International Journal of Computer Networks and Applications (IJCNA)
DOI: 10.22247/ijcna/2023/220740 Volume 10, Issue 2, March – April (2023)

RESEARCH ARTICLE

QMRNB: Design of an Efficient Q-Learning Model to Improve Routing Efficiency of UAV Networks via Bioinspired Optimizations

Anshu Vashisth
School of Computer Science and Engineering, Lovely Professional University, Punjab, India.
anshu.23500@lpu.co.in

Balraj Singh
School of Computer Science and Engineering, Lovely Professional University, Punjab, India.
balraj.13075@lpu.co.in

Ranbir Singh Batth
Menzies Institute of Technology, Melbourne, Australia.
ranbir@menzies.vic.edu.au

Received: 19 February 2023 / Revised: 14 April 2023 / Accepted: 23 April 2023 / Published: 29 April 2023

Abstract – The design of efficient routing strategies for Unmanned Aerial Vehicle (UAV) Networks is a multidomain task that involves analysis of node-level & network-level parameters, and mapping them with communication & contextual conditions. Existing path planning optimization models either showcase higher complexity or cannot be scaled for larger network scenarios. Moreover, the efficiency of these models also reduces w.r.t. the number of communication requests, which limits their scalability levels. To get a better result over these challenges, this article provides an idea to design an efficient Q-Learning model to improve the routing efficiency of UAV networks via bioinspired optimizations. The model initially collects temporal routing performance data samples for individual nodes and uses them to form coarse routes via Q-Learning optimizations. These routes are further processed via a Mayfly Optimization (MO) Model, which assists in the selection of optimal routing paths for high Quality of Service (QoS) even under large-scale routing requests. The MO Model can identify alternate paths via the evaluation of a high-density routing fitness function that assists the router in case the selected paths are occupied during current routing requests. This assists in improving temporal routing performance even under dense network conditions. Due to these optimizations, the model is capable of reducing the routing delay by 8.5%, improving energy efficiency by 4.9%, and reducing the routing jitter by 3.5% when compared with existing routing techniques by taking similar routing conditions.

Index Terms – UAV, Routing, Delay, Energy, Mayfly, Optimization, Jitter, efficiency, Complexity.

1. INTRODUCTION

Due to the continual movement of vehicles, the UAV (Unmanned Aerial Vehicle) routing protocol must deal with a variety of issues, including unequal node distribution, topological changes, and changes in the surrounding environment via Energy-aware Collaborative Routing (ECoR) [1, 2, 3, 4]. Q-learning (QL) was included to make UAV routing [5, 6] more adaptable and sensitive to the dynamic environment. Traditional reinforcement learning is referred to as Q-learning, and it is distinguished by the lack of a state transition model in favor of an assessment of the value of state-action pair combinations. The following are the five components of Q-learning: s, a, R, where s represents the state set of RL, a represents the action set of RL, and R and R represent the attenuation factor of future reward and the rate of learning in case of reinforcement learning [7, 8, 9], respectively. If an "a" operation is performed in a certain state, the hope will upgrade its state value table by inserting equation (1), where s is the succeeding state (Q-value table). Equation (1) tells about the updated and succeeding state of the Q-Learning table. Following a definite no. of repetitions, the Q-table will determine the optimal action for each state. This operation will guarantee that the node in question gets the maximum reward available for the current set of iterations.

$$Q(s,a) = (1-\alpha)Q(s,a) + \alpha\big(R + \gamma \, max \, Q(s\,,a\,)\big) \ldots (1)$$

RESEARCH ARTICLE

The research contribution of the proposed Q Learning model could lead to improve routing efficiency in UAV networks which could have various advantages in terms of throughput, latency, and energy consumption. By incorporating bioinspired optimization Q Learning model gives better performance than normal. The performance can also measure in terms of the number of states required to take the action, which reduces the complexity of the model as well as increase the scalability of the model. The proposed research work contributes to the field of UAV networks by giving novel mechanisms to improve performance and also to make efficient use of resources used. Overall contribution can vary as per real-time application requirements such as logistics, surveillance, etc. The proposed model used unique characteristics of the network in terms of dynamic topology and energy consumption. The q-Learning model is a reinforcement algorithm that can handle every situation from its learning database. The model can take decisions even if the condition does not belong to his data base. It can work for stochastic transition. This paper presents an efficient Q Learning Model for improving routing efficiency in UAV networks.

In section 2 literature work has been done in which various latest techniques have been discussed in detail to show the strength of the proposed model. In section 3 proposed methodology is discussed by using continuous pattern analysis of dynamic collision-aware UAV networks. The building of an effective Q-Learning model to increase the routing efficiency of UAV networks using bioinspired optimizations is suggested in section 3 of this article as a solution to efficiency during the collision. In section 4, the suggested model's performance was assessed using large-scale network scenarios and compared to that of conventional routing models. This paper concludes with some network-specific observations regarding the suggested model and suggestions for ways to further enhance its functionality in various network scenarios in section 5.

2. LITERATURE WORK

The paper [32] proposed a strategy to select optimal path selection using the Q learning model for the UAV network. The state space of the proposed model is temperature, humidity, and energy used. The author used the Voronoi diagram to the partition urban environment, where each partition represents a cell and the cell becomes a city for a region. The author has given weight to each cell based on building density and calculated the future courses of possible action. The result of the proposed method is evaluated in simulation in the urban environment and has been found that gives better performance in terms of energy consumption. But the main challenge of this proposed model was the dependency on accurate input data e.g. building density, temperature, humidity, etc. Due to incomplete or inaccurate input data proposed model will not give a better result. Secondly proposed model is only valid for the urban area but there can need for other types of environments. Lastly, the use of the Voronoi diagram and assignment of weight to each cell is also a computational task that increases complexity and can lead to latency when the urban area becomes denser.

