information to be sent to the base station. The protocol named PEGASIS works better concerning the lifespan of the network contrasted with the LEACH¹⁴.

In the UAV-aided networks, the energy efficacy may be efficiently processed by enhancing both the deep learning and the machine learning tasks in the UAV framework¹⁵. When considering the machine learning tasks to define the critical multi-UAV deployment issues. Most importantly, the “Multi-Agent Reinforcement Learning (MARL)” techniques have been developed in most of the tasks to tune the energy efficacy of the systems. The technique called “distributed Q-learning” concentrated on tuning the UAV’s energy usage without concentrating the energy efficacy of the systems¹⁶. In order to define this issue, the “Deep Reinforcement Learning (DRL)” model is suggested to tune the energy efficacy of the constantly winged UAVs, that moves circularly and can’t perform the hovering such as the “rotary-winged UAVs”¹⁷. The UAVs offer the wireless coverage to the mobile and the static places. But, the UAV’s power is constrained and faces trouble from the neighbourhood cells of UAVs¹⁸. A “double deep Q-network”-aided structure can enhance the energy efficacy and overall throughput of the network in the UAV-aided terrestrial frameworks¹⁹.

The contributions of the designed energy efficiency framework for the UAV-assisted network are presented below.

- The proposed work implements the energy efficient framework for the UAV-aided networks by employing adaptive deep learning technique that enriches the energy efficiency and minimizes the interference rates.
- Proposed work uses ADRL-NLF technique by utilizing the classical DRL network, that highly increases the energy efficiency rate and performs the rapid communication. Here, the HEVHC algorithm is supported to optimize the parameters of DRL.
- The proposed work uses HEVHC algorithm by influencing the traditional EVO and HCO algorithms, that assists to optimize the DRL network’s parameters.
- Proposed method optimizes the energy efficiency in multi agent environments with maximum reward and minimum penalty
- Proposed work provides optimization in distance detection, energy efficiency, and computational complexity reduction for ultrareliable and low latency communications.
- Proposed work enhances the stability of the ground station for multi agent UAV assisted networks.
- Proposed work investigates the presented mechanism’s effectiveness and robustness with numerous conventional approaches and with several performance metrics.

The structure of the implemented energy efficiency framework for the UAV-assisted networks is provided as follows. The classical works for the UAV-assisted networks are illustrated in Division II. Then, the system model and problem description for UAV-assisted networks with the description of the proposed model are presented in Division III. In addition, the implemented HEVHC algorithm for the presented network is explained in Division

IV. Also, the ADRL network with a novel loss function, utilizing the multi-objective function is demonstrated in Division V. The solutions and the explanations of the implemented work are given in Division VI. Lastly, the designed work's summary is given in Division VII.

Existing works

Related works

In 2023, Omoniwa et al.²⁰ have presented a direct collaborative mechanism called “Communication-enabled Multi-Agent Decentralized Double deep Q-Network (CMAD-DDQN)”. This was a collaborative approach that permits the UAV to distribute their telemetry explicitly through the conventional 3GPP guidelines by interacting with their neighbourhoods. This approach enhanced the energy efficacy of the system, without reducing the network's functionality gain. The presented work outperformed the other conventional baseline tasks.

In 2022, Omoniwa et al.²¹ have implemented a task for enhancing the energy efficacy of the system by jointly tuning every 3D trajectory of the UAV, power utilization, and the amount of linked users when considering the interference. Hence, experts recommended a Multi-Agent Decentralized Double deep Q-Network (MAD-DDQN) mechanism. This mechanism performed better than the conventional tasks concerning the energy efficacy.

In 2021, Hu et al.²² have suggested a framework to examine the functionality of the UAV-aided spectrum-sharing approach. Experts implemented the transmit power of the UAV, normalized sensing threshold, and the sensing period to enhance the energy efficacy under the factor that the significant candidate was enough secured. The simulation outcomes have resulted in the suggested approach converging to the overall optimal measures and the energy-efficacy functionality compared with the other conventional approaches.

In 2020, Jia et al.²³ have presented an approach to raise the energy efficacy, by integrating the “Spatial Modulation (SM)” and “Non-Orthogonal Multiple Access (NOMA)” techniques and named a new scheme called “Spatial NOMA (S-NOMA)”. In addition, an energy allocation optimization approach concerning the energy efficacy of the S-NOMA approach was suggested. Moreover, the simulation solutions displayed that the recommended S-NOMA performed better compared with the classical NOMA.

In 2022, Tang et al.²⁴ have deployed an efficient approach to enhance the energy efficacy of the model by concerning both the propulsion energy of the UAV and the transmission power of the IoT, where the transmission resources of the IoT and the trajectory of the UAV were jointly tuned. The extensive simulation outcomes were offered to corroborate the efficacy of the suggested approach.

In 2021, Chen et al.²⁵ have recommended an energy transmission device utilizing the UAV as a flying base station. This work employs the DRL approach, where the experts suggested a DQN mechanism-aided resource allocation to enhance the energy efficacy of the model. The numerical experiments displayed that the designed approach could highly enhance the energy efficacy of the system, contrasted with the other conventional resource allocation tasks.

In 2018, Sikeridis et al.²⁶ have discovered a mechanism that integrated the UAV assistance with the “Wireless Powered Communication (WPC)” approaches to enhance the energy efficacy. The IoT systems generate coalitions by selecting their part in the framework. At last, a “non-cooperative game-theoretic” task was recommended to estimate the transmission energy of every IoT node. The functionality estimation of the recommended work was attained and the outcomes demonstrated its scalability, robustness, and energy efficacy.

In 2021, Zhang et al.²⁷ have suggested energy effective path optimization mechanism for UAV-aided IoT frameworks. The “action-confined off-policy and on-policy” and the module-free Reinforced Learning (RL) tasks were suggested and jointly employed to resolve the optimization issue. Experts had estimated the efficacy of the recommended work by contrasting it with the other dynamic mechanisms. The simulation had displayed that the presented work performed better than the other traditional approaches.

Research gaps and challenges

The energy efficiency in UAV networks includes challenges such as, the communication between the networks getting affected due to the continuous movement of the UAV network. This network has a low battery life that limits the payload and range of the network. The energy efficiency in the network depends upon the size of the RIS used. Several researches have been performed by the researchers on consuming energy in the UAV networks. The merits and disadvantages of a few existing networks are presented in Table 1. DRL²⁰ is capable of processing large data. This method provides better robustness. Still, it does not provide a better convergence rate. DDQN²¹ improves the system's optimization capability. This method solves high-dimensional problems and this method is aware of the noise. But, this method has high implementation costs and requires huge storage space. RL²² obtains increased performance in energy efficiency. This method provides high robustness. Yet, it is time consuming and expensive. S-NOMA²³ provides enhanced performance in low transmission power. Still, it is time consuming. This method has a lower error detection rate. Dinkelbach²⁴ has the ability to solve large scale problems. But, it is difficult to execute in multiple-ratio problems. UDN²⁵ improves the spectrum efficiency of the network. This method utilizes less energy. This method is not preferred because, it is expensive. This method also has low network capacity. RL²⁶ utilizes less energy and provides maximum performance. But, this method uses large hyperparameters that increase the computational complexity and cause overload issues. “off-policy and on-policy” RL schemes²⁷ improve the data transmission ability of the system and provide stable outcomes. Yet, this method has low generalization ability and has low convergence rate. DRL based block-length optimization and power control was proposed¹ for UAV based IIoT network. This work needs to be enriched with energy optimization with ADRL for multi-agent systems. However, this work is the motivation for the proposed work. Low-complexity Coordinate Descent Approximation Algorithm (CDAA) for UAV positioning and sub carrier allocation was proposed in the work² for ultra reliable and low latency communications for

Author [citation]	Methodology	Features	Challenges
Omoniwa et al. ²⁰	DRL	<ul style="list-style-type: none"> This method is capable of processing large data This method provides better robustness 	<ul style="list-style-type: none"> It does not provide a better convergence rate
Omoniwa et al. ²¹	DDQN	<ul style="list-style-type: none"> This method improves the system's optimization capability This method solves high-dimensional problems This method is aware of the noise 	<ul style="list-style-type: none"> This work has more development costs This method requires huge storage space
Hu et al. ²²	RL	<ul style="list-style-type: none"> This method obtains increased performance in energy efficiency This method provides high robustness 	<ul style="list-style-type: none"> It is time consuming It is expensive
Jia et al. ²³	S-NOMA	<ul style="list-style-type: none"> This method provides enhanced performance by consuming low transmission power 	<ul style="list-style-type: none"> It is time consuming This method has a lower error detection rate
Tang et al. ²⁴	Dinkelbach method	<ul style="list-style-type: none"> This method has the ability to solve large scale problems 	<ul style="list-style-type: none"> It is difficult to execute in multiple-ratio problems
Chen et al. ²⁵	UDN	<ul style="list-style-type: none"> This method enhances the spectrum efficiency of the network This method utilizes less energy 	<ul style="list-style-type: none"> It is expensive This method has low network capacity
Sikeridis et al. ²⁶	RL	<ul style="list-style-type: none"> This method utilizes less energy This method offers maximum performance 	<ul style="list-style-type: none"> This method uses large hyperparameters that increase the computational complexity It causes overload issues
Zhang et al. ²⁷	Off-policy and on-policy RL schemes	<ul style="list-style-type: none"> This method improves the data transmission ability of the system This method provides stable outcomes 	<ul style="list-style-type: none"> This method has low generalization ability It has a low convergence rate
Ranjha, A. et al. ¹		<ul style="list-style-type: none"> Optimization of the block length allocation in IIoT networks of UAV Power control and allocation 	<ul style="list-style-type: none"> Uses Deep learning reinforcement model which is the inspiration for the proposed work Agent based optimization is used in this work. The proposed work uses optimization for the energy efficiency also This uses only DRL for distace optimization
Ranjha, A., et al. ²		<ul style="list-style-type: none"> Solar powered multi carrier low latency agents for UAV low-complexity coordinate descent approximation algorithm (CDAA) for UAV positioning and sub carrier allocation 	<ul style="list-style-type: none"> There is only position and sub carrier optimization and not addressing energy optimization The proposed work uses ADRL based energy optimization also
Ranjha, A., et al. ³		<ul style="list-style-type: none"> This work optimizes mission completion time subject to the constraints of resource allocation, including power, transmission CPU frequency usage, error rate decoding, block length estimation, communication bandwidth, and task partitioning as well as 3-D UAV positioning 	<ul style="list-style-type: none"> No measures for reward enhancement and the penalty reduction Proposed work uses ADRL-NLF adaptive model for the energy efficiency

Table 1. Features and challenges of traditional energy efficient techniques in UAV networks.

UAV networks. This work inspired the proposed work for the positioning aspects of the agents. Optimization of the resource allocation, including power, transmission CPU frequency usage, error rate decoding, block length estimation, communication bandwidth, and task partitioning as well as 3-D UAV positioning was proposed³ in the work for 6G networks. This work address power issues, task allocation, error handling, positioning and optimization of the agents. The proposed work is an enhancement of this work with ADRL with penalty reduction and reward increase.

So, it is necessary to develop an effective framework to resolve several complications in existing techniques and also to strengthen the energy efficiency rate in the UAV network using deep learning approaches.

Materials and Methods

System model

The suggested work considered a group of mobile²¹ and the static candidates ξ placed in a certain region. Every candidate $c \in \xi$ at a specific time s is placed in the coordinate (a_c^s, b_c^s) . This work considered the service unavailability from the traditional terrestrial framework because of the enhanced network load or the disasters. A group of quadrotors of the UAVs is placed in the region to offer wireless coverage to the candidates on the ground. An assisting UAV $u \in N$ at a specific time s , is placed in the coordinate (a_u^s, b_u^s, d_u^s) . Without the generality loss, it is assumed that a confirmed “line-of-sight (LOS)” channel situation because of the UAV’s aerial places. One of the factors of the signal quality is referred as the “Signal to Interferences plus Noise Ratio (SINR)”. It is referred as the rate of the power of the specific interest signals plus the power of noise. Every candidate $c \in \xi$ at a specific time s can be linked to the one UAV $u \in N$ that offers the powerful downlink SINR. Hence, the SINR for the specific time s is formulated in Eq. (1).

$$\gamma_{c,u}^s = \frac{\chi Q(e_{c,u}^t)^{-\delta}}{\sum_{x \in \text{int}} \chi Q(e_{c,u}^t)^{-\delta} + \sigma^2} \quad (1)$$

Here, the variables δ and χ are the “path loss exponent and the attenuation” factor that features the wireless channel correspondingly. At the receiver’s “additive white Gaussian noise power” is specified as σ^2 . The distance among the variables c and u at the corresponding time s , which is indicated as $e_{c,u}^t$. The interfering UAV’s set is

denoted as $\lambda^{\text{int}} \in N$. The interfering UAV's index in the group λ^{int} is denoted as z . The UAV's transmit power is given as Q . The proposed work developed the mobile candidate's mobility by employing the "Gauss Markov Mobility (GMM)" approach, that permits the candidates to dynamically vary their places. The UAVs should tune their flight paths to offer everlasting connectivity to the candidates. With the assumption that the receiving information rate of the ground candidates and the channel bandwidth W_B can be formulated utilizing the "Shannon's" expression represented in the Eq. (2).

$$\mathfrak{N}_{c,u}^s = W_B \log_2 (1 + \gamma_{c,u}^t) \tag{2}$$

In the interference-restricted model, the coverage is troubled by the SINR. Thus, this value is estimated in the UAV's connectivity score $u \in N$ at a particular time s in Eq. (3).

$$V_u^s = \sum_{c \in \xi} B_u^s(c) \tag{3}$$

In this, the factor $B_u^s(c) \in [0, 1]$ indicates, whether the candidate c is linked to the UAV u for a particular time s . $B_u^s(c) = 1$ if $\gamma_c^s = \gamma_{c,u}^s > \gamma^{\text{th}}$, or else $B_u^s(c) = 0$, here the variable γ^{th} is already defined the threshold of the SINR. Similarly, the variable $\mathfrak{N}_{c,u}^s = 0$ if the candidate c is not linked to the UAV u .

In the flight tasks, the UAV $u \in N$ at the corresponding time s consumes the power f_{su} . The overall energy usage of the UAV f_T is defined as the total in the communication f_C and the propulsion f_P energies $f_T = f_P + f_C$. Since the variable f_C is much smaller than the factor f_P , the model avoided the variable f_C . Equation (4) offers an enclosed form analytical propulsion energy usage approach for the "rotary wing UAV" at a specific time s .

$$Q(s) = \kappa_0 \left(1 + \frac{3H^2}{D_{\text{tip}}^2} \right) + \kappa_c \left(\sqrt{1 + \frac{H^4}{2h_0^4} + \frac{H^2}{2h_0^2}} \right)^{\frac{1}{2}} + \frac{\rho}{2} htBH^3 \tag{4}$$

Here, the variables κ_0 and κ_c are the flight constants of the UAV. The tip speed of the rotor blade is denoted as D_{tip} and the velocity of the mean hovering is pointed as h_0 . The rotor solidity and the drag ratio are indicated as t and h accordingly. The region of the rotor disc is pointed as B for the specific time s and the density of the air is pointed as ρ . Finally, formulated the mean propulsion power against the overall time sequences as $\frac{1}{S} \sum_{s=1}^S Q(s)$, and the overall power utilized by the UAV u for the specific time s is expressed in Eq. (5).

$$f_u^s = \alpha_s \cdot Q(s) \tag{5}$$

Here, the variable α_s measured in every time sequence period. The UAV's u energy efficiency is explained as the rate of information throughput and the power utilized in the specific time period s as shown in Eq. (6).

$$\eta_u^s = \frac{\sum_{c \in \xi} \mathfrak{N}_{c,u}^s}{f_u^s} \tag{6}$$

Thus, the system model of the suggested work has been mathematically given.

Problem description

The aim of the proposed task is to strengthen the overall energy efficiency of the system by jointly tuning its amount of linked candidates, 3D trajectory, and the power utilized by the UAVs assisting the ground candidates²¹, within the restricted power budget. Enhancing the amount of linked candidates L_u^s raises the overall amount of information $\sum_{c \in \xi} \mathfrak{N}_{c,u}^s$. The UAV u send the signal in the corresponding time s for a specific amount of utilized energy f_u^s , which also enhances the overall energy efficiency η_{tot} . Thus, the issue of optimization is expressed in the below expressions.

$$\forall u \in N : \max_{y_u^s, z_u^s, x_u^s, f_u^s, L_u^s} \eta_{\text{tot}} = \frac{\sum_{s=1}^S \sum_{u \in N} \sum_{c \in \xi} \mathfrak{N}_{c,u}^s}{\sum_{s=1}^S \sum_{u \in N} f_u^s} \tag{7}$$

$$s.t. \gamma_{c,u}^s \geq \gamma^{\text{th}}, \forall B_u^s(c) \in [0, 1], c, u, s, \tag{8}$$

$$f_u^s \leq f_{\text{max}}, \forall u, s, \tag{9}$$

$$y_{\text{min}} \leq y_u^s \leq y_{\text{max}}, \forall u, s, \tag{10}$$

$$z_{\text{min}} \leq z_u^s \leq z_{\text{max}}, \forall u, s, \tag{11}$$

$$x_{\min} \leq x_u^s \leq x_{\max}, \forall u, s, \quad (12)$$

Here, the UAV's maximum energy level is denoted as f_{\max} . Further, the variables $y_{\max}, z_{\max}, x_{\max}$ and $y_{\min}, z_{\min}, x_{\min}$ are the maximum and the minimum 3D coordinates of $y, z,$ and x accordingly. Various wireless transmitters sending the similar frequency bands are in the enclosed presence to one another, resulting in the interference likelihood of the issue, Eq. (7) which is referred as the "NP-complete". This issue Eq. (7) is a non-convex problem, hence having numerous optimums. For this concern, rectifying Eq. (7) with traditional optimization tasks is critical. Most importantly, the issue Eq. (7) will continue very critical as various UAVs are installed in the distributed wireless platform. Thus it is complex to locate the optimal cooperative mechanisms to strengthen the energy efficiency of the model fulfilling the coverage works in the dynamic actions. This is frequently due to the UAVs becoming selfish and attaining the aim of enhancing their separate energy efficiency, while reducing the transmission outage and the power utilization, opposing the collective aim of enhancing the energy efficacy of the model. In the certain cases, the cooperative techniques may be applicable when collective and the separate UAVs trouble the required interests. The DRL works well in the decision-creating approaches in various platforms. Thus, the suggested work utilized the adaptive DRL technique to resolve the energy efficiency optimization issue of the system.

Description of proposed framework

The installation of UAVs to offer the wireless coverage to the ground candidates has attained important experimental attention. The UAVs play an important part in assisting the IoT by offering the connectivity to more systems, either mobile or static. Significantly, the UAVs have multiple real-world developments, varying from helping the transmission in the disaster-caused regions to the rescue, search, and for the surveillance tasks. Most importantly, the UAVs can be installed in the situations of conventional terrestrial framework downtime or the congestion of the network. However, to offer omnipresent tasks to the various ground candidates, the UAV demands robust mechanisms to tune their flight paths, when offering the coverage. There have been important experiments on optimizing the energy efficiency in the UAV frameworks. But, the traditional approaches depend on the middle ground controller for the decision-making of the UAV, by making it impractical for installing in the critical situations, because of the important amount of transmitted data between the controller and the UAV. In addition, it may be complex to locate the user places in certain occurrences. Machine learning is highly supportive for defining the complex UAV installation issues. Modern developments concentrate on tuning the energy efficacy of the systems by optimally determining the UAV's trajectory only on the static ground candidates and ignoring the mobile candidates. Some other approaches avoid the intrusion impact from the neighbourhood cells, considering an intrusion-free network module. In addition, several tasks consider the ground candidate's overall spatial knowledge region through the controller that simultaneously scans the perimeter of the network and offers the real-time updates to the UAVs for making the decisions. But, this consideration may be inappropriate in the emergency situations, since it demands important data transmission among the controller and the UAV. In addition, it is not possible to locate the regions of the candidates in the emergency situations. Thus, it is significant to develop the framework by considering these problems. Figure 1 depicts the implemented energy efficiency framework for the UAV-aided networks.