In the paper [33] author provides a deep insight into the Q learning model by considering position-aware routing protocols. The individual use of the Q learning model has some disadvantages in it like sensitivity and accurate position information. So to get better performance in terms of routing efficiency and energy consumption one has to use another model with the Q learning model to make it efficient in the adversarial environment too.

In the paper [34] authors proposed an algorithm to optimize the path by finding the optimal energy consumption method with the given condition like the limited battery life of the device and complex dynamic environment. The proposed model has been a combination of two techniques MACO (Multi-Objective Ant Colony Optimization) and MEA (Multi-Objective Evolutionary Algorithm). Here MACO is used to optimize path planning and MEA is used to optimize energy consumption. Both MACO and MEA use the graph to check the tradeoff between flight time and safety. Author evaluated the proposed model using experimental and theoretical grounds. On the theoretical ground proposed model outperforms and on the experimental ground model significantly reduces energy consumption. The main challenge of this proposed model is computational complexity because the generation of the graph and the hybrid combination of the two models made this model a bit lengthy and complex. Secondly, convergence issues, where sometimes an algorithm fails to find an optimal solution and ended at a suboptimal solution.

The paper [35] proposed a novel deep learning approach to optimize the communication between two nodes. The author used DBS (Drone Base Station) mechanism with the Q learning model to improve the efficiency of the network by considering parameters like network topology, link quality, etc. The author simulates the result and found that model gives a significantly better result in terms of energy consumption. The main challenge of this method was training complexity. In this proposed method Q learning model has been trained which can be time consuming and computationally intensive. The proposed model is tested in simulation so there can be a chance that the model cannot outperform in real time applications where training is required in high volume.

The paper [36] presented a systematic review of the classical approach used for the optimization of path and energy consumption in UAV networks. The authors compared the classical approach with the reinforcement learning approach

RESEARCH ARTICLE

by considering their strengths and weakness. The authors reviewed that classical techniques like ant colony optimization, particle swarm optimization, etc can be used to optimize path and energy consumption. But the techniques like the Q learning model and deep Q learning can add more advantages to classical techniques in terms of adapting the change in the environment. The author concluded that the hybrid approach could give better results under various conditions in an adversarial environment.

The paper [37] provides descriptive knowledge about various path planning existing techniques used in UAV networks. Various approaches have been used like geometric, sampling based, machine learning based, etc. whereas in paper [38] provides challenges that occurred in UAV networks. In this paper Classification of UAVs, and networks have been done by considering their multiple parameters like navigation info, flight time, flight angle, etc. Connectivity and Coverage have been discussed in detail as a challenge to UAV networks.

To achieve the objectives of reinforcement learning, the set of neighbor nodes is considered, and the base stations (BS) is treated as the fixed destination node that broadcasts hello packets regularly via the use of Q -learning-based topology-aware routing (QTAR) [10, 11, 12, 13, 14]. By the aforementioned principles, the receiving node is required to update the Q-value table of its device: the higher the Q-value, the closer the device is to RS. When this method is used, the routing to static destination nodes is enhanced.

The work in [15, 16, 17, 18] utilizes the conventional Adhoc on Demand Multipath Distance Vector (AOMDV) routing protocol in addition to the Q-learning approach. The hopes can upgrade the Q-value information stored in their respective local memory by exchanging the hello and RREQ packets that need their route information. Using the AODV routing algorithm established in [19, 20], an excellent degree of performance was achieved in a case with restricted mobility. In [21, 22, 23, 24], unmanned aerial vehicles (UAVs) have been used to aid VANET (Vehicular Adhoc Networks) in determining the most efficient route for data transmission. However, the following are some of the most common problems that emerge with such routing systems: Each node on the ground is responsible for keeping its Q-value table, regardless of whether it has neighbors or not. Because (1) ground nodes only employ locally stored information to decide the next hop [25, 26, 27, 28] and (2) both the size of the Q-value info and the stored info are subject to rapid change [29, 30, 31], this leads in increased bandwidth use and a slower convergence speed of Q-values & their alternatives.

As a result, conventional route optimization techniques either have increased complexity or are not saleable for situations involving bigger networks. Additionally, these models' efficiency declines as the number of communication requests increases, which restricts the extent of their scalability.

3. PROPOSED HYBRID BIOINSPIRED MODEL WITH CONTINUOUS PATTERN ANALYSIS FOR DYNAMIC COLLISION-AWARE ROUTING IN UAV NETWORKS

As per the review of existing routing models that are used for UAV Networks, it has been found that existing path planning optimization models provide and showcase higher complexity or cannot be scaled for larger network scenarios. Moreover, the efficiency of these models reduces w.r.t. the number of communication requests, which limits their scalability levels. To get rid of these challenges, section 3 discusses the design of an efficient Q-Learning model to improve the routing efficiency of UAV networks via bioinspired optimizations. As per Figure 1, it has been found that the model begins with collecting temporal routing performance data samples for individual nodes, and uses them to form coarse routes via Q-Learning optimizations. These routes are further processed by a Mayfly Optimization (MO) Model, which assists in the selection of optimal routing paths for high Quality of Service (QoS) even under large-scale routing requests. The MO Model can identify alternate paths via the evaluation of a high-density routing fitness function that assists the router in case the selected paths are occupied during current routing requests. This assists in improving temporal routing performance even under dense network conditions.

Thus, the model initially uses Q-Learning to identify different routes between a given source & destination pair of nodes. Equation (2) calculate the optimal distance between the source and destination. This is done by initially calculating a reference distance between these nodes via equation (2),

$$d_{ref} = \sqrt{\begin{matrix}(x_{src} - x_{dest})^2 + (y_{src} - y_{dest})^2 \\ +(z_{src} - z_{dest})^2\end{matrix}} \dots (2)$$

Where, $x, y \,\&\, z$ are the Cartesian locations of these nodes, while $src \,\&\, dest$ are the IP addresses of source & destination nodes.

Now, based on the calculated distance between the source and destination, select all nodes that satisfy equation (3),

$$d_{src,i} < d_{ref} \,\&\, d_{i,dest} < d_{ref} \dots (3)$$

Where, $d_{src,i}$ is the distance between the selected node & the source node, while $d_{i,dest}$ represents the distance between selected & destination node sets. For all these nodes, evaluate their Q values via equation (4),

$$Q = \sum_{i=1}^{N_p} \frac{\sum NC_i}{\sum LQ_i} \dots (4)$$

RESEARCH ARTICLE

This Q value is updated via equation (5), where this Q value gives an idea about that how well an action is taken at a particular instance based on its previous state.

$$Q(New) = Q(Old) + L_r * \frac{NC}{Max(NC)} + Max(Q) \ldots (5)$$

Where N_p represents the number of nodes in the current path, L_r is a stochastic learning rate, while NC & LQ are the node communication metric and link quality metric, which is estimated via equations 6 & 7 as follows. Equation (6) and (7) helps to find communication quality by considering optimization parameter to make use of randomness using a Bioinspired algorithm like Mayfly.

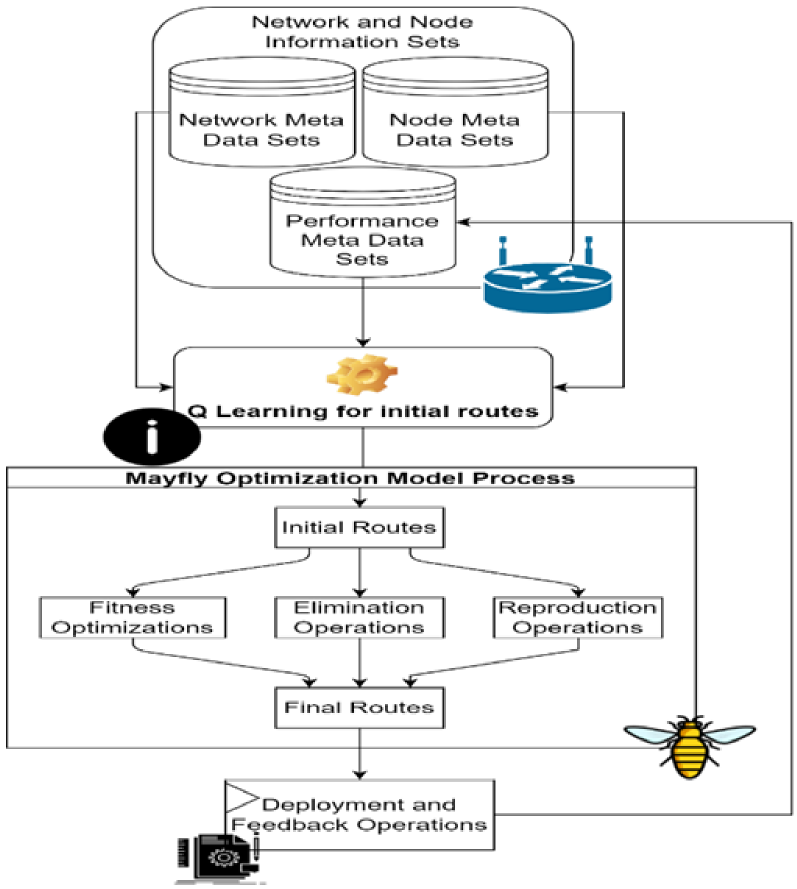

Figure 1 Proposed Routing Model for UAV Networks

$$NC = \frac{1}{N_h} \sum_{i=2}^{N_h} d_{i-1,i} * \left[\frac{THR_{i-1}}{Max(THR)}\right] \ldots (6)$$

Where, d is the distance, while THR is the temporal throughput, which is evaluated via equation (8), while N_h is the total number of hops used for the routing operations. The output of the function and routing operation performed as per the number of hops used is calculated by equation (8).

$$LQ = \frac{1}{N_h} \sum_{i=1}^{N_h} \frac{100}{PDR_i} + \frac{e_i}{Max(e)} \ldots (7)$$

Where, PDR & e represent the packet delivery ratio and energy consumption during previous communications which are estimated at equations (9) & (10), which are updated after individual routing operations.

$$THR = \sum_{t=t1}^{t2} \frac{NN(t)}{Max(NN) * (t2 - t1)} \ldots (8)$$

Where $NN(t)$ represents the total number of packets passed during the given time intervals.

$$PDR = \sum_{t=t1}^{t2} \frac{NN(t)}{NN_d(t) * (t2 - t1)} \ldots (9)$$

Where NN_d represents the total number of packets dropped during the given time intervals.