The presented work focuses on implementing an innovative energy efficacy approach in the UAV-aided framework, utilizing the DRL. The optimization of energy efficiency in the UAV presents good wireless coverage to the mobile and the static ground candidates. Nowadays, the RL tasks optimize the rate of energy efficacy of

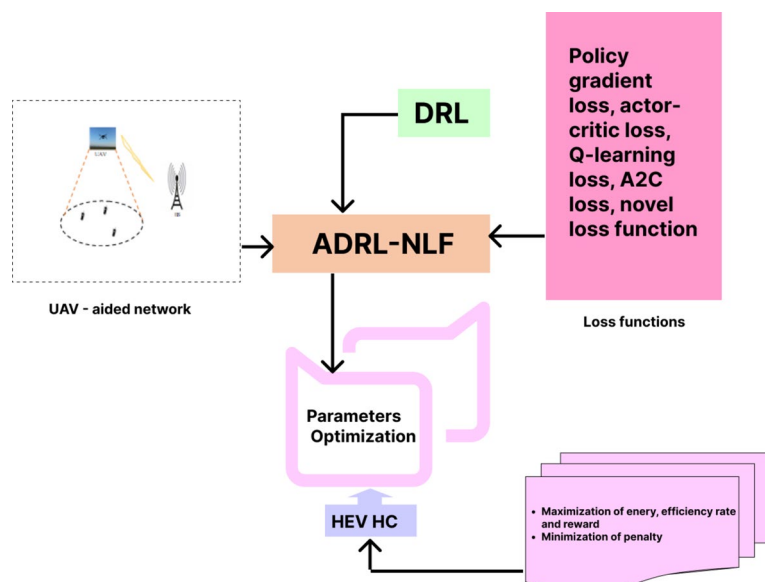


Fig. 1. The representation of the suggested energy efficiency framework for the UAV-assisted networks.

the model by utilizing the 2D trajectory strategy that efficiently prevents the rate of interference obtained in the neighbourhood of the UAV cells. The important theme of the suggested task is to strengthen the energy efficiency rate of the UAV framework by processing the joint optimization in the overall 3D trajectory of the UAV, energy employed during the intrusion accounting, and linking the candidate accounts. Thus, an effective ADRL-NLF approach is implemented to offer a better energy efficacy rate for the UAV structure. In addition, the attributes of the ADRL are optimally utilizing the HEVHC algorithm. Besides, diverse research is conducted to monitor the efficacy rate of the suggested energy efficiency approach of the UAV-aided network against the conventional energy efficacy approach in the UAV framework.

The proposed algorithm and the workflow is presented in the Fig. 2. It contains three vertices in which the functioning of the proposed algorithm is illustrated. The right most hybrid model of HCO-EVO algorithm solves the operational delay, computational complexity and optimization problems, this hybrid model works on the determination of particle stability, distance data about the candidates on ground station, the minimum and maximum number of iterations, and executes the social search operator of the HCO algorithm for identifying the best particle positions. This ADRL is integrated with the HEVHC algorithm for the reduction of the loss and error rate. These hyper parameters are additionally achieved through the integration of the ADRL with the HEVHC algorithms. This integration reduces the penalty and increases the reward in the estimation of the best agents, by reducing the loss, error, search time, increasing the energy efficiency, better groundstation control, reduction of the error rates, loss, optimization of the search process, and power consumption. Thus the proposed HEVHC algorithm has loss function accomplishment with optimization enhancement using the hybrid modeling of the EVO-HCO.

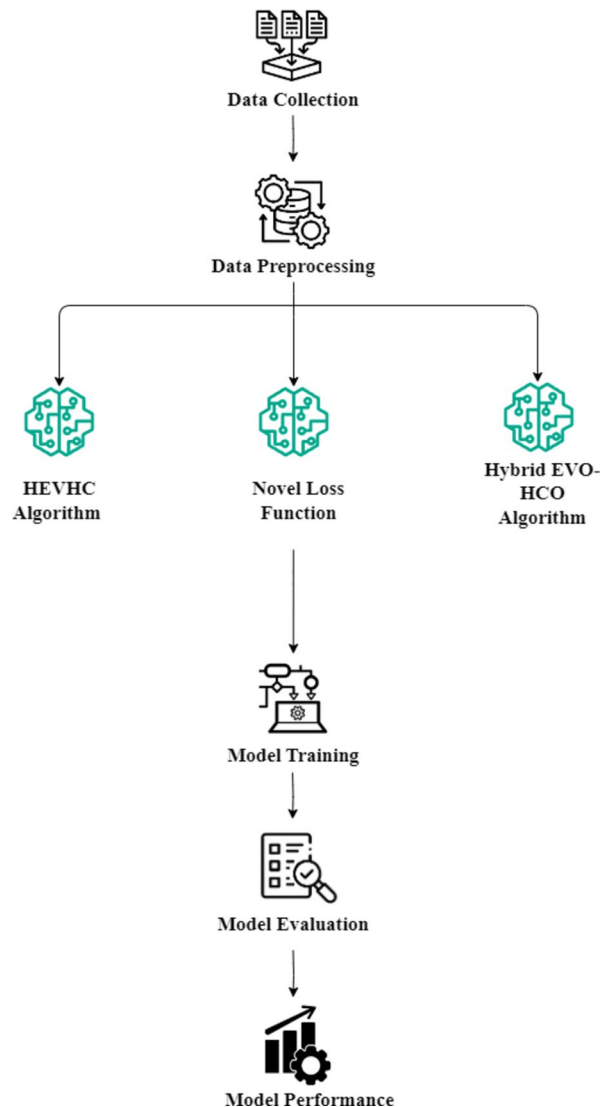


Fig. 2. Workflow of the proposed algorithm and process sequencing.

Hybrid energy valley and hermit crab for energy-efficient UAV-assisted network-Conventional EVO

The EVO²⁸ is motivated by the modern principles concerning the stability and diverse sections of particle collision. The formulation of the EVO is described here.

In the starting stage, the initialization approach is performed in that the solution members S_i , which are considered to be the particles with diverse stages of the stability in the search place that is considered to be a particular section of the universe.

$$S_i = \begin{bmatrix} S_1 \\ S_2 \\ \vdots \\ S_i \\ \vdots \\ S_m \end{bmatrix} = \begin{bmatrix} s_1^1 & s_1^2 & \dots & s_1^j & \dots & s_1^x \\ s_2^1 & s_2^2 & \dots & s_2^j & \dots & s_2^x \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ s_i^1 & s_i^1 & \dots & s_i^j & \dots & s_i^x \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ s_m^1 & s_m^2 & \dots & s_m^j & \dots & s_m^x \end{bmatrix} \quad (13)$$

$$\left. \begin{matrix} i = 1, 2, \dots, m \\ j = 1, 2, \dots, x \end{matrix} \right\}$$

$$s_i^j = s_{i,\min}^j + rdm \cdot (s_{i,\max}^j - s_{i,\min}^j), \left\{ \begin{matrix} i = 1, 2, \dots, m \\ j = 1, 2, \dots, x \end{matrix} \right. \quad (14)$$

Here, the particle’s overall count is denoted as m , and the specified issue’s dimension is indicated as x . The variable s_i^j specifies the j^{th} decision factor for estimating the starting place of the i^{th} member. The j^{th} factor’s “lower and upper” boundaries are considered as $s_{i,\min}^j$ and $s_{i,\max}^j$. The distributed arbitrary factor is given as rdm and it is ranged from 0 to 1.

In the next phase, the particle’s “Enrichment Bound (EB)” is estimated and is employed for taking the variations among the “neutron-poor and neutron-rich” particles. The objective measure calculation for all the particles is carried out and estimated as the particle’s “Neutron Enrichment Level (NEL)”. This is presented in Eq. (15).

$$eb = \frac{\sum_{i=1}^m nel_i}{m}, i = 1, 2, \dots, m \quad (15)$$

Here, the variable nel_i is the particle’s NEL and the factor eb is the particle’s EB.

In the next stage, the particle’s stability phases are established according to the objective function and that is expressed in Eq. (16).

$$sl_i = \frac{nel_i - bs}{ws - bs}, i = 1, 2, \dots, m \quad (16)$$

Here, the i^{th} particle’s stability level is denoted as sl_i . The “worst and best” stability stages of the particle are specified as ws and bs accordingly. In the solution member, the emitted rays are taken as the decision factors, that are substituted and removed by the rays presented in the member or the particle with the best stability stage S_{bs} . This is shown in Eq. (17).

$$S_i^{new1} = S_i(S_{bs}(s_i^j)), \left\{ \begin{matrix} i = 1, 2, \dots, m \\ j = Alpha\ index\ II \end{matrix} \right. \quad (17)$$

Here, the currently produced universe particle is indicated as S_i^{new1} , and the present place vector of the i^{th} particle is pointed as S_i . The variable $Alpha\ index\ I$ is the arbitrary integer with the limit of $[1, x]$. Then, the attribute $Alpha\ index\ II$ explains which rays to be discharged with the limit of $[1, Alpha\ index\ I]$.

The total distance among the specified particle and the other is determined in Eq. (18).

$$X_i^k = \sqrt{(s_2 - s_1)^2 + (r_2 - r_1)^2}, \left\{ \begin{matrix} i = 1, 2, \dots, m \\ k = 1, 2, \dots, m - 1 \end{matrix} \right. \quad (18)$$

Here, the factor X_i^k is the overall length of i^{th} the particle and the k^{th} nearby particles. Then, the search place coordinates are pointed as (s_1, r_1) and (s_2, r_2) .

Utilizing these situations, the place upgrading approach for creating the second solution member is formulated in Eq. (19).

$$S_i^{new2} = S_i(S_{ng}(s_i^j)), \left\{ \begin{matrix} i = 1, 2, \dots, m \\ j = Gamma\ index\ II \end{matrix} \right. \quad (19)$$

Here, the currently produced particle is specified as S_i^{new2} . The nearby particle of the i^{th} particle is pointed as S_{ng} .

The approach copies the tendency of the particles to attain the stability range. Most of the aware particles are placed near the boundary, and several of them have better ranges of stability. This is formulated in Eq. (20) and Eq. (21).

$$S_{cp} = \frac{\sum_{i=1}^m S_i}{m}, \quad i = 1, 2, \dots, m \tag{20}$$

$$S_i^{new1} = S_i + \frac{(a_1 \times S_{bs} - a_2 \times S_{cp})}{sl_i}, \quad i = 1, 2, \dots, m \tag{21}$$

Here, S_i^{new1} and S_i are future and present place vectors of the i^{th} particles. The variables a_1 and a_2 are the arbitrary integers in the limit of $[0, 1]$ that estimates the number of particle's motion. The place vector of the middle of the particles is denoted as S_{cp} .

To enhance the exploration and the exploitation stages of the approach, another place updation is performed as shown in Eq. (22).

$$S_i^{new2} = S_i + \frac{(a_3 \times S_{bs} - a_4 \times S_{cng})}{sl_i}, \quad i = 1, 2, \dots, m \tag{22}$$

Here, the future and present place vectors of the i^{th} particles are denoted as S_i^{new2} and S_i . The variables a_3 and a_4 , which are the arbitrary integers, with the limit of $[0, 1]$, estimates the number of particle's motion. An arbitrary motion in the search place is estimated for taking these movement's sorts are given in Eq. (23).

$$S_i^{new} = S_i + a, \quad i = 1, 2, \dots, m \tag{23}$$

The variable a is the arbitrary integer in the limit of $[0, 1]$, estimates the number of particle's motion. Algorithm 1 offers the conventional EVO's pseudo-code.

```

Represent the iteration parameters and the population matrix
Formulate the objective function
For t=1 to T_max
    Determine the stability phases of the particles utilizing Eq. (17)
    Estimate the distance among the certain particles employing Eq. (18)
    Upgrade the new positions by employing Eq. (21) and Eq. (22)
End
Return the better solutions
    
```

Algorithm 1. Conventional EVO

Conventional HCO

The classical HCO²⁹ algorithm is motivated by the hermit crabs that normally live in herms or groups of hundred or more. The model of the HCO is mathematically given.