RESEARCH ARTICLE

$$e = \sum_{t=t1}^{t2} \frac{e_{start} - e_{complete}}{(t2 - t1)} \dots (10)$$

Where e_{start} & $e_{complete}$ are the energy levels of the nodes during the routing process. These Q values are sorted in descending order, and then N stochastic nodes are selected from this list via equation (11),

$$N = STOCH(L_r * N_n, N_n) \dots (11)$$

Where, N_n represents the total no. of hopes used in the list, while L_r is estimated (L_r is a stochastic learning rate) via equation (12),

$$L_r = \frac{N_n}{N_t} \dots (12)$$

Where N_t represents the total number of nodes in the network that are deployed in the UAV network sets. Based on this process, a set of NM Mayflies (routes) are generated, and optimized via the following Mayfly Optimization (MO) Model process,

From the set of Q learning solutions, N stochastic solutions are selected via equation (13),

$$N = STOCH(LM * NM, NM) \dots (13)$$

Where LM is the learning metric for the MO Model process.

For each of these solutions, a fitness value function is calculated via equation (14),

$$f = \frac{1}{NM} \sum_{i=1}^{NM} Q_i \dots (14)$$

This process is repeated for NM different Mayflies, and then a fitness threshold is estimated via equation (15),

$$f_{th} = \frac{1}{NM} \sum_{i=1}^{NM} f_i * LM \dots (15)$$

Mayflies with $f > f_{th}$ are discarded & reproduced in the next iteration, while others crossover to the next set of iterations.

These Mayflies are regenerated for NI iterations, and the Mayflies with the lowest fitness levels are selected for routing the UAV nodes. The selected Mayfly will contain multiple routing configurations, out of which the configuration with minimum fitness is selected for routing operations. In case the current route is busy or the path is not available, then the next higher fitness path is selected to route the UAV Nodes. Due to this, the model can optimize routing delay, energy, throughput, and packet delivery ratios during real-time route formation operations. The performance during routing is updated in the database, and the process is repeated for consecutive routing processes. This assists in the continuous improvement of routing performance under real-time scenarios. This performance was validated under different network conditions and compared with other models in the next section of this text.

4. STATISTICAL ANALYSIS

The proposed QMRNB model collects temporal routing performance data samples for individual nodes and uses them to form coarse routes through Q-Learning optimizations. These routes are then processed by a Mayfly Optimization (MO) Model, which aids in the selection of optimal routing paths for high Quality of Service (QoS) even when large-scale routing requests are being processed. The MO Model is capable of identifying alternate paths through the evaluation of a high-density routing fitness function, which assists the router if the selected paths are occupied during current routing requests. This helps to enhance temporal routing performance even in dense network environments. To validate its performance, the model was evaluated under a standard set of UAV configurations in Network Simulator 2 (NS 2.34), with the network parameters indicated in Table 1 as follows:

Table 1 Set of Simulation Configurations for Emulating Different Network Scenarios

Parameters for the UAV Network	Values used for the parameter sets
UAV propagation model	Wireless with inter-layer routing
MAC Model	802.16a
Queue Type	Priority queues with drop-tailing operations
Total UAV Nodes	5000
Size of the UAV Network	4 kms x 4 kms
Energy Model	Idle: 5 mW Receive: 10 mW Transmit: 15 mW Sleep: 1 mW Transition: 2.5 mW
Transition Delays	0.01 s
Energy set during initialization of UAV Nodes	2500 mW

Based on these configurations, the delay needed for routing was estimated via equation (16),

$$D = \frac{1}{NM} \sum_{i=1}^{NM} ts_{reach} - ts_{start} \dots (16)$$

RESEARCH ARTICLE

Where ts_{reach} & ts_{start} are the timestamps for reaching the destination and starting the routing process. This delay performance was evaluated w.r.t. different Number of Movements (NM), which were varied between 2k to 40k, and compared with ECOR [3], QL [5], & QTAR [14] in table 2 as follows:

Table 2 The Average Delay for Different Numbers of Routing Evaluations

NM	D (s) ECOR [3]	D (s) QL [5]	D (s) QTAR [14]	D (s) QM RNB
2k	15.95	17.09	18.03	12.23
4k	18.13	19.62	20.74	14.10
6k	20.96	22.79	24.09	16.39
8k	24.42	26.54	28.00	19.04
10k	28.42	30.73	32.33	21.96
12k	32.77	35.27	37.00	25.06
14k	37.21	39.94	41.79	28.19
16k	41.46	44.53	46.53	31.23
18k	45.39	48.91	51.06	34.11
20k	48.94	52.98	55.30	36.81
25k	52.22	56.79	59.26	39.33
28k	55.27	60.26	62.89	41.69
30k	58.17	63.53	66.31	43.93
35k	61.50	67.18	70.10	46.45
38k	64.98	70.94	74.02	49.07
40k	68.71	74.96	78.19	51.84

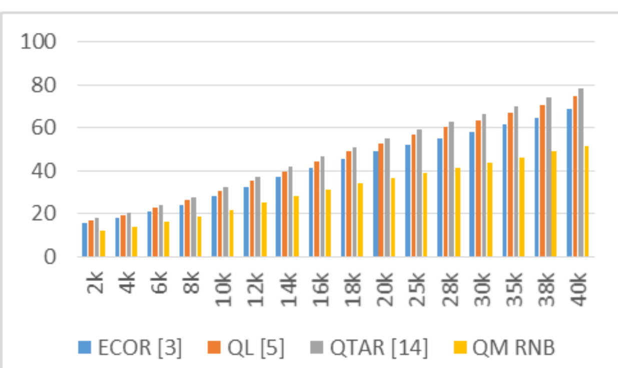

Figure 2 The Average Delay for Different Numbers of Routing Evaluations

Based on this estimation and its visualization in Figure 2, it has been found that the proposed model can reduce the delay needed for routing by 9.5% when compared with ECOR [3], 14.5% when compared with QL [5], and 18.9% when compared with QTAR [14], which makes it very important and critical for various real-time routing applications. This delay is reduced due to the use of distance levels, and temporal delay performance during the selection of routing paths. Similarly, the energy needed during these routing operations was evaluated via equation (17) as follows,

$$E = \frac{1}{NM}\sum_{i=1}^{NM} E_{src}(start)_i - E_{src}(complete)_i \ldots (17)$$