Representation

The search agent's population matrix Q in every execution presents and is upgraded or subjected to the upcoming creation. The matrix Q is generated in Eq. (24).

$$Q(i) = \begin{bmatrix} s_1^i & \dots & s_d^i \\ \vdots & \ddots & \vdots \\ s_1^m & \dots & s_X^m \end{bmatrix}_{M \times X} = \begin{bmatrix} \vec{S}_1(i) \\ \vdots \\ \vec{S}_M(i) \end{bmatrix}_{M \times 1} \tag{24}$$

Here, the overall amount of the search agents are denoted as M , and the overall encoded amount is indicated as X . Then, the variable s represents the region of the search member in the dimension x and the search member's feature vector is given as \vec{S} .

Initialization

The search member's starting population is created, which is arbitrarily utilizing the programming language's "pseudo-random number generator". This results in the HCO, presented in the Eq. (25).

$$s_x^{m0} = s_x^- + \alpha (s_x^+ - s_x^-), \quad 1 \leq x \leq X, 1 \leq m \leq M \quad (25)$$

Here, the factors s_x^+ and s_x^- stand the "upper and lower" boundaries of every search agent. Further, the "pseudo-random" integer is denoted as α with the limit [0, 1].

Pre-processing

Every search member with the attributes troubling the max/min boundaries is adjusted by allocating their value similar to their nearby boundary. Three controlling attributes called distraction indicator d , shell availability's perceived risk c , and the overall shells b , which are set as zero.

Evaluation

Every search member is related to the value and characterized by the vector \vec{S} , is mapped into the answer space. In this stage, the population matrix Q is combined with a new attribute denoting the objective values obtained by the search members, and the same is shown in Eq. (26).

$$\prod(i) = [Q(i), Z(i)] = \begin{bmatrix} \vec{S}_1(i) & f\left(\vec{S}_1(i)\right) \\ \vdots & \vdots \\ \vec{S}_M(i) & f\left(\vec{S}_M(i)\right) \end{bmatrix} \quad (26)$$

Here, the decoder function is denoted as f , and the state-reward matrix is indicated as \prod . Then, the reward matrix is pointed as Z .

Comparison

Here, the rewards obtained during the new state estimation of the agents are contrasted with the historical rewards in the previous stage.

Post-processing

Here, the shell availability's perceived risk is estimated in Eq. (27).

$$c = 1 - \alpha/M \quad (27)$$

Solitary search operator

This phase is only utilized for those search members with the successful label. The solitary search member is formulated in Eq. (28).

$$\vec{S}_m(i+1) = \frac{1}{G} \sum_{g=1}^G \left(\vec{S}_g'(i) + \gamma_g \left| \delta_1 \vec{S}_g'(i) - \vec{S}_m(i) \right| \right) \quad (28)$$

In this, the variables $\vec{S}_m(i+1)$ and $\vec{S}_m(i)$ denote the successful search member in two simultaneous executions. The elitism integer's G perceived place is denoted as $\vec{S}_g'(i)$. In order to estimate $\vec{S}_g'(i)$, Eq. (29) is employed.

$$\vec{S}_g'(i) = \vec{S}_g(i) + \beta |A_g(i) L_{1 \times X}, 1 \leq g \leq G \quad (29)$$

Here, the utmost distance is denoted as $A_g(i)$. The arbitrary integer is specified as β with the limit [-1, 1]. Furthermore, the unit matrix is indicated as $L_{1 \times X}$.

The utmost travelable length of the molecule in every direction is estimated in Eq. (30).

$$A_g(i) = \sqrt{2X_g(i)\Delta t}, \quad 1 \leq g \leq G \quad (30)$$

Here, the factors $X_g(i)$ and Δt are measured in Eq. (31) and Eq. (32).

$$X_g(i) = \frac{\left| f\left(\vec{S}_g(i)\right) - f\left(\vec{S}_m(i)\right) \right|}{\left| f + f\left(\vec{S}_m(i)\right) - f^-\left(\vec{S}_m(i)\right) \right|}, 1 \leq g \leq G \tag{31}$$

And: $\Delta t = i/I$ (32).

The rate of distraction for every hermit to forward the Odor’s dominant source is estimated in Eq. (33).

$$\gamma_g = (1 - i/I)\beta(2c - 1)(1 - \rho_g), 1 \leq g \leq G \tag{33}$$

Here,

$$\rho_g(i) = \frac{\left| f\left(\vec{S}_G(i)\right) - f\left(\vec{S}_g(i)\right) \right|}{\left| f + f\left(\vec{S}_m(i)\right) - f^-\left(\vec{S}_m(i)\right) \right|}, 1 \leq g \leq G \tag{34}$$

At last, the variable γ_1 displays the wind efficiency in the habitat by Eq. (35).

$$\gamma_1 = d + 2\alpha(1 - d) \tag{35}$$

After utilizing the dominant smell source’s places, the places are upgraded utilizing Eq. (36) and Eq. (37).

If $\alpha_3 < 0.5$:

$$\vec{S}_g(i + 1) = \vec{S}_g(i) + \alpha_1(\vec{S}^-(i) + \alpha_2(\vec{S}^+(i) - \vec{S}^-(i))) \tag{36}$$

else :

$$\vec{S}_g(i + 1) = \vec{S}_g(i) - \alpha_1(\vec{S}^-(i) + \alpha_2(\vec{S}^+(i) - \vec{S}^-(i))) \tag{37}$$

In this, the overall α variables are considered as arbitrary integers in the limit [0, 1].

Social search operator

This phase applies to those search members with the label called fail. Here, the transition matrix indicates the likelihood of hermit crab shell exchange. This state is expressed in Eq. (38).

$$V = \begin{bmatrix} I & P \\ Z & W \end{bmatrix} \tag{38}$$

In this, the variable I is the 2×2 matrix, which relates with absorbing the places of the largest and smallest hermit crabs in the chain of vacancy. Further, the factor W is the $(Y - 2) \times (Y - 2)$ matrix, which relates to the exchanges of shells among the hermit crabs in the chain of vacancy. Then, the variable P is referred as the zero matrix, and the transition matrix is indicated as Z . For example, the transition matrix for the hermit crabs in the chain of vacancy is formulated in Eq. (39).

$$V = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ w & 0 & 0 & q & 0 \\ 0 & 0 & w & 0 & q \\ 0 & p & 0 & w & 0 \end{bmatrix} \tag{39}$$

This is the same as the “Markov chain” approach.

In the end, to copy the stochastic shell replacement task, the social search member for every hermit crab in the chain of vacancy is given in Eq. (40).

$$\vec{S}_m(i + 1) = \left(\vec{S}_m(i) + \beta \left| \delta_2 \vec{S}_g(i) - \vec{S}_m(i) \right| \right) \tag{40}$$

Here,

$$\vec{S}_m'(i) = q \vec{S}_{m+1}(i) + w \vec{S}_{m-1}(i)(i) \quad (41)$$

and

$$\delta_2 = d + 2\alpha(1 - d) \quad (42)$$

Further, the largest and smallest search member's update is presented in Eq. (43) and Eq. (44).

$$\vec{S}_i(i+1) = \begin{cases} \text{Reinitialize,} & \alpha_2 < 0.5 \\ (1 - (2c - 1))(1 + \beta)\delta_2 \vec{S}_Y(i), & \alpha_2 \geq 0.5 \end{cases} \quad (43)$$

and

$$\vec{S}_Y(i+1) = \begin{cases} \vec{S}_Y(i) + \beta L_{1 \times X}, & \alpha_3 < 0.5 \\ \vec{S}_G(i) - \beta \left| \vec{S}_G(i) - \vec{S}_Y(i) \right|, & \alpha_3 \geq 0.5 \end{cases} \quad (44)$$

The traditional HCO's pseudo-code is displayed in Algorithm 2.

<p>Represent the iteration parameters and the population matrix</p> <p>Formulate the objective function</p> <p>For $t=1$ to T_{\max}</p> <p style="padding-left: 20px;">Perform the pre-processing and the post-processing tasks</p> <p style="padding-left: 20px;">Execute the solitary search operator phase by utilizing Eq. (28)</p> <p style="padding-left: 20px;">Execute the social search operator phase by applying Eq. (43) and Eq. (44)</p> <p style="padding-left: 20px;">Upgrade the current positions</p> <p>End</p>
--

Algorithm 2. Traditional HCO

Recommended HEVHC for parameter tuning

With the conventional EVO and HCO, the recommended HEVHC algorithm is implemented. The traditional EVO is motivated by the physics principles based on the distinct parts of particle degradation. This algorithm outranks the other optimization algorithms, when analyzing the unconstrained mathematical functions. It offers better outcomes with the little rank means and converges to the overall best answer. Moreover, the classical HCO utilizes the hermit crab's swarm intelligence. This algorithm assists the search members privately and performs similar to the RL approach. In this method, the failed and the successful agents are handled distinctly. It only needs a small amount of parameters, to decide the failed and successful agents, and it is very robust. However, the traditional EVO is not producing the better solutions for the complex optimization issues. Also, the conventional HCO has more computational burdens. In order to resolve these issues a new HEVHC algorithm is presented. This algorithm utilizes the both conventional EVO and HCO approaches and resolves the issues presented in the

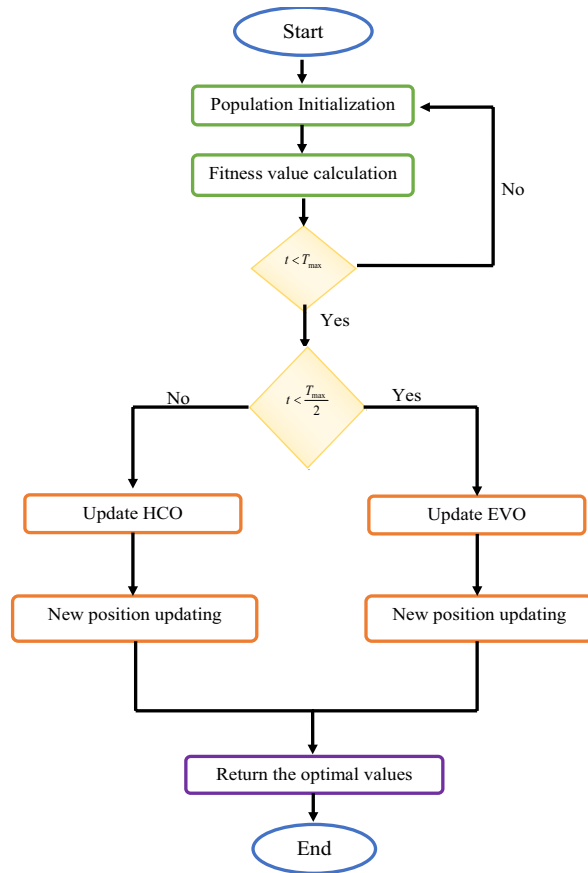


Fig. 3. The flowchart of the recommended HEVEC for parameter tuning.

both approaches. The suggested HEVHC algorithm performs based on the iteration counts. The formulation of the suggested HEVHC algorithm is shown in Eq. (45).

$$\begin{aligned}
 & \text{if } t < \frac{T_{\max}}{2} \\
 & \quad \text{Update via EVO} \\
 & \text{else} \\
 & \quad \text{Update via HCO}
 \end{aligned} \tag{45}$$

If the present iteration value is less than the half of maximum iteration, then the position update on is performed, utilizing the conventional EVO, or by employing the traditional HCO. Here, the factors t and T_{\max} stand the present and the maximum iterations.