Where $E(start)$ & $E(complete)$ are the energy levels of the source node during the start & completion of the routing process. This energy consumption can be observed in Table 3 as follows:

Table 3 Average Energy Consumed During the Different Number of Routing Evaluations

NM	E (mW) ECOR [3]	E (mW) QL [5]	E (mW) QTAR [14]	E (mW) QM RNB
2k	37.05	48.33	29.84	21.96
4k	38.98	50.78	31.34	23.06
6k	40.83	53.27	32.88	24.21
8k	42.75	55.90	34.50	25.40
10k	44.77	58.64	36.18	26.63
12k	46.87	61.46	37.90	27.89
14k	49.08	64.33	39.62	29.14
16k	51.34	67.17	41.32	30.37
18k	53.60	69.94	42.97	31.56
20k	55.80	72.61	44.57	32.71
25k	57.93	75.21	46.13	33.84
28k	60.01	77.80	47.70	34.99
30k	62.05	80.41	49.30	36.14
35k	64.09	83.10	50.92	37.33
38k	66.15	85.78	52.56	38.53
40k	68.22	88.48	54.20	39.72

Based on this estimation and its visualization in Figure 3, it has been found that the proposed model can reduce the energy needed for routing by 15.9% when compared with ECOR [3],

RESEARCH ARTICLE

19.2% when compared with QL [5], and 14.5% when compared with QTAR [14], which makes it very critical for various low-energy routing applications. This energy is reduced due to the use of distance levels, and temporal energy performance during the selection of routing paths. Similarly, the throughput during these routing operations can be observed from Table 4 as follows:

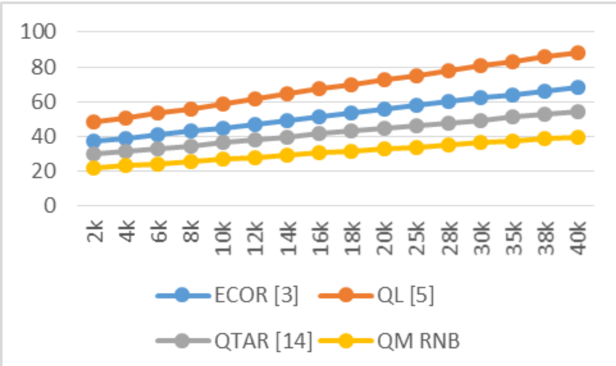

Figure 3 Average Energy Consumed During the Different Number of Routing Evaluations

Table 4 Average Throughput During Different Routing Operations

NM	THR (vpm) ECOR [3]	THR (vpm) QL [5]	THR (vpm) QTAR [14]	THR (vpm) QM RNB
2k	112.86	86.32	92.25	138.86
4k	113.71	87.11	93.00	140.00
6k	114.86	87.63	93.75	141.14
8k	116.00	88.42	94.50	142.29
10k	116.86	89.21	95.25	143.43
12k	117.71	89.74	96.00	144.57
14k	118.57	90.53	96.75	145.71
16k	119.43	91.32	97.50	146.86
18k	120.57	92.11	98.25	148.00
20k	121.71	92.89	99.00	149.14
25k	122.57	93.68	99.75	150.29
28k	123.43	94.47	100.50	151.43
30k	124.29	95.00	101.50	152.57
35k	125.10	95.79	102.34	153.71
38k	126.15	96.50	103.12	154.86
40k	127.15	97.12	103.84	156.00

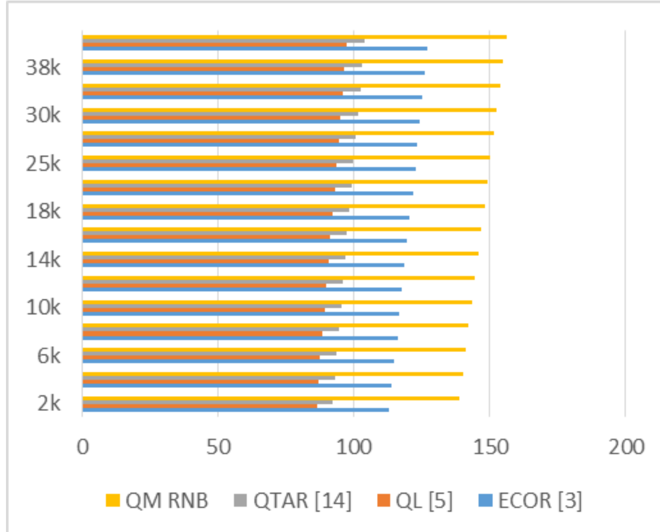

Figure 4 Average Throughput During Different Routing Operations

Based on this estimation and its visualization in Figure 4, it has been found that the proposed model can improve the routing throughput by 8.5% when compared with ECOR [3], 15.4% when compared with QL [5], and 12.5% when compared with QTAR [14], which makes it very important for multiple high data rate routing applications. This throughput is increased due to the use of packet delivery levels, and temporal throughput performance during the selection of routing paths. Due to these optimizations, the proposed model is capable of deployment for a wide variety of real-time UAV routing scenarios.