The pseudo-code of the presented HEVHC is given in Algorithm 3. Figure 3 presents the suggested HEVHC algorithm for parameter tuning.

```

Represent the iteration parameters and the population matrix
Formulate the objective function
For  $t=1$  to  $T_{\max}$ 
  For  $j=1$  to  $N_p$ 
    If  $t < \frac{T_{\max}}{2}$ 
      Perform the EVO
      Determine the stability phases of the particles utilizing Eq. (17)
      Estimate the distance among the certain particles employing Eq. (18)
      Upgrade the new positions by employing Eq. (21) and Eq. (22)
    Else
      Perform the HCO
      Perform the pre-processing and the post-processing tasks
      Execute the solitary search operator phase by utilizing Eq. (28)
      Execute the social search operator phase by applying Eq. (43) and Eq. (44)
      Upgrade the current positions
    End If
  Find out the best solutions
End
End
Do the repetitive executions
Offer the optimal solutions

```

Algorithm 3. Implemented HEVHC

Adaptive deep reinforcement learning with novel loss function using multi-objective function-Deep reinforcement learning

In the multi-UAV platform, it is complex to attain a high amount of information samples for the training, where the task of training is computation-intensive and time consuming. In order to solve these issues the DRL³⁰ offers an efficient answer. The DRL is the method, where experts have issues with numerous amounts of possible answers or ways to manage those issues. The aim is to choose an action from the existing options, with the goal of attaining the optimal outcome. The objective of the DRL is to generate a connection among the actions and situations to attain the better solutions. The DRL employs the “Deep Neural Network (DNN)”, to learn the world with the capacity to perform on that learning. The variation of DRL from the other machine learning methods, is that the network does not explain the outcome, rather the network understands from the experience. In this DRL, there is no human intervention applied, and it is a closed-loop approach. That means the result of the one step generates the fundamental input to the upcoming stage. The DRL locates the development in issues, where the overall issue is distributed over various steps or stages and linear decision making is required. The “Markov Decision Process (MDP)” is employed to design the RL issues. The MDP is normally explained as four-tuples (T, B, γ, g) where,

T is the group of the overall environment stages, the variable $t_u \in T$ denotes the agent state at the corresponding time u .

B is the group of processing agent’s actions, the factor $b_u \in B$ is the action considered by the agent at a certain time u

The term $\gamma : T \times B \rightarrow H$ is referred to as the reward function here the term $\rho \sim \gamma(t, b)$, denotes the instant reward value attained by the agent processing action b_u at a specific time t_u .

The factor $g : T \times B \times H \rightarrow [0, 1]$ is the probability distribution function of the state transition, the attribute $t_{u+1} \sim (t_u, b_u)$ denotes the likelihood that the agent does the action b_u , from the transition to the upcoming stages t_{u+1} in the phase t_u .

In the DRL, the strategy $\pi : T \rightarrow H$, is the link of the stage space to the action place. It is formulated as the agent choosing action b_u in the phase t_u conducting the action.

Consider the instant reward for every time sequence in the future should be multiplied with the support of the distance attribute ρ then from the time u to the end time U , the total reward is formulated in Eq. (46).

$$R_u = \sum_{u'=u}^U \gamma^{u'-u} \gamma u' \tag{46}$$

Here, the variable $\gamma \in [0, 1]$, is employed to weigh the future reward impacts. The cumulative return attained by the agent is formulated in Eq. (47).

$$P^\pi(t, b) = F[R_u | t_u = t, b_u = b, \pi] \tag{47}$$

Here, the state action value function is denoted as $P(t, b)$ and the variable π^* specifies the optimal policy. There can be more than one optimal mechanism, yet they distribute a state action value function, as per the Eq. (48).

$$P^*(t, b) = \max_{\pi} F[R_u | t_u = t, b_u = b, \pi] \tag{48}$$

This is referred to as the “optimal state-action value function” that forwards the “Bellman optimal expression” as shown in Eq. (49).

$$P^*(t, b) = F_{t'-T} \left[s + \gamma \max_{b'} P(t', b') | t, b \right] \tag{49}$$

In the conventional RL, the function of Q-value is normally resolved by the iterative “Bellman” expression as given in Eq. (50).

$$P_{c+1}(t, b) = F_{t'-T} \left[s + \gamma \max_{b'} P(t', b') | t, b \right] \tag{50}$$

In that if the variable is $c \rightarrow \infty$ then $P_c \rightarrow P^*$. The structure of the DRL is depicted in Fig. 4.

Adaptive DRL with novel loss function

The DRL resolves the complex issues and handles the uncertainty. However, the involvement of hidden neurons and the epochs is high. This results in slow processing and overfitting. Hence, optimization of these attributes is necessary. Moreover, the overall iteration of this network needs to be optimized to obtain the rapid outcomes. For this purpose, the HEVHC algorithm is recommended in the presented work.

Moreover, by varying the loss functions such as “Advantage Actor-Critic (A2C) loss, policy gradient loss, Q-learning loss, novel loss function, and actor-critic loss”, the network is processed.

A2C loss: It is supported to decrease the variance by employing a critical approach of the policy gradient.

Policy gradient loss: It is an amount of loss, that specifies when everything has been executed.

Q-learning loss: It specifies the variation among the expectation and the observation. It contrasts the Q-target and the Q-value prediction.

Actor-critic loss: It is the fault caused by the agent in forecasting the present state value.

Novel loss function: In this operation, the policy gradient loss and Q-learning loss are added with the help of HEVHC to regularize the Q-values and the policy gradient.

Hence, the ADRL network with novel loss function is constructed with the support of implemented HEVHC. In this network, the input attributes such as node names, topology, distance matrix, node position (x, y), and the traversal between nodes are given to produce the optimized network.

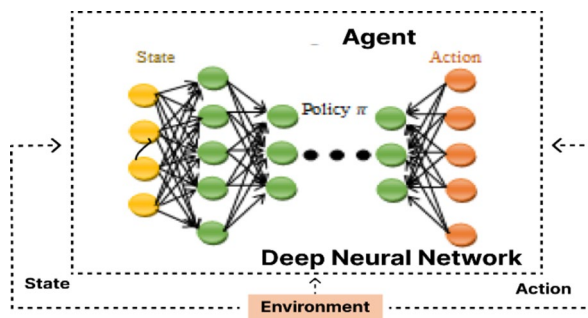


Fig. 4. The structure of the DRL.

The framework of the ADRL-NLF is presented in Fig. 5.

Objective formulation and its definition

The suggested HEVHC is employed to tune the parameters presented in the DRL. The parameter optimization for the task is given in Eq. (51).

$$ob = \arg \max_{\{hn^{DRL}, ep^{DRL}, it^{DRL}\}} \left[eer + re + \frac{1}{pe} \right] \tag{51}$$

Here, the variable hn^{DRL} specifies the DRL model's hidden neuron count, and the factor ep^{DRL} denotes the DRL model's epoch size. Then, the term it^{DRL} stands for the DRL model's number of iterations. The hidden neuron count of DRL is varied from 5 to 255 and the DRL model's epoch size is varied from 5 to 50. In addition, the number of iterations for this network is changed from 100 to 10,000. Furthermore, with the aid of the recommended HEVEC, the energy-efficient rate eer and the rewards re are maximized. Then, the penalty pe is minimized by the designed HEVEC. These attributes are explained as follows.

Energy efficient Rate eer : "It is referred as the ratio of overall transmitted data count to the transmission energy's weighted sum".

Rewards re : The goal of the agent is to understand the policy that enhances the energy efficiency of the system and reduces the UAV's energy consumption and outage of the users. Thus, the reward formulation of every agent is introduced in this work for every time step.

Penalty pe : It is the calculation on the present agent, to estimate how far it is from the target monitoring node of the UAV.

Figure 6 provides the solution encoding diagram of the presented task.

Results and discussions

Experimental setup

The proposed work was executed in the "MATLAB 2020a" platform and produced the optimal solutions. The implemented algorithm has 250 iterations and 10 populations. Also, the task had a chromosome length of 3. Moreover, the presented work is analyzed with diverse traditional algorithms such as "Egret Swarm Optimization Algorithm (ESOA)³¹, Coronavirus Mask Protection Algorithm (CMPA)³², Energy Valley Optimizer (EVO)¹⁷, and Hermit Crab Optimizer (HCO)¹⁸" to prove its effectiveness. The experimental setup of the suggested work is presented in Table 2.

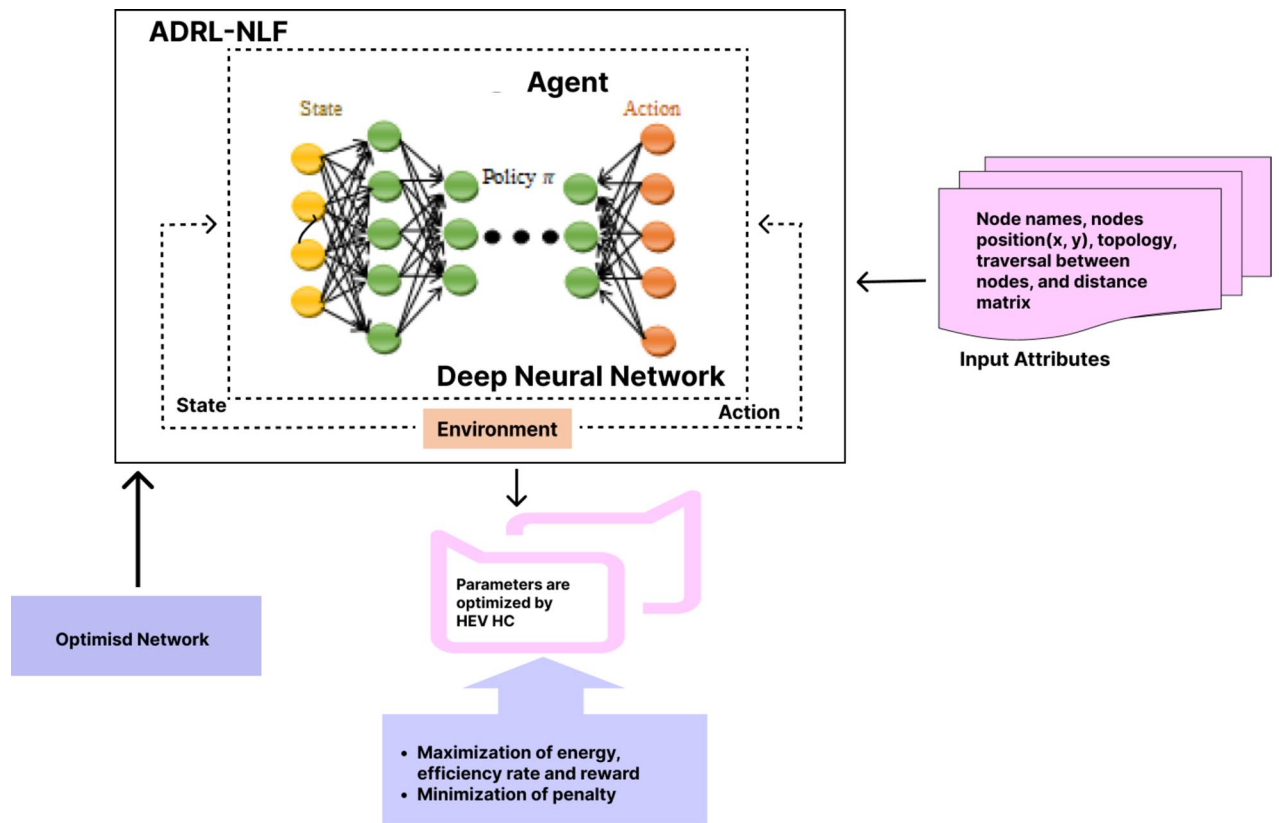


Fig. 5. The framework of the ADRL-NLF with the aid of designed HEVEC.

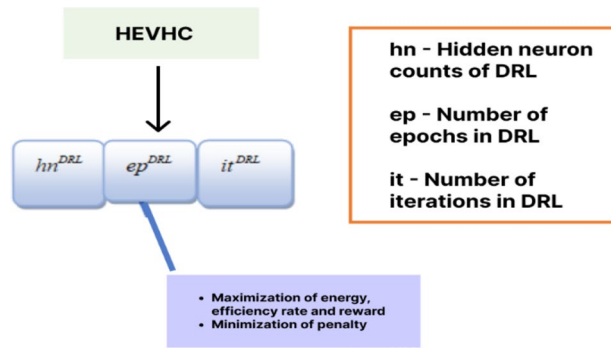


Fig. 6. The solution encoding diagram of the implemented framework.

Sl.No	Parameters	Value
1.	UAV step distance	[0–20] m
2.	Path loss exponent	2
3.	Bandwidth	1 MHz
4.	SINR threshold/Noise power	5 dB/ – 130 dBms
5.	Maximum transmit power	20 dBm
6.	Nominal battery capacity	16,000 mAh
7.	Weight per UAV/Number of UAVs	16 kg/[2–12]
8.	Velocity/Ground user direction	[0, 15]m/s / [0, 2 π]
9.	Model/ Number of Ground Users	GMM/400
10.	Policy/.maxStep/Episodes	ϵ -greedy/1500/250
11.	Batch size/Replay memory size	1024/10,000
12.	Activation function/Hidden layers	ReLU/2 (128,64)
13.	Discount factor/learning rate	0.95/0.0001
14.	Loss function/Optimizer	MSELoss/RMSprop

Table 2. Parameter setting.

Convergence examination of the implemented HEVHC algorithm

The presented HEVHC algorithm’s convergence is validated and displayed in Fig. 7, by influencing the iteration values. When concentrating on the 125th iteration, the proposed HEVHC’s convergence is strengthened by 96.16% of ESOA, 96.32% of CMPA, 96.4% of EVO, and 96.36% of HCO accordingly. From the evaluations, it is presented that the implemented HEVHC has a better capacity to recognize the optimal outcomes.

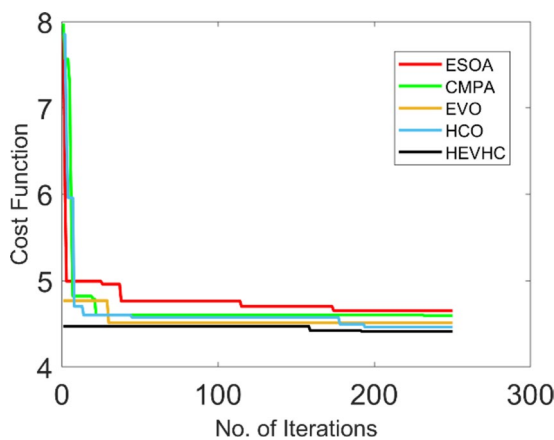


Fig. 7. The implemented HEVHC algorithm’s convergence validation over multiple traditional algorithms.

Statistical analysis of the HEVHC algorithm

Table 3 presents the statistical information about the proposed HEVHC algorithm by employing various statistical measures. When considering the median attribute, the HEVHC is advanced by 5.1% of ESOA, 2.8% of CMPA, 0.8% of EVO, and 2.2% of HCO correspondingly. Therefore, it is confirmed that the HEVHC algorithm has better functionality rates to support the designed energy-efficient framework.

Performance analysis of the energy efficiency framework in UAV-aided networks based on multiple loss functions

Figure 8 offers the performance evaluation of the implemented energy efficiency method in the UAV-aided networks over multiple conventional optimization tasks by adopting powerful loss functions. In Fig. 8a, when considering the novel loss function, the energy efficiency rate of implemented work is enriched by 11.11% of ESOA-ADRL-NLF, 8.88% of CMPA-ADRL-NLF, 22.22% of EVO-ADRL-NLF, and 5.55% of HCO-ADRL-NLF

“Terms”	ESOA ²⁸	CMPA ²⁹	EVO ¹⁷	HCO ¹⁸	HEVHC
“Best”	4.6514	4.596	4.5117	4.4627	4.4125
“Worst”	7.844	7.9765	4.7691	7.8536	4.4706
“Mean”	4.7625	4.6784	4.5415	4.607	4.4505
“Median”	4.7008	4.6008	4.5117	4.5743	4.4706
“Std”	0.23925	0.4291	0.082603	0.37152	0.02659

Table 3. The statistical analysis of the HEVHC algorithm over distinct conventional algorithms.

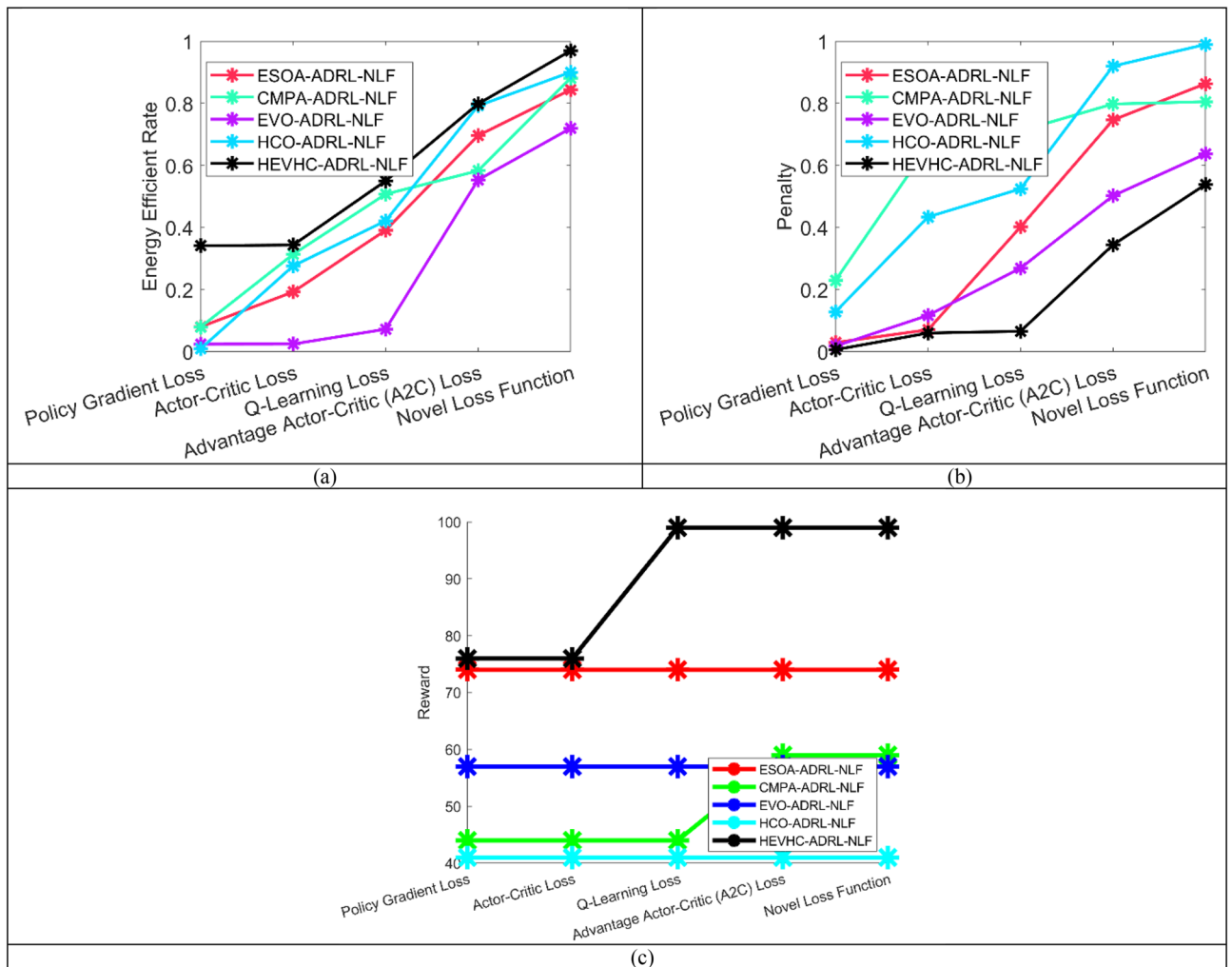


Fig. 8. The performance analysis of the implemented energy efficiency framework in UAV-aided networks over diverse classical approaches based on multiple loss functions concerning “(a) energy efficiency rate, (b) penalty, and (c) reward”.

appropriately. Thus, it is guaranteed that the designed process has higher energy efficiency than the other conventional mechanisms.