5. CONCLUSION AND FUTURE SCOPE

The proposed QMRNB model gathers samples of temporal routing performance data for individual nodes and uses Q-Learning optimizations to form coarse routes. A Mayfly Optimization (MO) Model then processes these routes, helping to choose the best routing paths for high Quality of Service (QoS) even when numerous routing requests are being handled simultaneously. If the chosen paths are already taken by current routing requests, the MO Model can find alternative routes by evaluating a high-density routing fitness function. This helps the router. Even in environments with dense network traffic, this aids in improving temporal routing performance. Based on the evaluation of routing speed, it can be seen that the proposed model can reduce the delay required for routing by 9.5% when compared to ECOR [3], 14.5% when compared to QL [5], and 18.9% when compared to QTAR [14], making it very useful for various real-time routing applications. The use of distance levels and the performance of the temporal delay during the selection of routing paths both help to reduce this delay. According to energy evaluation, the proposed model can reduce the energy

RESEARCH ARTICLE

required for routing by 15.9% when compared to ECOR [3], 19.2% when compared to QL [5], and 14.5% when compared to QTAR [14], making it extremely useful for a variety of low-energy routing applications. Due to the use of distance levels and temporal energy performance during the selection of routing paths, this energy is reduced. Based on the data-rate evaluation, it can be seen that the proposed model can increase routing throughput by 8.5% when compared to ECOR [3], 15.4% when compared to QL [5], and 12.5% when compared to QTAR [14], making it extremely useful for a variety of high data rate routing applications. The use of packet delivery levels and temporal throughput performance during routing path selection has increased this throughput. The proposed model can be used in real-time UAV routing scenarios as a result of these optimizations.

In the future, the performance of this model should be cross checked on larger UAV Networks and can be upgraded via the merging of low-complexity bio-inspired models including Genetic Algorithm (GA), Particle Swarm Optimization (PSO), etc. The performance can also be improved via the integration of transformer models that can predict collisions during routing, and enhance efficiency even under large-scale network scenarios.

REFERENCES

[1] N. Toorchi, F. Hu, E. S. Bentley, and S. Kumar, "Skeleton-Based Swarm Routing (SSR): Intelligent Smooth Routing for Dynamic UAV Networks," in IEEE Access, vol. 9, pp. 1286-1303, 2021, doi: 10.1109/ACCESS.2020.3043672.

[2] L. Hong, H. Guo, J. Liu, and Y. Zhang, "Toward Swarm Coordination: Topology-Aware Inter-UAV Routing Optimization," in IEEE Transactions on Vehicular Technology, vol. 69, no. 9, pp. 10177-10187, Sept. 2020, doi: 10.1109/TVT.2020.3003356.

[3] A. Mukherjee, S. Misra, V. S. P. Chandra, and N. S. Raghuwanshi, "ECoR: Energy-Aware Collaborative Routing for Task Offload in Sustainable UAV Swarms," in IEEE Transactions on Sustainable Computing, vol. 5, no. 4, pp. 514-525, 1 Oct.-Dec. 2020, doi: 10.1109/TSUSC.2020.2976453.

[4] L. Zhang, F. Hu, Z. Chu, E. Bentley, and S. Kumar, "3D Transformative Routing for UAV Swarming Networks: A Skeleton-Guided, GPS-Free Approach," in IEEE Transactions on Vehicular Technology, vol. 70, no. 4, pp. 3685-3701, April 2021, doi: 10.1109/TVT.2021.3061911.

[5] Z. Yang, H. Liu, Y. Chen, X. Zhu, Y. Ning, and W. Zhu, "UEE-RPL: A UAV-Based Energy Efficient Routing for Internet of Things," in IEEE Transactions on Green Communications and Networking, vol. 5, no. 3, pp. 1333-1344, Sept. 2021, doi: 10.1109/TGCN.2021.3085897.

[6] S. Jiang, Z. Huang, and Y. Ji, "Adaptive UAV-Assisted Geographic Routing with Q-Learning in VANET," in IEEE Communications Letters, vol. 25, no. 4, pp. 1358-1362, April 2021, doi: 10.1109/LCOMM.2020.3048250.

[7] T. Li et al., "A mean field game-theoretic cross-layer optimization for multi-hop swarm UAV communications," in Journal of Communications and Networks, vol. 24, no. 1, pp. 68-82, Feb. 2022, doi: 10.23919/JCN.2021.000035.

[8] X. Tan, Z. Zuo, S. Su, X. Guo, X. Sun and D. Jiang, "Performance Analysis of Routing Protocols for UAV Communication Networks," in IEEE Access, vol. 8, pp. 92212-92224, 2020, doi: 10.1109/ACCESS.2020.2995040.

[9] H. Fatemidokht, M. K. Rafsanjani, B. B. Gupta, and C. -H. Hsu, "Efficient and Secure Routing Protocol Based on Artificial Intelligence Algorithms with UAV-Assisted for Vehicular Ad Hoc Networks in Intelligent Transportation Systems," in IEEE Transactions on Intelligent Transportation Systems, vol. 22, no. 7, pp. 4757-4769, July 2021, doi: 10.1109/TITS.2020.3041746.