Performance analysis of the energy efficiency framework in UAV-aided networks based on the number of nodes

With the change in the number of nodes, the functionality of the implemented energy efficiency framework in UAV-aided networks is carried out over numerous classical approaches. This is presented in Fig. 9. The reward of the constructed work is enhanced by 47% of ESOA-ADRL-NLF, 36% of CMPA-ADRL-NLF, 22% of EVO-ADRL-NLF, and 58% of HCO-ADRL-NLF correspondingly when the node value is 100, as shown in Fig. 9b. Therefore, it is proved that the suggested model has higher robustness than the conventional approaches.

Performance analysis of the energy efficiency framework in UAV-aided networks based on the mobility models

According to numerous mobility models such as RWP, RW, GMM, and static, the performance of the framework is determined and shown in Fig. 10. The fairness index of the designed framework is escalated by 12% of ESOA-ADRL-NLF, 5.1% of CMPA-ADRL-NLF, 10.3% of EVO-ADRL-NLF, and 6.8% of HCO-ADRL-NLF correspondingly, when considering the mobility model as static in Fig. 10a. This finalizes that the implemented framework has higher scalability than the other models.

Performance analysis of the energy efficiency framework in UAV-aided networks based on the number of UAVs

Figure 11 offers the performance investigation of the implemented framework according to the number of UAVs against numerous classical tasks. The overall energy consumed in the presented framework is decreased by 41.5% of ESOA-ADRL-NLF, 91.5% of CMPA-ADRL-NLF, 49.9% of EVO-ADRL-NLF, and 93% of HCO-ADRL-NLF accordingly when the number of UAV is 6 in Fig. 11d. Therefore, it is highlighted that the designed framework has higher efficacy than others.

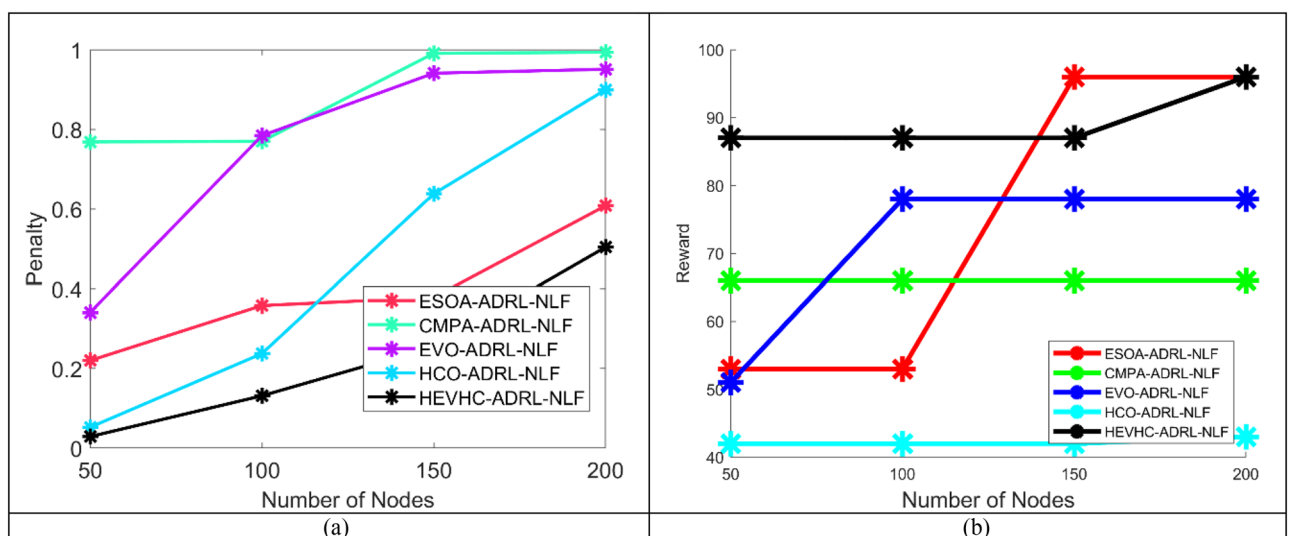


Fig. 9. The performance analysis of the implemented energy efficiency framework in UAV-aided networks over diverse classical approaches based on the number of nodes concerning “(a) Penalty, and (b) Reward”.

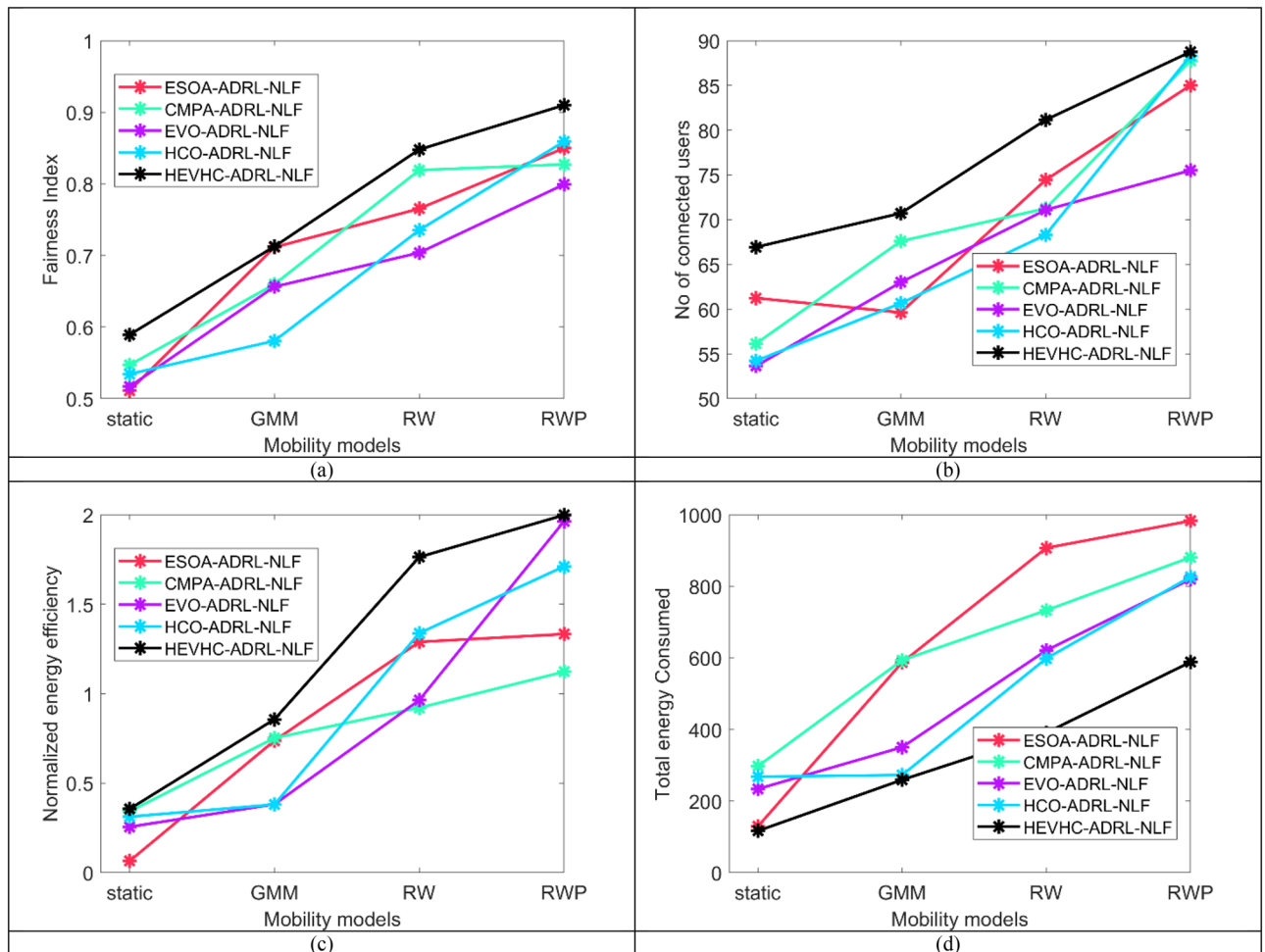


Fig. 10. The performance analysis of the implemented energy efficiency framework in UAV-aided networks over diverse classical approaches based on the mobility models concerning “(a) fairness index, (b) number of connected users, (c) normalized energy efficiency, and (d) total energy consumed”.

Conclusion

The proposed work has been focussed on developing a powerful energy efficiency approach in the UAV-aided framework utilizing the DRL. The optimization of energy efficiency in the UAV provided the superior wireless coverage to the mobile and the static ground candidate. Currently, the DRL tasks were exceptionally optimized for the model's energy efficacy rate by adopting the 2D trajectory strategy that avoided the rate of intrusion obtained in the neighbourhood UAV cells. The crucial concept of the designed approach was to enhance the rate of energy efficiency in the UAV framework by conducting the joint optimization, overall UAV 3D trajectory power employed during the intrusion accounting, and the mapped candidate counts. Thus, an effective ADRL-NLF task was developed to offer a better energy efficacy rate to the network of UAV. In addition, the attributes of the ADRL were optimally tuned utilizing the HEVHC algorithm. In addition, diverse validation was carried out to validate the supremacy of the implemented energy efficiency approach for the UAV-aided network against the conventional energy efficiency approach for the UAV networks. When focussing on the 150th node, the proposed energy efficiency framework in UAV-aided network's penalty was reduced by 99.74% of ESOA-ADRL-NLF, 99.48% of CMPA-ADRL-NLF, 99.47% of EVO-ADRL-NLF, and 99.86% of HCO-ADRL-NLF accordingly. Thus, the proposed energy efficiency framework in UAV-aided networks has achieved its supremacy over the other classical mechanisms. These enhancements can be aided to support drones in futuristic mission critical systems such as³³, for handling emergency situations for medical and other applications.

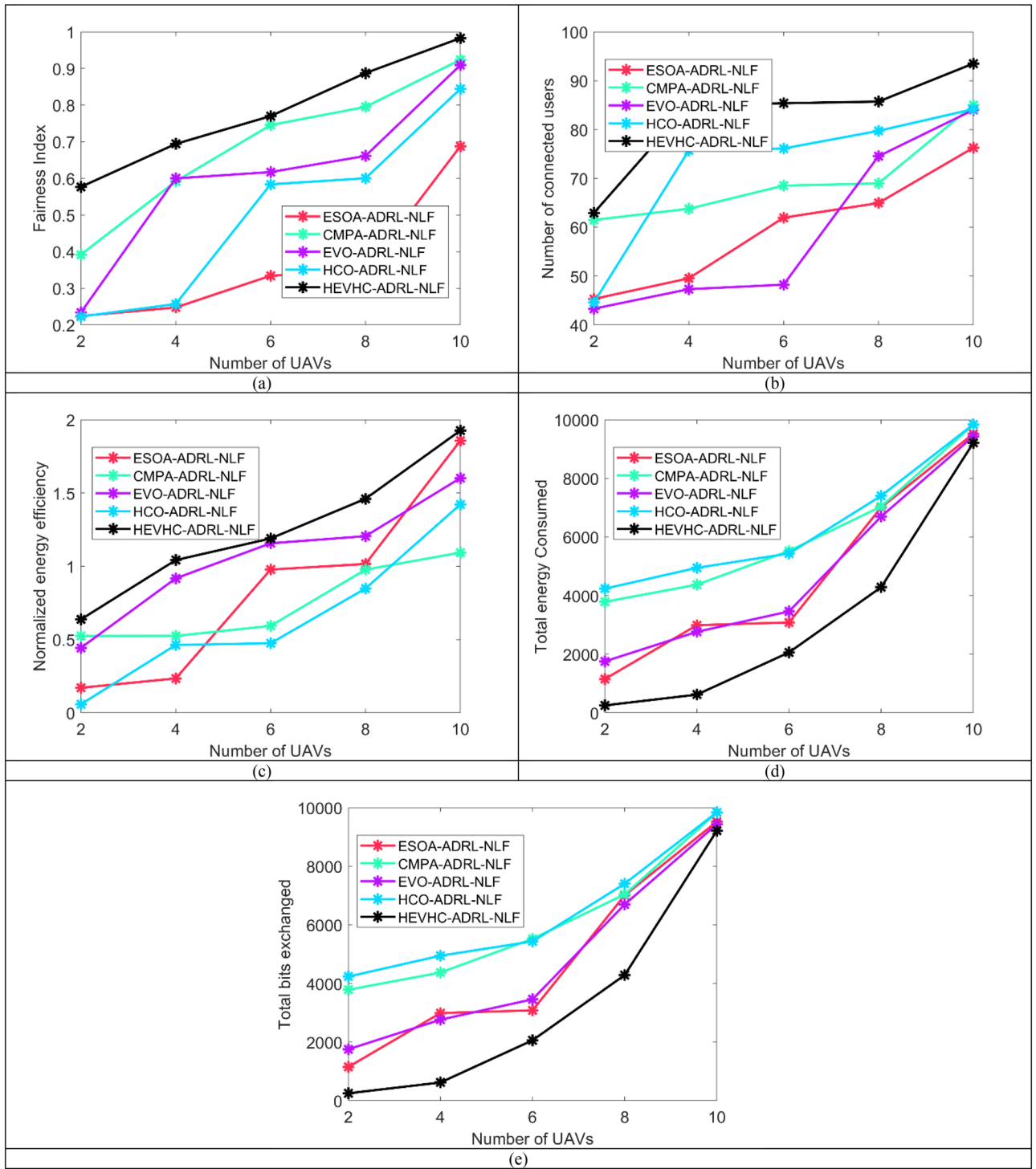


Fig. 11. The performance analysis of the implemented energy efficiency framework in UAV-aided networks over diverse classical approaches based on the number of UAVs concerning “(a) fairness index, (b) number of connected users, (c) normalized energy efficiency, (d) total energy consumed, and (e) total bits.

Data availability

The datasets used and/or analysed during the current study are available from the corresponding author on request.

Received: 25 February 2024; Accepted: 29 August 2024
 Published online: 27 September 2024

References

- Ranjha, A. *et al.* Toward Facilitating Power Efficient URLLC Systems in UAV Networks Under Jittering. In *IEEE Transactions on Consumer Electronics*, **70**(1), 3031–3041. <https://doi.org/10.1109/TCE.2023.3305550> (2024).
- Ranjha, A., Javed, M.A., Srivastava, G. & Asif, M. Quasi-optimization of resource allocation and positioning for solar-powered UAVs. *IEEE Transactions on Network Science and Engineering*, 1–10. <https://doi.org/10.1109/TNSE.2023.3282870> (2023).
- Ranjha, A., Naboulsi, D., El Emary, M. & Gagnon, F. Facilitating URLLC vis-à-vis UAV-Enabled Relaying for MEC Systems in 6-G Networks. *IEEE Transactions on Reliability*, 1–15. <https://doi.org/10.1109/TR.2024.3357356> (2024).
- Liu, C. H., Chen, Z., Tang, J., Xu, J. & Piao, C. Energy-efficient UAV control for effective and fair communication coverage: A deep reinforcement learning approach. *IEEE J. Sel. Areas Commun.* **36**(9), 2059–2070 (2018).
- Nguyen, K. K. *et al.* Real-time energy harvesting aided scheduling in UAV-assisted D2D networks relying on deep reinforcement learning. *IEEE Access* **9**, 3638–3648 (2020).
- Wei, D. *et al.* Computation offloading over multi-UAV MEC network: A distributed deep reinforcement learning approach. *Comput. Netw.* **199**, 108439 (2021).
- Hoang, T. M., Nguyen, B. C. & Kim, T. Performance analysis and optimization of UAV-assisted NOMA short packet communication systems. *ICT Express* **10**, 292–298 (2023).
- Wang, H., Ke, H. & Sun, W. Unmanned-aerial-vehicle-assisted computation offloading for mobile edge computing based on deep reinforcement learning. *IEEE Access* **8**, 180784–180798 (2020).
- Li, S., Hu, X. & Du, Y. Deep reinforcement learning for computation offloading and resource allocation in unmanned-aerial-vehicle assisted edge computing. *Sensors* **21**(19), 6499 (2021).
- Ebrahim, M. A., Ebrahim, G. A., Mohamed, H. K. & Abdellatif, S. O. A deep learning approach for task offloading in multi-UAV aided mobile edge computing. *IEEE Access* **10**, 101716–101731 (2022).
- Zhang, P. *et al.* Deep reinforcement learning based computation offloading in UAV-assisted edge computing. *Drones*, **7**, 213 (2023).
- Xu, Y.-H., Sun, Q.-M., Zhou, W. & Yu, G. Resource allocation for UAV-aided energy harvesting-powered D2D communications: A reinforcement learning-based scheme. *Ad Hoc Netw.* **136**, 102973 (2022).
- Xiong, Z. *et al.* UAV-assisted wireless energy and data transfer with deep reinforcement learning. *IEEE Trans. Cogn. Commun. Netw.* **7**(1), 85–99 (2020).
- Qin, Z. *et al.* Distributed UAV-BSs trajectory optimization for user-level fair communication service with multi-agent deep reinforcement learning. *IEEE Trans. Veh. Technol.* **70**(12), 12290–12301 (2021).
- Zhan, C. & Zeng, Y. Energy minimization for cellular-connected UAV: From optimization to deep reinforcement learning. *IEEE Trans. Wirel. Commun.* **21**(7), 5541–5555 (2022).
- Chen, D. *et al.* Mean field deep reinforcement learning for fair and efficient UAV control. *IEEE Internet Things J.* **8**(2), 813–828 (2020).
- Hu, J., Zhang, H., Song, L., Han, Z. & Vincent Poor, H. Reinforcement learning for a cellular internet of UAVs: Protocol design, trajectory control, and resource management. *IEEE Wirel. Commun.* **27**(1), 116–123 (2020).
- Qi, W., Song, Q., Guo, L. & Jamalipour, A. Energy-efficient resource allocation for UAV-assisted vehicular networks with spectrum sharing. *IEEE Trans. Veh. Technol.* **71**(7), 7691–7702 (2022).
- Ahmed, S., Chowdhury, M. Z. & Jang, Y. M. Energy-efficient UAV-to-user scheduling to maximize throughput in wireless networks. *IEEE Access* **8**, 21215–21225 (2020).
- Omoniwa, B., Galkin, B. & Dusparic, I. Communication-enabled deep reinforcement learning to optimize energy-efficiency in UAV-assisted networks. *Veh. Commun.* **43**, 100640 (2023).
- Omoniwa, B., Galkin, B. & Dusparic, I. Optimizing energy efficiency in UAV-assisted networks using deep reinforcement learning. *IEEE Wirel. Commun. Lett.* **11**(8), 1590–1594 (2022).
- Hu, H. *et al.* Optimization of energy efficiency in UAV-enabled cognitive IoT with short packet communication. *IEEE Sens. J.* **22**(12), 12357–12368 (2021).
- Min, J., Gao, Q., Guo, Q. & Xuemai, Gu. Energy-efficiency power allocation design for UAV-assisted spatial NOMA. *IEEE Internet Things J.* **8**(20), 15205–15215 (2020).
- Tang, X., Wang, W., He, H. & Zhang, R. Energy-efficient data collection for UAV-assisted IoT: Joint trajectory and resource optimization. *Chin. J. Aeronaut.* **35**(9), 95–105 (2022).
- Chen, X., Liu, Xu., Chen, Y., Jiao, L. & Min, G. Deep Q-network based resource allocation for UAV-assisted ultra-dense networks. *Comput. Netw.* **196**, 108249 (2021).
- Sikeridis, D., Tsiropoulou, E. E., Devetsikiotis, M. & Papavassiliou, S. Wireless powered public safety IoT: A UAV-assisted adaptive-learning approach towards energy efficiency. *J. Netw. Comput. Appl.* **123**, 69–79 (2018).
- Zhang, L., Celik, A., Dang, S. & Shihada, B. Energy-efficient trajectory optimization for UAV-assisted IoT networks. *IEEE Trans. Mob. Comput.* **21**(12), 4323–4337 (2021).
- Azizi, M., Aickelin, U., Khorshidi, H. A. & Shishehgharkhaneh, M. B. Energy valley optimizer: A novel metaheuristic algorithm for global and engineering optimization. *Sci. Rep.* **13**, 226 (2023).
- Tafakkori, K. & Tavakkoli-Moghaddam, R. Hermit crab optimizer (HCO): A novel meta-heuristic algorithm. *IFAC-PapersOnLine* **55**(10), 702–707 (2022).
- Agarwal, V. & Tewari, R. R. Improving energy efficiency in UAV attitude control using deep reinforcement learning. *J. Sci. Res.* **65**(3), 209–219 (2021).
- Chen, Z. *et al.* Egret swarm optimization algorithm: An evolutionary computation approach for model-free optimization. *Biometrics* **7**(4), 144 (2022).
- Yuan, Y. *et al.* Coronavirus mask protection algorithm: A new bio-inspired optimization algorithm and its applications. *J. Bionic Eng.* **20**(4), 1747–1765 (2023).
- Shankar, N., Nallakaruppan, M. K., Ravindranath, V., Senthilkumar, M. & Bhagavath, B. P. Smart IoMT framework for supporting UAV systems with AI. *Electronics* **12**(1), 86. <https://doi.org/10.3390/electronics12010086> (2023).

Author contributions

All authors contributed equally to the manuscript. All authors reviewed the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to S.S.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2024