[10] O. Thakoor, J. Garg and R. Nagi, "Multiagent UAV Routing: A Game Theory Analysis with Tight Price of Anarchy Bounds," in IEEE Transactions on Automation Science and Engineering, vol. 17, no. 1, pp. 100-116, Jan. 2020, doi: 10.1109/TASE.2019.2902360.

[11] H. Song, L. Liu, S. M. Pudlewski and E. S. Bentley, "Random Network Coding Enabled Routing Protocol in Unmanned Aerial Vehicle Networks," in IEEE Transactions on Wireless Communications, vol. 19, no. 12, pp. 8382-8395, Dec. 2020, doi: 10.1109/TWC.2020.3022399.

[12] M. Y. Arafat and S. Moh, "A Q-Learning-Based Topology-Aware Routing Protocol for Flying Ad Hoc Networks," in IEEE Internet of Things Journal, vol. 9, no. 3, pp. 1985-2000, 1 Feb.1, 2022, doi: 10.1109/JIOT.2021.3089759.

[13] F. Xiong et al., "Energy-Saving Data Aggregation for Multi-UAV System," in IEEE Transactions on Vehicular Technology, vol. 69, no. 8, pp. 9002-9016, Aug. 2020, doi: 10.1109/TVT.2020.2999374.

[14] X. Qiu, L. Xu, P. Wang, Y. Yang, and Z. Liao, "A Data-Driven Packet Routing Algorithm for an Unmanned Aerial Vehicle Swarm: A Multi-Agent Reinforcement Learning Approach," in IEEE Wireless Communications Letters, vol. 11, no. 10, pp. 2160-2164, Oct. 2022, doi: 10.1109/LWC.2022.3195963.

[15] M. Zhang, C. Dong, P. Yang, T. Tao, Q. Wu, and T. Q. S. Quek, "Adaptive Routing Design for Flying Ad Hoc Networks," in IEEE Communications Letters, vol. 26, no. 6, pp. 1438-1442, June 2022, doi: 10.1109/LCOMM.2022.3152832.

[16] P. K. Deb, A. Mukherjee, and S. Misra, "XiA: Send-It-Anyway Q-Routing for 6G-Enabled UAV-LEO Communications," in IEEE Transactions on Network Science and Engineering, vol. 8, no. 4, pp. 2722-2731, 1 Oct.-Dec. 2021, doi: 10.1109/TNSE.2021.3086484.

[17] P. K. Deb, A. Mukherjee, and S. Misra, "XiA: Send-It-Anyway Q-Routing for 6G-Enabled UAV-LEO Communications," in IEEE Transactions on Network Science and Engineering, vol. 8, no. 4, pp. 2722-2731, 1 Oct.-Dec. 2021, doi: 10.1109/TNSE.2021.3086484.

[18] J. Yoon, S. Doh, O. Gnawali, and H. Lee, "Time-Dependent Ad-Hoc Routing Structure for Delivering Delay-Sensitive Data Using UAVs," in IEEE Access, vol. 8, pp. 36322-36336, 2020, doi: 10.1109/ACCESS.2020.2974553.

[19] J. Fu, A. Núñez and B. De Schutter, "Real-Time UAV Routing Strategy for Monitoring and Inspection for Postdisaster Restoration of Distribution Networks," in IEEE Transactions on Industrial Informatics, vol. 18, no. 4, pp. 2582-2592, April 2022, doi: 10.1109/TII.2021.3098506.

[20] S. G. Manyam, K. Sundar and D. W. Casbeer, "Cooperative Routing for an Air–Ground Vehicle Team—Exact Algorithm, Transformation Method, and Heuristics," in IEEE Transactions on Automation Science and Engineering, vol. 17, no. 1, pp. 537-547, Jan. 2020, doi: 10.1109/TASE.2019.2931894.

[21] S. G. Manyam, K. Sundar and D. W. Casbeer, "Cooperative Routing for an Air–Ground Vehicle Team—Exact Algorithm, Transformation Method, and Heuristics," in IEEE Transactions on Automation Science and Engineering, vol. 17, no. 1, pp. 537-547, Jan. 2020, doi: 10.1109/TASE.2019.2931894.

[22] D. Shumeye Lakew, U. Sa'ad, N. -N. Dao, W. Na, and S. Cho, "Routing in Flying Ad Hoc Networks: A Comprehensive Survey," in IEEE Communications Surveys & Tutorials, vol. 22, no. 2, pp. 1071-1120, Second quarter 2020, doi: 10.1109/COMST.2020.2982452.

[23] A. Rovira-Sugranes, F. Afghah, J. Qu and A. Razi, "Fully-Echoed Q-Routing with Simulated Annealing Inference for Flying Adhoc Networks," in IEEE Transactions on Network Science and Engineering, vol. 8, no. 3, pp. 2223-2234, 1 July-Sept. 2021, doi: 10.1109/TNSE.2021.3085514.

[24] G. Sun, D. Qin, T. Lan, and L. Ma, "Research on Clustering Routing Protocol Based on Improved PSO in FANET," in IEEE Sensors Journal,

RESEARCH ARTICLE

vol. 21, no. 23, pp. 27168-27185, 1 Dec.1, 2021, doi: 10.1109/JSEN.2021.3117496.

[25] Y. Cui, Q. Zhang, Z. Feng, Z. Wei, C. Shi, and H. Yang, "Topology-Aware Resilient Routing Protocol for FANETs: An Adaptive Q-Learning Approach," in IEEE Internet of Things Journal, vol. 9, no. 19, pp. 18632-18649, 1 Oct.1, 2022, doi: 10.1109/JIOT.2022.3162849.

[26] C. Wang et al., "Elastic Routing Mechanism for Flying Ad Hoc Network," in IEEE Access, vol. 10, pp. 98712-98723, 2022, doi: 10.1109/ACCESS.2022.3206767.

[27] O. T. Abdulhae, J. S. Mandeep and M. Islam, "Cluster-Based Routing Protocols for Flying Ad Hoc Networks (FANETs)," in IEEE Access, vol. 10, pp. 32981-33004, 2022, doi: 10.1109/ACCESS.2022.3161446.

[28] O. T. Abdulhae, J. S. Mandeep and M. Islam, "Cluster-Based Routing Protocols for Flying Ad Hoc Networks (FANETs)," in IEEE Access, vol. 10, pp. 32981-33004, 2022, doi: 10.1109/ACCESS.2022.3161446.

[29] Z. Du et al., "A Routing Protocol for UAV-Assisted Vehicular Delay Tolerant Networks," in IEEE Open Journal of the Computer Society, vol. 2, pp. 85-98, 2021, doi: 10.1109/OJCS.2021.3054759.

[30] B. Liu, W. Zhang, W. Chen, H. Huang and S. Guo, "Online Computation Offloading and Traffic Routing for UAV Swarms in Edge-Cloud Computing," in IEEE Transactions on Vehicular Technology, vol. 69, no. 8, pp. 8777-8791, Aug. 2020, doi: 10.1109/TVT.2020.2994541.

[31] J. Baek, S. I. Han, and Y. Han, "Energy-Efficient UAV Routing for Wireless Sensor Networks," in IEEE Transactions on Vehicular Technology, vol. 69, no. 2, pp. 1741-1750, Feb. 2020, doi: 10.1109/TVT.2019.2959808.

[32] Souto, Anderson, Rodrigo Alfaia, Evelin Cardoso, Jasmine Araújo, and Carlos Francês. 2023. "UAV Path Planning Optimization Strategy: Considerations of Urban Morphology, Microclimate, and Energy Efficiency Using Q-Learning Algorithm" Drones 7, no. 2: 123. https://doi.org/10.3390/drones702012

[33] Alam, Muhammad Morshed, and Sangman Moh. 2022. "Survey on Q-Learning-Based Position-Aware Routing Protocols in Flying Ad Hoc Networks" Electronics 11, no. 7: 1099. https://doi.org/10.3390/electronics11071099

[34] Balasubramanian, E., Elangovan, E., Tamilarasan, P. et al. Optimal energy efficient path planning of UAV using hybrid MACO-MEA* algorithm: theoretical and experimental approach. J Ambient Intell Human Comput (2022). https://doi.org/10.1007/s12652-022-04098-z

[35] Fotouhi A, Ding M, Hassan M. Deep Q-Learning for Two-Hop Communications of Drone Base Stations. Sensors (Basel). 2021 Mar 11;21(6):1960. doi: 10.3390/s21061960. PMID: 33799546; PMCID: PMC7999891.

[36] Mannan, A, Obaidat, MS, Mahmood, K, Ahmad, A, Ahmad, R. Classical versus reinforcement learning algorithms for unmanned aerial vehicle network communication and coverage path planning: A systematic literature review. Int J Commun Syst. 2023; 36(5):e5423. doi:10.1002/dac.5423

[37] A. Vashisth and R. S. Batth, "An Overview, Survey, and Challenges in UAVs Communication Network," 2020 International Conference on Intelligent Engineering and Management (ICIEM), London, UK, 2020, pp. 342-347, doi: 10.1109/ICIEM48762.2020.9160197.

[38] A. Vashisth, R. Singh Batth, and R. Ward, "Existing Path Planning Techniques in Unmanned Aerial Vehicles (UAVs): A Systematic Review," 2021 International Conference on Computational Intelligence and Knowledge Economy (ICCIKE), Dubai, United Arab Emirates, 2021, pp. 366-372, doi: 10.1109/ICCIKE51210.2021.9410787.

Authors

Anshu Vashisth is a part-time Ph.D. Research Scholar in the School of Computer Science and Engineering at Lovely Professional University, Punjab, India. He has completed his M. Tech in Computer Science and Engineering in 2015 from Lovely Professional University, Punjab, India. His research interests include UAV communication, wireless communication, AI, and Machine Learning.

Balraj Singh is working at Lovely Professional University, Phagwara, Punjab, India. He received his Ph.D. from NIT, Jalandhar, Punjab, India. His research areas are UAV, Cloud Computing, Mobile computing, pervasive computing, expert system, wireless communication, and machine learning.

Ranbir Singh Batth is working at Menzies Institute of Technology, Melbourne Australia. He received his Ph.D. from IKG Punjab Technical University, Kapurthala, Punjab, India in 2018 and his Master's degree in Computer Engineering from Punjabi University, Patiala. His research interests include Wireless Communication, Cloud Computing, Vehicular Ad Hoc Networks, the Internet of Things, Machine Learning, Deep Learning, pervasive computing, Intelligent Transportation Systems, and Mobile computing. He also serves as an editorial member, guest editor

How to cite this article:

Anshu Vashisth, Balraj Singh, Ranbir Singh Batth, “QMRNB: Design of an Efficient Q-Learning Model to Improve Routing Efficiency of UAV Networks via Bioinspired Optimizations”, International Journal of Computer Networks and Applications (IJCNA), 10(2), PP: 256-264, 2023, DOI: 10.22247/ijcna/2023/220740.